CCD ARRAYS CAMERAS and DISPLAYS

second edition

Gerald C. Holst

CCD ARRAYS, CAMERAS, and DISPLAYS

Second edition

Gerald C. Holst

Copublished by

JCD Publishing
2932 Cove Trail
Winter Park, FL 32789

and

SPIE Optical Engineering Press
A Publication of SPIE—The International Society for Optical Engineering
Bellingham, Washington USA

Library of Congress Cataloging-in-Publication Data

Holst, Gerald C.
CCD arrays, cameras, and displays/ Gerald C. Holst. -- 2nd ed.
p. cm.
Includes bibliographical references (p.) and index.
ISBN 0-9640000-4-0 (hardcover)
1. Charge coupled devices. 2. Information display systems.
I. Title
TK7871.99.C45H65 1998
621.36'7--dc21 98-10770
CIP

Copublished by

JCD Publishing
2932 Cove Trail
Winter Park, FL 32789
Phone: 407/629-5370
Fax: 407/629-5370
ISBN: 0-9640000-4-0

SPIE - The International Society for Optical Engineering
P.O. Box 10
Bellingham, WA 98227-0010
Phone: 360/676-3290
Fax: 360/647-1445
WWW: http://www.spie.org
ISBN: 0-8194-2853-1

The second edition is dedicated to

Emily, Irma, and Deanne

PREFACE to the SECOND EDITION

The first edition was published in January 1996. The response was overwhelming and it sold out in two years. Rather than reprint that edition, it was important to update the information. After all, technology moves rapidly. The first edition also had some typographic errors. I thank those who were gracious enough to point them out to me. Those errors have been corrected.

One major change is that "CCD array" was changed to "solid state array" throughout the text. This change allows flexibility in sensor type (e.g., CCD, CID, and CMOS). As new sensors enter the market place, the appropriate chapters in this book are still applicable. Camera operation is nearly independent of the specific sensor. The book title remains the same because of the preponderance of CCD arrays. Chapters 2, 5, 7, 11, and 12 are essentially the same. The reference lists have been updated and they now include books and journal articles published in 1996 and 1997.

Calling a sample a pixel does not seem bad. But "pixel" is used to denote a detector, an element of a computer stored data array, and the primary element of a monitor. Unfortunately, there is no *a priori* relationship between the various device pixels. As indicated in Chapter 1, the various samples are called pixels, datels, and disels. This clarifies the relationship between what the detector array produces and what is presented on the monitor. The general assumption is that there is a one-to-one relationship between these elements - but there is no requirement that this be so.

Chapter 3 was renamed *Solid State Arrays*. Both CID and CMOS detector technologies have been added. Photodiodes require transfer gates to move the stored charge into the CCD transfer register. Although these gates were not shown in the first edition, their inclusion provides further understanding of how the detectors operate. The antibloom drain section was expanded and its action as an electronic shutter is described. Dark current reduction, so important for low illumination imaging, is explained in more detail.

In Chapter 4, additional factors that affect the spectral responsivity are given. Although dynamic range would seem to be a straight forward concept, it has numerous definitions. Manufacturers specify minimum and maximum illumination levels as either scene illumination or faceplate illumination (Chapter

6). Since the manufacturer has no control over the user's selection of lenses, the faceplate illumination is more popular. Minimum illumination may create a full video signal or some lower value (e.g., 30, 50, or 80% video). Numerical examples illustrate the differences. The signal-to-noise section has been expanded.

An electronic imaging system has two major components: that which responds to scene variations (detector, optics, and the observer) and that which responds to time varying signals (electronics and display). To better describe the two concepts, "MTF theory" was renamed "linear system theory" in Chapters 8 and 9. This makes the approach more general and the nomenclature associated with Fourier transforms has been clarified.

In keeping with the various diverse components of a solid state camera, MTF analysis is possible in four different spatial frequency domains. These domains are related by simple equations (Chapter 10). A primary consideration in system design is the relationship between the optical resolution and the detector response. A camera may be either detector-limited or optics-limited. The transition from one to the other has a profound effect on imagery. More detail can be found in *Sampling, Aliasing, and Data Fidelity*, G. C. Holst, JCD Publishing (1998).

February 1998 Gerald C. Holst

PREFACE to the FIRST EDITION

Charge-coupled devices (CCDs) were invented by Boyle and Smith in 1970. CCDs have matured to the point where they are widely affordable and have spawned an explosion of applications. Some applications include high definition television for program production, consumer camcorders, electronic still cameras, industrial machine vision cameras, and optical character readers including bar code scanners and fax machines. Applications are only limited by man's imagination.

Linear and time-delay-and-integrate (TDI) sensors are appropriate for applications where the objects are continuously moving as in the lumber, paper, textile, steel, and aluminum forming industries. These sensors provide non-contact measurement during fabrication as well as information that allows analysis of surface conditions.

The CCD architecture has three basic functions: (a) charge collection, (b) charge transfer, and (c) the conversion of charge into a measurable voltage. Since most sensors operating in the visible region use CCD architecture to read the signal, they are popularly called CCD cameras. For these devices, charge generation is often considered as the initial function of the CCD.

Although silicon CCDs respond from 10^{-4} μm to 1.1 μm, only imaging applications in the visible and near infrared are considered here. The long wavelength response may be limited by the semiconductor band gap (equivalent to 1.1 μm for silicon), diffusion length (photoelectrons recombine with holes before reaching a charge well), or infrared blocking filter (limits the response to the visible region). The short wavelength response may be limited by the polysilicon electrode transmittance (opaque below 0.4 μm) or a glass protective cover (opaque below about 0.3 μm).

This book is not a design book. It highlights those features and specifications that are most often reported in data sheets. It provides the necessary information to compare and analyze camera systems. The material in this book is representative of solid state camera operation and architecture. A specific device may not have all the features listed or may have additional features not listed.

Arrays and cameras may be specified by radiometric or photometric units (Chapter 2). Emphasis is placed on the relationship between the spectral content of the lighting compared to the sensor's spectral response. CCD array operation is described in Chapter 3 and array specifications (performance parameters) are described in Chapter 4. Array specifications, capabilities, and limitations affect

the overall camera specifications. The camera manufacturer cannot increase the inherent signal-to-noise ratio of the array. He strives to maintain the array specifications without introducing additional noise or adversely affecting image quality. In Chapters 5 and 6, camera operation and specifications are related to the array values.

A camera is of no use by itself. Its value is only known when an image is evaluated. Therefore, the observer is a critical component of the imaging system. While the camera manufacturer strives to provide the highest possible resolution, he has no control over the display used. It is up to the end-user to select an adequate display. There is no advantage to using a high quality camera if the display cannot faithfully reproduce the image. Flat panel displays are an emerging technology but cathode ray tubes (CRT) will probably dominate the display market due to low cost, high resolution, wide color gamut, and long life. Chapter 7 discusses CRT-based display performance metrics.

Staring arrays are inherently undersampled. The highest spatial frequency that can be faithfully reproduced is the Nyquist frequency (Chapter 8). Many performance parameters are evaluated at the Nyquist frequency. The selection of Nyquist frequency is for convenience. Spatial frequencies above Nyquist frequency do not disappear. They are just distorted. Spatial sampling creates ambiguity in target edge location and produces moiré patterns when periodic targets are viewed.

Most of the chapters discuss imaging system phenomenology with minimal mathematics. Many examples are included to illustrate basic principles and back-of-the-envelope calculations. Detailed math is found in Chapters 9 and 10 (*MTF Theory* and *System MTF*).

Many image quality metrics are linked to the system's modulation transfer function (MTF) and therefore are listed in Chapter 11. Image quality, as specified in this book, does not include data compression effects, image processing, or machine vision algorithms since these are nonlinear operators. These processes can only be evaluated on a case-by-cases basis. As digital image processing capability is incorporated into cameras, they may be called sensors with built-in computers.

The military is interested in detecting, recognizing, and identifying targets at long distances, and the minimum resolvable contrast (MRC) is an appropriate figure of merit (Chapter 12). The MRC is a system performance metric that includes both resolution (via the MTF) and sensitivity measures. It is a measure of an observer's ability to detect a bar target embedded in noise.

Although this book is titled *CCD Arrays, Cameras, and Displays*, nearly all of the concepts apply to every staring array including charge injection devices (CIDs) and active pixels sensors (APS). The electronics cannot "know" what created the signal nor the spectral content associated with the signal. Similarly, without prior training, the observer will not know what sensor was used to collect the imagery. The MTFs are generic.

Emphasis is placed upon monochrome systems. Color camera design is quite complex and represents a separate book. As such, the topics on color systems are cursory. Chapter 3 introduces basic color array operation and Chapter 5 provides some aspects of color cameras. The literature is rich with visual data on which color camera operation is based. See, for example, J. Peddie, *High-Resolution Graphics Display Systems*, Chapter 2, Windcrest/McGraw-Hill, New York, NY (1993) or A. R. Robertson and J. F. Fisher "Color Vision, Representation, and Reproduction," in *Television Engineering Handbook*, K. B. Benson, ed., Chapter 2, McGraw Hill, New York, NY (1985).

This book is for the first time CCD buyer. It explains the basic operating concepts and what the specifications represent. The book is also for the experienced buyer for it provides the background material necessary to specify cameras. The electro-optical system analyst will use the MTF and MRC theory to describe performance. For the system integrator, it provides information on the entire system: scene-to-observer. For the manager, purchasing agent, and salesperson, it is a handy reference.

The author extends his deepest gratitude to all his coworkers and students who have contributed to the ideas in this book. They are too many to mention by name. The author especially thanks all those who read draft copies of the manuscript: Constantine Anagnostopoulos, Eastman Kodak; Peter Barten, Barten Consultancy; Mike Ensminger, Ball Aerospace; Chris Fritz, Texas Instruments; Paul Gallagher, EG&G Reticon; Ron Hamilton, Cohu; Richard Hofmeister, Pulnix; Herb Huey, Northrop; Lindon Lewis, Ball Aerospace; Terry Lomheim, The Aerospace Corporation; Hugh Masterson, Mitre Corporation; Ishai Nir, Princeton Instruments; Harold Orlando, Northrop; Mike Pappas, Northrop; Kenneth Parulski, Eastman Kodak; Mike Roggemann, Air Force Institute of Technology; Luke Scott, NV&ESD; Mark Sartor, Xybion; and Mark Stafford, Dalsa. Although these reviewers provided valuable comments, the accuracy of the text is the sole responsibility of the author. Jeff Odhner and Steve Kraemer provided the graphic arts.

February 1996 Gerald C. Holst

CCD ARRAYS, CAMERAS, and DISPLAYS

Second edition

TABLE of CONTENTS

SYMBOL LIST

SYMBOL	DEFINITION	UNITS
α_{ABS}	spectrally averaged absorption coefficient	1/m
$\alpha_{ABS}(\lambda)$	spectral absorption coefficient	1/m
α_G	dark current factor	numeric
α_{ERROR}	target displacement in TDI systems	mm
α_m	display aspect ratio	numeric
β_1	eye filter 1	1/electrons^2
β_m	eye filter m	1/electrons^2
γ	gamma	numeric
ΔV	TDI velocity error	mm/s
$\Delta\lambda$	wavelength interval	μm
$\Delta\rho$	target-background reflectance difference	numeric
Δf_e	noise equivalent bandwidth	Hz
$\Delta n_{MINIMUM}$	minimum number of photoelectrons	numeric
Δn_{pe}	target-background photoelectron difference	numeric
Δt	time separation between two pulses	s
ϵ	charge transfer efficiency	numeric
η	quantum efficiency	numeric
$\eta_{CATHODE}$	intensifier photocathode quantum efficiency	numeric
η_{CCD}	CCD quantum efficiency	numeric
$\eta_{PHOTOCATHODE}$	intensifier photocathode quantum efficiency	numeric
η_{SCREEN}	intensifier screen quantum efficiency	numeric
λ	wavelength	μm
λ_{AVE}	average wavelength	μm
λ_{MAX}	maximum wavelength	μm
λ_{MIN}	minimum wavelength	μm
λ_O	selected wavelength	μm
λ_p	wavelength of peak response	μm
ρ	spectrally averaged reflectance	numeric
$\rho(\lambda)$	spectral reflectance	numeric
ρ_B	background reflectance	numeric
ρ_T	target reflectance	numeric
σ_{ATM}	atmospheric absorption	1/m
σ_H	component of three-dimensional noise model	rms
σ_{HV}	component of three-dimensional noise model	rms
σ_{IIT}	image intensifier tube 1/e spot size	mm

σ_{IS}	display 1/e spot size in image space	mm
σ_R	rms value of random motion	rms mm
σ_{SPOT}	display spot 1/e intensity size	mm
σ_T	component of three-dimensional noise model	rms
σ_{TH}	component of three-dimensional noise model	rms
σ_{TVH}	component of three-dimensional noise model	rms
σ_V	component of three-dimensional noise model	rms
$\Phi_{CATHODE}$	flux incident onto to intensifier photocathode	W
Φ_V	luminous flux	lumens
a_L	distance target has moved on detector	mm
$A_{CATHODE}$	area of photocathode	m^2
A_D	detector area	m^2
A_I	area of source in image plane	m^2
A_O	lens area	m^2
A_{PIXEL}	projected CCD pixel on intensifier photocathode	m^2
A_S	source area	m^2
c	speed of light, $c = 3 \times 10^8$	m/s
c_1	first radiation constant $c_1 = 3.7418 \times 10^8$	W-$\mu m^4/m^2$
c_2	second radiation constant $c_2 = 1.4388 \times 10^4$	μm-K
c_3	third radiation constant $c_3 = 1.88365 \times 10^{27}$	photons-μm^3/s-m^2
C	sense node capacitance	farad
C_O	target's inherent contrast	numeric
C_R	target's contrast at entrance aperture	numeric
CTE	charge transfer efficiency	numeric
d	detector width	mm
d_{CC}	detector pitch	mm
d_{CCH}	horizontal detector pitch	mm
d_{CCV}	vertical detector pitch	mm
d_H	horizontal detector size	mm
d_O	selected target dimension	mm
d_T	target detail	mm
d_V	vertical detector size	mm
D	observer to display distance	m
D_O	optical diameter	mm
DAS_H	horizontal detector angular subtense	mrad
DAS_V	vertical detector angular subtense	mrad
E_G	detector band gap	eV
E_T	impurity band gap	eV
f_{2D}	two-dimensional spatial frequency	cycles/mrad
f_{BOOST}	boost frequency	Hz
f_c	cutoff frequency of an ideal circuit	Hz
f_{CLOCK}	pixel clock rate	Hz
f_{DC}	detector cutoff frequency	cycles/mm
f_{e3dB}	one-half power frequency	Hz
f_e	electrical frequency	Hz

f_{EYE}	spatial frequency at eye	cycles/deg
f_i	image spatial frequency	cycles/mm
f_N	Nyquist frequency	cycles/mm
f_o	selected spatial frequency	cycles/mm
f_{OC}	optics cutoff frequency	cycles/mm
f_{PEAK}	peak frequency of eye MTF	cycles/deg
f_{RASTER}	raster frequency	cycles/mm
f_S	sampling frequency	cycles/mm
f_v	electrical frequency in the video domain	Hz
f_{v3dB}	one-half power frequency (video domain)	Hz
f_{vN}	Nyquist frequency in the video domain	Hz
f_{vS}	sampling frequency in the video domain	Hz
f_x	horizontal spatial frequency	cycles/mm
f_y	vertical spatial frequency	cycles/mm
fl	focal length	m
F	f-number	numeric
F_{MAX}	maximum frame rate	Hz
F_R	frame rate	Hz
G	on-chip amplifier gain	numeric
G_1	off-chip amplifier gain	numeric
G_{MCP}	microchannel gain	numeric
h	Planck's constant, $h = 6.626 \times 10^{-34}$	J-s
h_c	target critical dimension	m
H	target height	m
$H_{MONITOR}$	monitor height	m
HFOV	horizontal field-of-view	mrad
J_D	dark current density	A/cm^2
k	Boltzmann's constant, $k = 1.38 \times 10^{-23}$	J/K
k_1	a constant	numeric
$K_H(f)$	horizontal summary noise factor	
K_M	luminous efficacy, $K_M = 683$ (photopic)	lumens/W
k_{MCP}	microchannel excess noise	numeric
$K_V(f)$	vertical summary noise factor	
L_B	background luminance	$lumen/m^2\text{-}sr$
L_D	Depletion length	μm
L_{DIFF}	Diffusion length	μm
L_e	spectral radiant sterance	$W/(m^2\text{-}\mu m\text{-}sr)$
L_q	spectral photon sterance	$photons/(s\text{-}m^2\text{-}\mu m\text{-}sr)$
L_T	target luminance	$lumen/m^2\text{-}sr$
m	index	numeric
M_e	spectral radiant exitance	$W/(m^2\text{-}\mu m)$
M_{OPTICS}	optical magnification	numeric
$M_P(\lambda)$	spectral power	$W/\mu m$
M_q	spectral photon exitance	$photons/(s\text{-}m^2\text{-}\mu m)$

n	index	numeric
$n_{CATHODE}$	photons incident onto intensifier photocathode	numeric
n_{DARK}	number of dark electrons	numeric
$n_{DETECTOR}$	number of photons incident onto detector	numeric
n_e	number of electrons	numeric
n_{EYE}	number of photons incident onto eye	numeric
n_{IMAGE}	number of photons incident onto image plane	numeric
n_{LENS}	number of photons incident onto lens	numeric
n_{MCP}	number of photons incident on microchannel plate	numeric
n_{pe}	number of photoelectrons	numeric
$n_{pe\text{-}B}$	background photoelectrons	numeric
$n_{pe\text{-}T}$	target photoelectrons	numeric
$n_{PHOTON\text{-}CCD}$	number on photons incident onto CCD	numeric
n_{READ}	pixels between active array and sense node	numeric
n_{SCREEN}	photons incident onto intensifier screen	numeric
n_{WINDOW}	number of photons incident onto intensifier	numeric
N	index	numeric
N_{50}	Johnson's 50% probability	cycles
$N_{DETECTORS}$	total number of detectors	numeric
N_H	number of horizontal detectors	numeric
N_{LINE}	number of raster lines	numeric
N_T	equivalent number of cycles on target	cycles
N_{TDI}	number of TDI elements	numeric
N_{TRANS}	number of charge transfers	numeric
N_{TV}	display resolution	TVL/PH
N_V	number of vertical detectors	numeric
N_{WELL}	charge well capacity	numeric
P	display line spacing or pixel spacing	mm/line
PAS_H	horizontal pixel angular subtense	mrad
PAS_V	vertical pixel angular subtense	mrad
q	electron charge, $q = 1.6 \times 10^{-19}$	coul
R	range to target	m
R_1	distance from lens to source	m
R_2	distance from lens to detector	m
R_{AVE}	spectral averaged responsivity	$V/(\mu J\text{-}cm^{-2})$ or $DN/(\mu J\text{-}cm^{-2})$
R_e	average spectral responsivity	A/W
$R_e(\lambda)$	spectral responsivity	A/W
R_P	peak responsivity	A/W
$R_{PHOTOMETRIC}$	responsivity	V/lux
R_q	spectrally averaged quantum efficiency	numeric
$R_q(\lambda)$	spectral quantum efficiency	numeric
R_R	raster resolution	lines/mm
R_{TVL}	horizontal display resolution	TVL/PH
R_V	responsivity	A/lumen
$R_{VERTICAL}$	vertical display resolution	lines
S	display spot size (FWHM intensity)	mm

t_{ARRAY}	time to clock out the full array	s
t_{CLOCK}	time between pixels	s
t_e	eye integration time	s
$t_{H\text{-}LINE}$	time to read one pixel line	s
t_{INT}	integration time	s
t_{LINE}	video active line time	s
T	absolute temperature	Kelvin
T/#	T-number	numeric
T_{ASPECT}	target aspect ratio	numeric
T_{ATM}	spectrally averaged atmospheric transmittance	numeric
$T_{ATM}(\lambda)$	spectral atmospheric transmittance	numeric
T_{fo}	fiber optic bundle transmittance	numeric
T_{illum}	absolute temperature of illuminating source	K
$T_{IR\text{-}FILTER}$	spectrally averaged transmittance of IR filter	numeric
T_{OPTICS}	spectrally averaged optical transmittance	numeric
$T_{OPTICS}(\lambda)$	spectral optical transmittance	numeric
$T_{RELAY\ LENS}$	relay lens transmittance	numeric
T_{WINDOW}	transmittance of intensifier window	numeric
u_d	horizontal spatial frequency (display space)	cycles/mm
u_i	horizontal spatial frequency (image space)	cycles/mm
u_{iC}	horizontal optical cutoff (image space)	cycles/mm
u_{iD}	horizontal detector cutoff (image space)	cycles/mm
u_{iN}	horizontal Nyquist frequency (image space)	cycles/mm
u_{ob}	horizontal spatial frequency (object space)	cycles/mm
U	photoresponse nonuniformity	numeric
v_d	vertical spatial frequency (display space)	cycles/mm
v_i	vertical spatial frequency (image space)	cycles/mm
v_{iD}	vertical detector cutoff (image space)	cycles/mm
v_{iN}	vertical Nyquist frequency (image space)	cycles/mm
$V(\lambda)$	photopic or scotopic eye response	numeric
V_{CAMERA}	camera output voltage	V
V_{GRID}	voltage on CRT grid	V
V_{LSB}	voltage of the least significant bit	V
V_{MAX}	maximum signal	numeric
V_{MIN}	minimum signal	numeric
V_{NOISE}	noise voltage after on-chip amplifier	V rms
V_{OUT}	voltage after on-chip amplifier	V
V_R	viewing ratio	numeric
V_{RESET}	reset voltage after on-chip amplifier	V
V_{SCENE}	video voltage before gamma corrector	V
V_{SIGNAL}	signal voltage after on-chip amplifier	V
V_{VIDEO}	video voltage after gamma corrector	V
VFOV	vertical field-of-view	mrad
W	target width	m
$W_{MONITOR}$	monitor width	m
x	horizontal distance	m
y	vertical distance	m

$\langle n_1 \rangle$	noise source 1	rms electrons
$\langle n_{ADC} \rangle$	quantization noise	rms electrons
$\langle n_{CCD\text{-}DARK} \rangle$	dark current CCD noise in an ICCD	rms electrons
$\langle n_{CCD\text{-}PHOTON} \rangle$	noise before CCD in an ICCD	rms electrons
$\langle n_{CCD} \rangle$	CCD noise in an ICCD	electrons rms
$\langle n_{DARK} \rangle$	dark current shot noise	rms electrons
$\langle n_{FLOOR} \rangle$	noise floor	rms electrons
$\langle n_{FPN} \rangle$	fixed pattern noise	rms electrons
$\langle n_{MCP} \rangle$	microchannel noise	rms electrons
$\langle n_m \rangle$	noise source m	rms electrons
$\langle n_{OFF\text{-}CHIP} \rangle$	off-chip amplifier noise	rms electrons
$\langle n_{ON\text{-}CHIP} \rangle$	on-chip amplifier noise	rms electrons
$\langle n_{PATTERN} \rangle$	pattern noise	rms electrons
$\langle n_{PC\text{-}DARK} \rangle$	intensifier dark current shot noise	rms electrons
$\langle n_{PC\text{-}SHOT} \rangle$	intensifier photon shot noise	rms electrons
$\langle n_{pe} \rangle$	photon shot noise	rms electrons
$\langle n_{PRNU} \rangle$	photoresponse nonuniformity noise	rms electrons
$\langle n_{RESET} \rangle$	reset noise	rms electrons
$\langle n_{SCREEN} \rangle$	intensifier screen noise	rms electrons
$\langle n_{SHOT} \rangle$	shot noise	rms electrons
$\langle n_{SYS} \rangle$	system noise	rms electrons
$\langle V_N \rangle$	noise voltage	rms V

1

INTRODUCTION

Charge-coupled devices (CCDs) were invented by Boyle and Smith[1,2] in 1970. Since then, considerable literature[3-7] has been written on CCD physics, fabrication, and operation. However, the array does not create an image by itself. It requires an optical system to image the scene onto the array's photosensitive area. The array requires a bias and clock signals. Its output is a series of analog pulses that represent the scene intensity at a series of discrete locations.

Devices may be described functionally according to their architecture (frame transfer, interline transfer, etc.) or by application. Certain architectures lend themselves to specific applications. For example, astronomical cameras typically use full frame arrays whereas consumer video systems use interline transfer devices.

The heart of the solid state camera is the solid state array. It provides the conversion of light intensity into measurable voltage signals. With appropriate timing signals, the temporal voltage signal represents spatial light intensities. When the array output is amplified and formatted into a standard video format, a solid state camera is created. Because CCDs were the first solid state detectors, cameras are popularly called CCD cameras even though they may contain charge injection devices (CIDs) or complementary metal-oxide-semiconductors (CMOS) as detectors. These are solid state cameras.

The array specifications, while the first place to start an analysis, are only part of the overall system performance. The system image quality depends on all the components. Array specifications, capabilities, and limitations are the basis for the camera specifications. Camera manufacturers cannot change these. A well-designed camera will not introduce additional noise nor adversely affect image quality provided by the array.

A camera is of no use by itself. Its value is only known when an image is evaluated. The camera output may be directly displayed on a monitor, stored on video tape or disk for later viewing, or processed by a computer. The computer may be part of a machine vision system, be used to enhance the imagery, or be used to create hard copies of the imagery. If interpretation of image quality is performed by an observer, then the observer becomes a critical component of

the imaging system. Consideration of human visual system attributes should probably be the starting point of the camera design. But some machine vision systems may not create a "user friendly" image. These systems are designed from a traditional approach: resolution, high signal-to-noise ratio, and ease of operation.

Effective camera design requires an orderly integration of diverse technologies and languages associated with radiation physics, optics, solid state sensors, electronic circuitry, and image processing algorithms. Each field is complex and is a separate discipline.

Electro-optical imaging system analysis is a mathematical construct that provides an optimum design through appropriate tradeoff analyses. A comprehensive model includes the target, background, the properties of the intervening atmosphere, the optical system, detector, electronics, display, and the human interpretation of the displayed information. While any of these components can be studied in detail separately, the electro-optical imaging system cannot. Only complete end-to-end analysis (scene-to-observer interpretation) permits system optimization.

1.1. SOLID STATE DETECTORS

CCD refers to a semiconductor architecture in which charge is read out of storage areas. The CCD architecture has three basic functions: (a) charge collection, (b) charge transfer, and (c) the conversion of charge into a measurable voltage. The basic building block of the CCD is the metal-oxide-semiconductor (MOS) capacitor. The capacitor is called a gate. By manipulating the gate voltages, charge can be either stored or transferred. Charge generation in most devices occurs under a MOS capacitor (also called a photogate). For some devices (notably interline transfer devices) photodiodes create the charge. After charge generation, the transfer occurs in the MOS capacitors for all devices.

Because most sensors operating in the visible region use a CCD type architecture to read the signal, they are popularly called CCD cameras. For these devices, charge generation is often considered as the initial function of the CCD. More explicitly, these cameras should be called *solid state cameras with a CCD readout.*

CCDs and detectors can be integrated either monolithically or as hybrids. Monolithic arrays combine the detector and CCD structure on a single chip. The most common detectors are sensitive in the visible region of the spectrum. They use silicon photogates or photodiodes and are monolithic devices. CCDs have been successfully used for infrared detectors such as Schottky barrier devices that are sensitive to 1.2 μm to 5 μm radiation.

Hybrid arrays avoid some pitfalls associated with growing different materials on a single chip and provide a convenient bridge between well-developed but otherwise incompatible technologies. HgCdTe (sensitive to 8 to 12 μm radiation) is bump bonded to a CCD readout using indium as the contact and, as such, is a hybrid array.

With charge-injection devices (CID), the pixels consist of two MOS capacitors whose gates are separately connected to rows and columns. Usually the column capacitors are used to integrate charge while the row capacitors sense the charge after integration. With the CID architecture, each pixel is addressable, i.e., it is a matrix-addressable device.

CID readout is accomplished by transferring the integrated charge from the column capacitors to the row capacitors. After this nondestructive signal readout, the charge moves back to the columns for more integration or is injected (discarded) back into the silicon substrate. By suspending charge injection, the user initiates "multiple frame integration" (time lapse exposure) and can view the image on a display as the optimum exposure develops. Integration may proceed for a few milliseconds up to several hours. With individual capacitors on each sensing pixel, blooming is not transported so overloads cannot propagate.

Recently, active pixels have been introduced. These devices are fabricated with complementary-metal-oxide-semiconductor (CMOS) technology. The advantage is that one or more active transistors can be integrated into the pixel. As such, they become fully addressable (can read selected pixels) and can perform on-chip image processing.

1.2. IMAGING SYSTEM APPLICATIONS

There are four broad applications: general imagery (which includes both professional television broadcast and consumer camcorder systems), machine vision, scientific, and military. Trying to appeal to all four applications, manufacturers use words such as low noise, high frame rate, high resolution,

reduced aliasing, and high sensitivity. These words are simply adjectives with no specific meaning. They only become meaningful when compared (i.e., camera A has low noise compared to camera B).

Table 1-1 lists several design categories. While the requirements vary by category, a camera may be used for multiple applications. For example, a consumer video camera often is adequate for many scientific experiments. A specific device may not have all the features listed or may have additional features not listed. The separation between professional broadcast, consumer video, machine vision, scientific, and military devices becomes fuzzy as technology advances.

Color cameras are used for professional television, camcorder, and film replacement systems. With machine vision systems, color cameras are used to verify the color consistency of objects such as printed labels or paint mixture colors. While color may not be the primary concern, it may be necessary for image analysis when color is the only information to distinguish boundaries.

While consumers demand color camera systems, this is not true for other applications. Depending on the application, monochrome (black and white) cameras may be adequate. A monochrome camera has higher sensitivity and therefore is the camera of choice in low light level conditions.

Image enhancement applies to those systems that have a human observer and helps the observer extract data. Some images belong to a small precious data set (e.g., remote sensing imagery from interplanetary probes). The images must be processed repeatedly to extract every piece of information. Some images are part of a data stream that are examined once (e.g., real-time video) and others have become popular and are used routinely as standards. These include the three-bar or four-bar test patterns, Lena, and the African mandrill.

The camera cannot perfectly reproduce the scene. The array spatially samples the image, and noise is injected by the array electronics. Spatial sampling creates ambiguity in target edge location and produces moiré patterns when viewing periodic targets. While this is a concern to scientific and military applications, it typically is of little consequence to the average professional television broadcast and consumer markets.

Table 1-1
DESIGN GOALS

DESIGN CATEGORY	GENERAL IMAGERY	MACHINE VISION	SCIENTIFIC	MILITARY
Image processing algorithms	Gamma correction	Application specific	Menu-driven multiple options	Application specific
Image processing time	Real time	Application specific with emphasis on high speed operation	Real time not usually required	Real time
Resolution	Matched to video format (e.g., EIA 170)	For a fixed field-of-view, increased resolution is desired	High resolution	High resolution
Dynamic range	8 bits/color	8 bits/color	Up to 16 bits	10 or 12 bits
Sensitivity (high signal-to-noise ratio)	High contrast targets (noise not necessarily a dominant design factor)	Application specific - not necessarily an issue because lighting can be controlled	Low noise operation	Low noise operation

Table 1-1 listed several design goals by application category. To some extent these goals are incompatible, thereby dictating design compromises. The demand for machine vision systems is increasing dramatically. Smaller target detail can be discerned with magnifying optics. However, this reduces the field-of-view. For a fixed field-of-view, higher resolution (more pixels) cameras are required to discern finer detail. When selecting an imaging system, the environment, camera, data storage, and final image format must be considered (Table 1-2).

Table 1-2
SYSTEM DESIGN CONSIDERATIONS

ENVIRONMENT	CAMERA	TRANSMISSION and STORAGE	DISPLAY
Target size Target reflectance Distance to target Atmospheric transmittance Lighting conditions	Noise Frame rate Sensitivity Detector size Array format Dynamic range Color capability	Data rate Type of storage Storage capacity Video compression	Hard copy Soft copy Resolution

Finally, the most convincing evidence of system performance is a photograph of an image. Every time an image is transferred to another device, that device's tonal transfer and modulation transfer function will affect the displayed image. Video recorders degrade resolution. Imagery may look slightly different on different displays. Similarly, hard copies produced by different printers may appear different, and there is no guarantee that the hard copy will look the same as the soft copy in every respect. Hard copies should only be considered as representative of system capability. Even with its over 100-year history, wet film developing and printing must be controlled with extreme care to create "identical" prints.

1.2.1. GENERAL IMAGERY

Cameras for the professional broadcast television and consumer camcorder markets are designed to operate in real time with an output that is consistent with a standard broadcast format. The resolution, in terms of array size, is matched to the bandwidth recommended in the standard. An array that provides an output of 768 horizontal by 484 vertical pixels creates a satisfactory image for conventional television. Eight bits (256 intensity levels or gray levels) provide an acceptable image in the broadcast and camcorder industry. The largest consumer market presently is the camcorder market.

Because solid state cameras have largely replaced image vacuum tubes, the terminology associated with these tubes is also used with solid state cameras. For example, compared to image vacuum tubes, solid state cameras have no image burn-in, no residual imaging, and usually are not affected by microphonics.

The current accepted meaning of high definition television (HDTV) is a television system providing approximately twice the horizontal and vertical resolution of present NTSC, PAL, and SECAM systems (discussed in Section 5.2.1., *Video Timing*). HDTV is envisioned as a system that provides high resolution on large displays and projection screens. The goal is to have world-wide compatibility so that HDTV receivers can display NTSC, PAL, and SECAM transmitted imagery. With today's multimedia approach, any new standard must be compatible with a variety of imaging systems ranging from 35-mm film to the various motion picture formats.

The Federal Communication Commission (FCC) adopted a HDTV transmission standard in December 1996 and broadcasting should begin between 1998 and 2000. HDTV camera descriptions and receivers will appear in many journals over the next few years.

1.2.2. MACHINE VISION

In its simplest version, a machine vision system consists of a light source, camera, and computer software that rapidly analyzes digitized images with respect to location, size, flaws, and other preprogrammed data. Unlike other types of image analysis, a machine vision system also includes a mechanism that immediately reacts to images that do not conform to the parameters stored in the computer. For example, defective parts are taken off a production line conveyor belt.

Machine vision functions include location, inspection, gauging, counting, identification, recognition, and motion tracking. These systems do not necessarily need to operate at a standard frame rate. For industrial inspection, linear arrays operating in the time delay and integration (TDI) mode can be used to measure objects moving at a high speed on a conveyor belt.

While a multitude of cues are used for target detection, recognition, or identification, machine vision systems cannot replace the human eye. The eye processes intensity differences over 11 orders of magnitude, color differences, and textual cues. The solid state camera, on the other hand, can process limited

data much faster than the human. Many operations can be performed faster, cheaper, and more accurately by machines than by humans. Machine vision systems can operate 24 hours a day without fatigue. They operate consistently whereas variability exists among human inspectors. Furthermore, these cameras can operate in harsh environments that may be unfriendly to humans (e.g., extreme heat, cold, or ionizing radiation).

1.2.3. SCIENTIFIC APPLICATIONS

For scientific applications, low noise, high responsivity, large dynamic range, and high resolution are dominant considerations. To exploit a large dynamic range, scientific cameras may digitize the signal into 12, 14, or 16 bits. Array linearity and analog-to-digital converter linearity are important. Resolution is specified by the number of detector elements, and scientific arrays may have 5000×5000 detector elements.[8] Theoretically, the array can be any size but manufacturing considerations may ultimately limit the array size.

Low noise means low-dark current and low readout noise. The dark current can be minimized by cooling the CCD. Long integration times can increase the signal value so that the readout noise is small compared to the photon shot noise.

Although low light level cameras have many applications, they tend to be used for scientific applications. There is no industry-wide definition of a "low light level" imaging system. To some, it is simply a solid state camera that can provide a usable image when the lighting conditions are less than 1 lux. To others, it refers to an intensified camera and is sometimes called a low light level television (LLLTV) system. An image intensifier amplifies a low light level image so that it can be sensed by a solid state camera. The image-intensifier/CCD camera combination is called an intensified CCD or ICCD. The image intensifier provides tremendous amplification but also introduces additional noise.

1.2.4. MILITARY APPLICATIONS

The military is interested in detecting, recognizing, and identifying targets at long distances. This requires high resolution, low noise sensors. Target detection is a perceptible act. A human determines if the target is present. The military uses the minimum resolvable contrast (MRC) as a figure of merit.

Solid state cameras are popular because of their ruggedness and small size. They can easily be mounted on remotely piloted vehicles. They are replacing wet-film systems used for mapping and photo interpretation.

1.3. CONFIGURATIONS

Imaging systems for the four broad application categories may operate in a variety of configurations. The precise setup depends on the specific requirements. Figure 1-1 is representative of a closed circuit television system where the camera output is continuously displayed. The overall image quality is determined by the camera capability (which is based on the array specifications), the bandwidth of the video format used (e.g., EIA 170, NTSC, PAL, or SECAM), the display performance, and the observer. EIA 170 was formerly called RS 170 and the NTSC standard is also called RS 170A.

Figure 1-2 illustrates a generic transmission system. The transmitter and receiver must have sufficient electronic bandwidth to provide the desired image quality. For remote sensing, the data may be compressed before the link. Compression may alter the image and the effects of compression will be seen on selected imagery. However, compression effects are not objectionable in most imagery.

For remote applications the image may be recorded on a separate video recorder[9,10] (Figure 1-3) or may be obtained with a camcorder. The most popular player is the video cassette recorder (VCR). The recorder further modifies the image by reducing the signal-to-noise ratio (SNR) and image sharpness. As the desire for portability increases, cameras and recorders will shrink in size.[10-13]

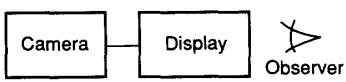


Figure 1-1. A closed circuit television system. Image quality is determined by an observer.

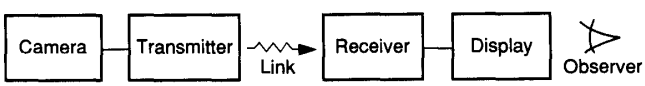


Figure 1-2. Generic remote transmission system. The transmitter and receiver must have adequate electronic bandwidth to support the image quality created by the solid state camera.

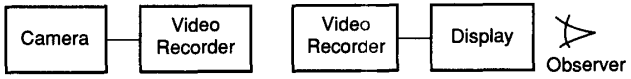


Figure 1-3. Imagery can be stored on video tape. However, the recorder circuitry may degrade the image quality.

For scientific applications, the camera output is digitized and then processed by a computer (Figure 1-4). After processing, the image may be presented on a monitor (soft copy), printed (hard copy), or stored. The digital image can also be transported to another computer via the Internet, local area net, floppy disc, or magnetic tape. For remote applications, the digital data may be stored on a digital recorder (Figure 1-5).[14-16]

Perhaps the most compelling reason for adopting digital technology is the fact that the quality of digital signals remains intact through copying and reproduction unless the signals are deliberately altered. Digital signal "transmission" was first introduced into tape recorders. Because a bit is either present or not, multiple generation copies retain high image quality.

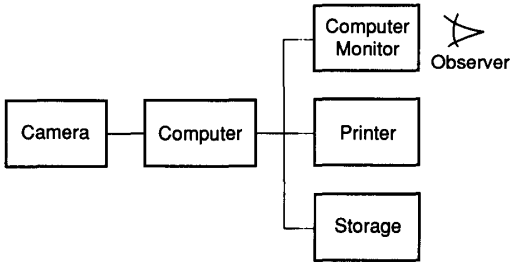


Figure 1-4. Most imagery today is enhanced through image processing. The camera may view a scene directly or may scan a document. The computer output hard copy may be used in newspapers, advertisements, or reports.

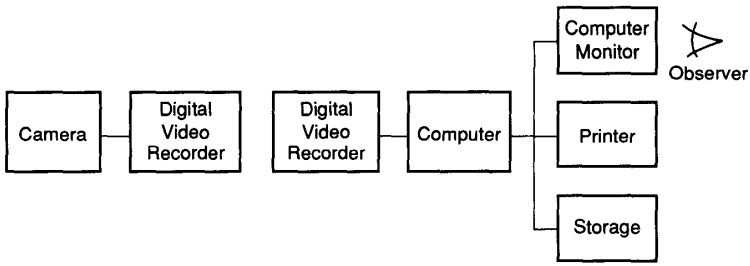


Figure 1-5. Digital systems can provide very high quality imagery. Electronic bandwidth limitations may impose data compression requirements. Data compression may alter the image but this alteration may not be obvious on general imagery.

In comparison, analog recorders, such as in video home systems (VHS), provide very poor quality after just a few generations. The first digital recorder used a format called D-1 and it became known as the CCIR 601 component digital standard. Digital recording formats now include[16] D-1, D-2, D-3, D-4, D-5, and D-6. Error correcting codes can be used to replace missing bits.

New digital recorders combine advanced technologies in electronics, video compression, and mechanical transport design. Real-time digital video systems operate at data rates that are faster than most computers. While errors in the imagery are never considered desirable, the eye is very tolerant of defects. In comparison, a computer is considered worthless if the error rate was high.

When it comes to displaying images (either hard copy or soft copy), the range of the camera digitizer is often greater than the display device. A camera may offer 12, 14, or 16 bits whereas the display may only be 8 bits. Here, a look-up table is employed to match the camera output to the display. This may be a simple linear relationship or specific expansion of a part of the camera's gray scale. False color can also be introduced to enhance the displayed image. False color is often useful for the human observer but serves no function in machine vision systems.

While a machine vision system does not require a monitor (Figure 1-6), a monitor is often used during system setup and for diagnostic evaluation. That is, a computer algorithm compares an image to a stored standard. If the target does not compare favorably to the standard, then the process is changed. For example, this may mean sending a rejected item for rework or changing the light intensity.

Figure 1-6. Machine vision systems do not require a monitor. The computer output controls the manufacturing process. Monitors are used for setup and diagnostic evaluation.

1.4. IMAGE QUALITY

Image quality is a subjective impression ranking of imagery from poor to excellent. It is a somewhat learned skill. It is a perceptual ability, accomplished by the brain, affected by and incorporating other sensory systems, emotions, learning, and memory. The relationships are many and not well understood. Perception varies between individuals and over time. There exist large variations in an observer's judgment as to the correct rank ordering from best to worst and therefore image quality cannot be placed on an absolute scale. Visual psychophysical investigations have not measured all the properties relevant to imaging systems.

Many formulas exist for predicting image quality. Each is appropriate under a particular set of viewing conditions. These expressions are typically obtained from empirical data in which multiple observers view many images with a known amount of degradation. The observers rank order the imagery from worst to best and then an equation is derived that relates the ranking scale to the amount of degradation.

Early metrics were created for film-based cameras. Image quality was related to the camera lens and film modulation transfer functions (MTFs). With the advent of televisions, image quality centered on the perception of raster lines and the minimum SNR required for *good* imagery. Here, it was assumed that the studio camera provided a perfect image and only the receiver affected the image quality.

The system MTF is the major component of system analysis. It describes how sinusoidal patterns propagate through the system. Because any target can be decomposed into a Fourier series, the MTF approach indicates how imagery will appear on the display.

Many tests have provided insight into image metrics that are related to image quality. Most metrics are related to the system MTF, resolution, or the signal-to-noise ratio. In general, images with higher MTFs and less noise are judged as having better image quality. There is no single *ideal* MTF shape that provides best image quality.

Digital processing is used for image enhancement and analysis. Because the pixels are numerical values in a regular array, mathematical transforms can be applied to the array. The transform can be applied to a single pixel, group of pixels, or the entire image. Many image processing designers think of images as an array of numbers that can be manipulated with little regard to who is the final interpreter.

With a solid state camera system, the lens, array architecture, array noise, and display characteristics all affect system performance. Only an end-to-end assessment will determine the overall image quality. For example, there is no advantage to using a high quality camera if the display cannot produce a faithful image. Often, the display is the limiting factor in terms of image quality and resolution. No matter how good the camera is, if the display resolution is poor, then the overall system resolution is poor. Only if the display's contrast and spatial resolution are better than the camera will the camera image quality be preserved.

For consumer applications, system resolution is important with SNR being secondary. Scientific applications may place equal importance on resolution and MTF. The military, interested in target detection, couples the system MTF and system noise to create the perceived signal-to-noise ratio (SNR_P).

$$SNR_P = k \frac{MTF_{SYS}\,\Delta I}{(system\ noise)} \frac{1}{(eye\ spatial filter)(eye\ temporal filter)}, \quad (1\text{-}1)$$

where ΔI is the intensity difference between the target and its immediate background. The factor k is a proportionality constant that depends on the specific camera characteristics (aperture diameter, focal length, quantum efficiency, etc.).

When the perceived SNR is above a threshold value, the target is just perceived. Selecting a threshold SNR and solving the equation for ΔI provides the minimum resolvable contrast (MRC). When the MRC is coupled with the target description and atmospheric transmittance effects, the target range can be predicted.

1.5. PIXELS, DATELS, DISELS, and RESELS

The overall system may contain several independent sampling systems. The array spatially samples the scene, the computer may have its own digitizer, and the monitor may have a limited resolution. A monitor "pixel" may or may not represent a "pixel" in camera space. The designer and user must understand the differences among the sampling lattices.

Sampling theory describes the requirements that lead to the reconstruction of a digitized signal. The original analog signal is digitized, processed, and then returned to the analog domain. For electrical signals, each digitized value is simply called a sample and the digitization rate is the sampling rate.

Electronic imaging systems are more complex and several sampling lattices are present. The detector output represents a sampling of the scene. The detector output is digitized and placed in a memory. After image processing, the data are sent to a display medium. Although the display medium provides an analog signal, it is typically digitally controlled.

Each device has its own minimum sample size or primary element. Calling a sample a pixel or pel (picture element) does not seem bad. Unfortunately, there is no *a priori* relationship between the various device pixels. The various digital samples in the processing path are called pixels, datels, and disels (sometimes called dixels). For analog systems, the minimum size is the resel (Table 1-3). Each "-el" is mapped onto the next (Figure 1-7). However, there is no standard definition for a resel. For optical systems it may be the Airy disk or Rayleigh criterion. For electronic circuits, it is related to the bandwidth but the definition has not been standardized. When a conversion takes place between the analog and digital domains, the resel may be different from the digital sample size. The analog signal rather than the digital sample may limit the system resolution. In oversampled systems, the resel consists of many samples.

Table 1-3
THE "-ELS"

ELEMENT	DESCRIPTION
Pixel or pel (picture element)	A sample created by a detector.
Datel (data element)	Each datum is a datel. Datels reside in a computer memory.
Disel (display element)	The smallest element (sample) that a display medium can access.
Resel (resolution element)	The smallest signal supported by an analog system.

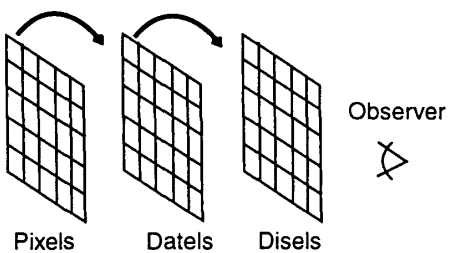


Figure 1-7. Each array is mapped onto the next. The number of elements in each array may be different. Not every array exists in every system.

For staring arrays, the total number of pixels is equal to the number of detectors. The detector's spatial response is determined by the detector's size (e.g., photosensitive area). If the extent is d_H in the horizontal direction and the optics focal length is fl, then the horizontal detector-angular-subtense (DAS_H) is

$$DAS_H = \frac{d_H}{fl}. \tag{1-2}$$

Similarly, in the vertical direction,

$$DAS_V = \frac{d_V}{fl}. \tag{1-3}$$

Staring arrays are often specified by the detector center-to-center spacing (pitch). The horizontal pixel-angular-subtense (PAS_H) is

$$PAS_H = \frac{d_{CCH}}{fl}, \tag{1-4}$$

where d_{CCH} is the horizontal pitch. Similarly, in the vertical direction

$$PAS_V = \frac{d_{CCV}}{fl}. \tag{1-5}$$

The fill factor is the ratio of areas:

$$Fill\ factor = \frac{d_H d_V}{d_{CCH} d_{CCV}}. \tag{1-6}$$

With a staring array that has a 100% fill factor, the PAS is equal to the DAS (e.g., $d_H = d_{CCH}$ and $d_V = d_{CCV}$).

If the Airy disk (a resel) is much larger than the PAS, then the optical resel determines the system resolution. If the number of pixels across this resel is large, the sampling theorem can be satisfied. If the electronic imaging system output is in a digital format, then the number of datels (samples) equals the number of pixels. If the camera's analog output is digitized, then the number of datels is linked to the frame grabber's digitization rate. This number can be much greater than the number of pixels. This higher number does not create more resolution in terms of the "-els." However, the higher sampling frequency may improve the overall system MTF.

Image processing algorithms operate on datels. A datel may not represent a pixel or a resel. Because there must be two samples to define a frequency, the Nyquist frequency associated with pixels, datels, and disels may be different. If there are more datels than pixels, the Nyquist frequency associated with an image processing algorithm is higher than that for the detector array. Most image processing books illustrate datels and call them pixels. The image processing specialist must understand the differences between the sampling lattices and take into account who will be the final data interpreter.

After image processing, the datels are output to a display medium. For monitors, each datel is usually mapped, one-to-one, onto each disel. Monitors are often specified by the number of addressable pixels (defined as disels in this text). Although the number of addressable pixels (disels) may be large for cathode ray tube based monitors, the monitor resel may be limited by the electron beam diameter. With most display media, the finite-sized spots of two adjacent disels overlap to provide an analog (continuous) image.

Finally, the system designer must be aware of which subsystem limits the overall system resolution. In some respects, the starting point for system design should begin with the final interpreter of the data. The minimum "-el" should be discernible by the interpreter to ensure maximum transfer of information. However, the observer may not find this image aesthetically pleasing.

1.6. REFERENCES

1. W. S. Boyle and G. E. Smith, "Charge Coupled Semiconductor Devices," *Bell Systems Technical Journal*, Vol. 49, pp. 587-593 (1970).
2. G. F. Amelio, M. F. Tompsett, and G. E. Smith, "Experimental Verification of the Charge Coupled Concept," *Bell Systems Technical Journal*, Vol. 49, pp. 593-600 (1970).
3. M. J. Howes and D. V. Morgan, eds., *Charge-Coupled Devices and Systems*, John Wiley and Sons, NY (1979).
4. C. H. Sequin and M. F. Tompsett, *Charge Transfer Devices*, Academic Press, NY (1975).
5. E. S. Yang, *Microelectronic Devices*, McGraw-Hill, NY (1988).
6. E. L. Dereniak and D. G. Crowe, *Optical Radiation Detectors*, John Wiley and Sons, NY (1984).
7. A. J. P. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices*, Kluwer Academic Publishers, Dordrecht, The Netherlands (1995).
8. S. G. Chamberlain, S. R. Kamasz, F. Ma, W. D. Washkurak, M. Farrier, and P.T. Jenkins, "A 26.3 Million Pixel CCD Image Sensor," in *IEEE Proceedings of the International Conference on Electron Devices*, pp. 151-155, Washington, D.C. December 10, 1995.
9. Z. Q. You and T. H. Edgar, *Video Recorders: Principles and Operation*, Prentice Hall, New York, NY (1992).
10. M. Hobbs, *Video Cameras and Camcorders*, Prentice Hall, New York, NY (1989).
11. K. Tsuneki, T. Ezaki, J. Hirai, and Y. Kubota, "Development of the High-band 8 mm Video System," *IEEE Transactions on Consumer Electronics*, Vol. 35(3), pp. 436-440 (1989).
12. M. Oku, I. Aizawa, N. Azuma, S. Okada, K. Hirose, and M. Ozawa, "High Picture Quality Technologies for and S-VHS Portable VCR," *SMPTE Journal*, Vol. 98(9), pp. 636-639 (1989).
13. T. Kawamura, S. Kasai, T. Tominaga, H. Sato, and M. Inatsu, "A New Small-Format VTR Using an 8 mm Cassette," *SMPTE Journal*, Vol. 96(5), pp. 466-472 (1987).
14. J. Watkinson, *The D-3 Digital Video Recorder*, Focal Press, Oxford, England (1992).
15. K. Suesada, K. Ishida, J. Takeuchi, I. Ogura, and P. Livingston "D-5: 1/2-in. Full Bit Rate Component VTR Format, *SMPTE Journal*, Vol. 103(8), pp. 507-516 (1994).
16. J. Hamalainen, "Video Recording Goes Digital," *IEEE Spectrum*, Vol. 32(4), pp. 76-79 (1995).

2

RADIOMETRY and PHOTOMETRY

The radiometric/photometric relationship between the scene and camera is of primary importance. It determines the output signal. The second radiometric/photometric consideration is the match between the display's spectral output and the eye's spectral response. Display manufacturers have considered this in detail and this reduces the burden on the system designer. However, the displayed image typically is not a precise reproduction of the scene. While the displayed image usually represents the scene in spatial detail, the intensity and color rendition may be different.

Scenes in the visible and near infrared are illuminated by the sun, moon, starlight, night glow, or artificial sources. Since both the target and its background are illuminated by the same source, targets are detected when differences in reflectance exist. The camera's output voltage depends on the relationship between the scene spectral content and the spectral response of the camera.

Historically, cameras were designed to operate in the visible region only. For those systems, it was reasonable to specify responsivity in photometric units (e.g., volts/lux). However, when the spectral response extends past the visible region as with silicon photodetectors, the use of photometric units can be confusing and even misleading.

Radiometry describes the energy or power transfer from a source to a detector. When normalized to the eye's response, photometric units are used. Radiometric and photometric quantities are differentiated by subscripts: e is used for radiometric units, q for photons, and v is used for photometric units.

The symbols used in this book are summarized in the *Symbol List* (page xviii) which appears after the *Table of Contents*.

2.1. RADIATIVE TRANSFER

Spectral radiant sterance, L_e, is the basic quantity from which all other radiometric quantities can be derived. It contains both the areal and solid angle concept[1] that is necessary to calculate the radiant flux incident onto a system. It is the amount of radiant flux, (watts), per unit wavelength (micrometers) radiated into a cone of incremental solid angle (steradians) from a source whose area is measured in meters (Figure 2-1)

$$L_e(\lambda) = \frac{\partial^2 \Phi(\lambda)}{\partial A_S \, \partial \Omega} \quad \frac{W}{m^2\text{-}\mu m\text{-}sr} . \tag{2-1}$$

Similarly, L_q is the spectral photon sterance expressed in photons/(s-m^2-μm-sr). The variables L_e and L_q are invariant for an optical system that has no absorption or reflections. That is, L_e and L_q remain constant as the radiation transverses through the optical system. Table 2-1 provides the standard radiometric units.

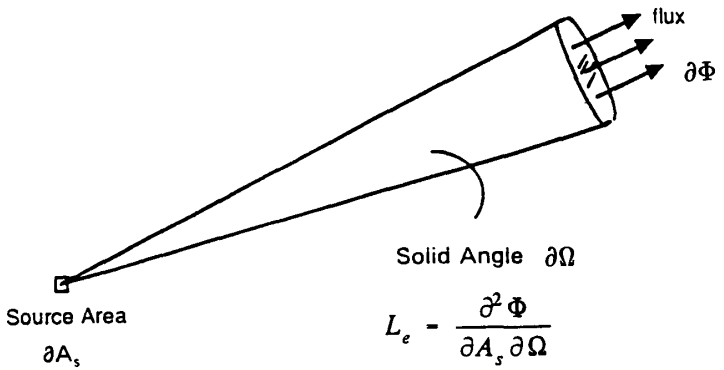


Figure 2-1. Radiant sterance.

Table 2-1
STANDARD SPECTRAL RADIOMETRIC UNITS

DEFINITION	SYMBOL	UNITS
Radiant energy	Q_e	Joule (J)
Radiant flux	Φ_e	Watts (W)
Spectral radiant intensity	I_e	W/(m^2-μm-sr)
Spectral radiant exitance (from a source) (also radiant emittance)	M_e	W/(m^2-μm)
Spectral radiant incidance (onto a target) (also irradiance)	E_e	W/(m^2-μm)
Spectral radiant sterance (also radiance)	L_e	W/(m^2-μm-sr)

For Lambertian sources, the radiance is emitted into a hemisphere. Then the spectral radiant exitance is related to the spectral radiant sterance by

$$L_e(\lambda) = \frac{M_e(\lambda)}{\pi} \quad \frac{W}{m^2-\mu m-sr} \tag{2-2}$$

and

$$L_q(\lambda) = \frac{M_q(\lambda)}{\pi} \quad \frac{photons}{s-m^2-\mu m-sr} . \tag{2-3}$$

2.2. PLANCK'S BLACKBODY LAW

The spectral radiant exitance of an ideal blackbody source whose absolute temperature is T (Kelvin), can be described by Planck's blackbody radiation law

$$M_e(\lambda,T) = \frac{c_1}{\lambda^5}\left(\frac{1}{e^{(c_2/\lambda T)} - 1}\right) \quad \frac{W}{m^2\text{-}\mu m}, \tag{2-4}$$

where the first radiation constant is $c_1 = 3.7418 \times 10^8$ W-μm^4/m^2, the second radiation constant is $c_2 = 1.4388 \times 10^4$ μm-K, and λ is the wavelength expressed in micrometers. The value T is also called the color temperature. Figure 2-2 illustrates Planck's spectral radiant exitance in logarithmic coordinates. Since a photodetector responds linearly to the available power, linear coordinates may provide an easier representation to interpret (Figure 2-3). Each curve has a maximum at λ_{PEAK}. Wien's displacement law provides $\lambda_{PEAK}\,T = 2898$ μm-K. A source must have an absolute temperature above about 700 K to be perceived by the human eye.

The spectral photon exitance is simply the spectral radiant exitance divided by the energy of one photon (hc/λ):

$$M_q(\lambda,T) = \frac{\lambda}{hc} M_e(\lambda,T) \tag{2-5}$$

or

$$M_q(\lambda,T) = \frac{c_3}{\lambda^4}\left(\frac{1}{e^{(c_2/\lambda T)} - 1}\right) \quad \frac{photons}{s\text{-}m^2\text{-}\mu m}, \tag{2-6}$$

where h is Planck's constant ($h = 6.626 \times 10^{-34}$ J-s), c is the speed of light ($c = 3 \times 10^8$ m/s), and the third radiation constant is $c_3 = 1.88365 \times 10^{27}$ photons-μm^3/s-m^2. Figures 2-4 and 2-5 provide the spectral photon exitance in logarithmic and linear coordinates, respectively, for sources typically used for CCD calibration. These curves have a maximum at $\lambda_{PEAK}\,T = 3670$ μm-K.

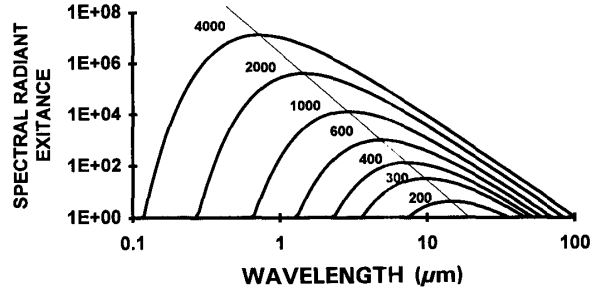


Figure 2-2. Planck's spectral radiant exitance plotted in logarithmic coordinates. The units are W/(m^2-μm). The light line is Wien's law.

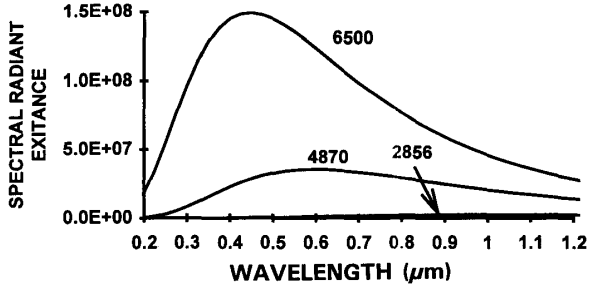


Figure 2-3. Planck's spectral radiant exitance plotted in linear coordinates for typical sources used for calibrating CCD arrays. The units are W/(m^2-μm). The maximum values occur at 1.01, 0.60, and 0.45 μm for color temperatures of 2856, 4870, and 6500K, respectively.

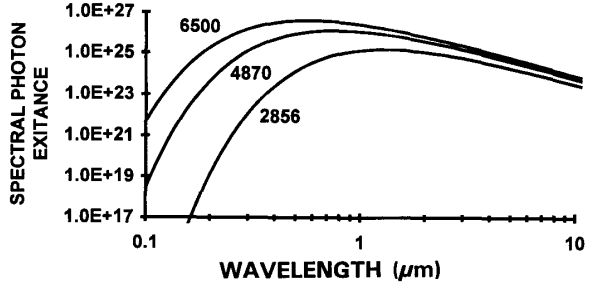


Figure 2-4. Planck's spectral photon exitance for T = 2856, 4870, and 6500 K. The maximum values occur at 1.29, 0.75, and 0.56 μm, respectively. The units are photons/(s-m^2-μm).

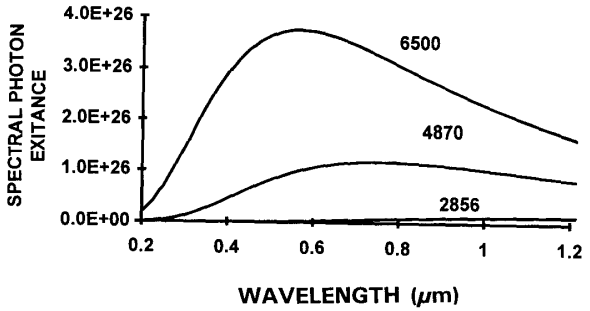


Figure 2-5. Planck's spectral photon exitance plotted in linear coordinates. The units are photons/(s-m^2-μm).

2.3. PHOTOMETRY

Photometry describes the radiative transfer from a source to a detector where the units of radiation have been normalized to the spectral sensitivity of the eye. Photometry applies to all systems that are sensitive to visible radiation. The luminous flux emitted by a source is

$$\Phi_v = K_M \int_{0.38\mu m}^{0.75\mu m} V(\lambda)\, M_p(\lambda)\, d\lambda \quad lumens\ , \tag{2-7}$$

where $M_p(\lambda)$ is the power expressed in units of W/μm and K_M is the luminous efficacy for photopic vision. It is 683 lumens/W at the peak of the photopic curve ($\lambda \approx 0.55$ μm) and 1746 lumens/W for the scotopic region at $\lambda \approx 0.505$ μm. Although both photopic and scotopic units are available, usually only the photopic units are used (Table 2-2 and Figure 2-6). At very low light levels (less than 5×10^{-3} lux), the eye's rods operate (scotopic response). For light levels above 5×10^{-2} lux, the eye's cones (photopic response) respond.

Table 2-2
PHOTOPIC and SCOTOPIC EYE RESPONSE

Wavelength nm	Photopic V(λ)	Scotopic V'(λ)	Wavelength nm	Photopic V(λ)	Scotopic V'(λ)
380		0.00059	570	0.952	0.2076
390	0.00012	0.00221	580	0.870	0.1212
400	0.0004	0.00929	590	0.757	0.0655
410	0.0012	0.03484	600	0.631	0.03315
420	0.0040	0.0966	610	0.503	0.01593
430	0.0116	0.1998	620	0.381	0.00737
440	0.023	0.3281	630	0.265	0.00335
450	0.038	0.455	640	0.175	0.00150
460	0.060	0.567	650	0.107	0.00067
470	0.091	0.676	660	0.061	0.00031
480	0.139	0.793	670	0.032	
490	0.208	0.904	680	0.017	
500	0.323	0.982	690	0.0082	
510	0.503	0.997	700	0.0041	
520	0.710	0.935	710	0.0021	
530	0.862	0.811	720	0.00105	
540	0.954	0.650	730	0.00052	
550	0.995	0.481	740	0.00025	
560	0.995	0.3288	750	0.00012	

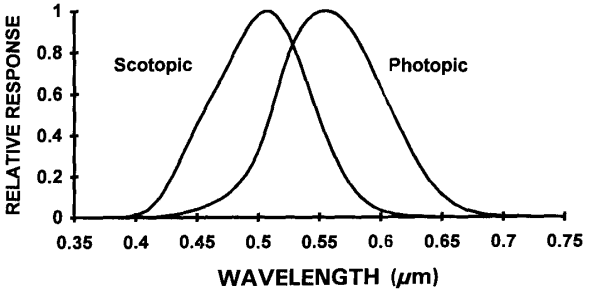


Figure 2-6. Photopic and scotopic eye responses.

2.3.1. UNITS

Unfortunately, there is overabundance of terminology being used in the field of photometry (Table 2-3). The SI units are recommended. Luminance emittance and illuminance are sometimes used as alternatives to luminous exitance and luminous incidance, respectively. Brightness and luminance may be used in place of luminous sterance. For convenience, luminance sterance has been labeled as the normalized intensity or nit. For Lambertian sources, the luminance is emitted into a hemisphere whose solid angle is 2π. The luminous exitance of a Lambertian surface is simply the luminance exitance given in Table 2-3 divided by π. For Lambertian sources, the luminance exitance terms are the apostilb (abs), Lambert (L), and foot-Lambert (fL) for the SI, CGS and English systems, respectively.

Figure 2-7 illustrates the geometric relationship between the SI, CGS, and English luminous incidance units. The numeric relationship is provide in Table 2-4.

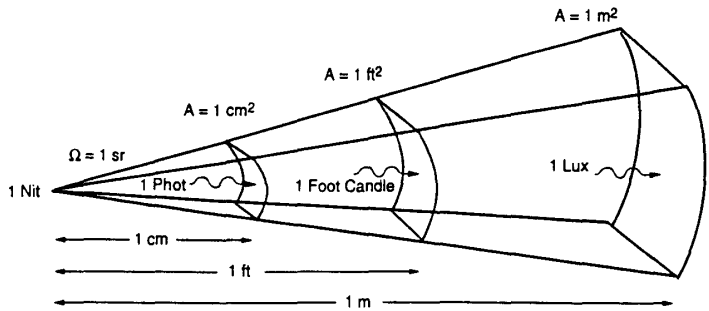


Figure 2-7. Geometric relationship between the SI, CGS, and English luminous incidance units. The solid angle is one steradian. (Not to scale).

Table 2-3
STANDARD PHOTOMETRIC UNITS
(The SI units are recommended)

DEFINITION	SYMBOL	SI and MKS UNITS	CGS UNITS	ENGLISH UNITS
Luminous energy	Q_v	Talbot (T)	Talbot (T)	Talbot (T)
Luminous flux	Φ_v	Lumen (lm)	Lumen (lm)	Lumen (lm)
Luminous intensity	I_v	Candela (cd) lm/sr	Candela (cd) lm/sr	Candela (cd) lm/sr
Luminous exitance (from a source) (also luminous emittance)	M_v	Lux (lx) lumen/m^2	Phot (ph) lumen/cm^2	Footcandle (fc) lumen/ft^2
Luminous incidance (onto a target) (also illuminance)	E_v	Lux (lx) Lumen/m^2	Phot (ph) lumen/cm^2	Footcandle (fc) lumen/ft^2
Luminous sterance (also brightness and luminance)	L_v	Nit Lumen/m^2-sr or cd/m^2	Stilb (sb) cd/cm^2	candela/ft^2

Table 2-4
CONVERSION BETWEEN SI, CGS, and ENGLISH UNITS

	PHOT	FOOTCANDLE	LUX
1 Phot =	1	929	1×10^4
1 Footcandle =	1.076×10^{-3}	1	10.764
1 Lux =	1×10^{-4}	0.0929	1

2.3.2. TYPICAL ILLUMINATION LEVELS

Natural lighting levels can vary by over 9 orders of magnitude (Table 2-5). The minimum level is limited by night glow and the maximum level is provided by the sun. At very low light levels (less than 5×10^{-3} lux), the eye's rods operate (scotopic response). For light levels above 5×10^{-2} lux, the eye's cones (photopic response) respond. Between these two values, both rods and cones are operating and the eye's response is somewhat between the two values. This composite response is called the mesopic response. Table 2-6 provides typical artificial lighting levels.

The eye automatically adapts to the ambient lighting conditions to provide an optimized image. Cameras may need neutral density filters if the light level is too high. Cameras may have an automatic iris and shutter speed to optimize the image on the detector array. If the light level is too low, an intensified camera may be necessary.

Table 2-5
NATURAL ILLUMINANCE LEVELS

SKY CONDITION	EYE RESPONSE	AVERAGE LUMINOUS INCIDANCE (lux)
Direct sun	Photopic	10^5
Full daylight	Photopic	10^4
Overcast sky	Photopic	10^3
Very dark day	Photopic	10^2
Twilight	Photopic	10
Deep twilight	Photopic	1.0
Full moon	Photopic	10^{-1}
Quarter moon	Mesopic	10^{-2}
Moonless, clear night (starlight)	Scotopic	10^{-3}
Moonless, overcast (night glow)	Scotopic	10^{-4}

Table 2-6
TYPICAL ARTIFICIAL ILLUMINANCE LEVELS

LOCATION	AVERAGE LUMINOUS INCIDANCE (lux)
Hospital operating theater	10^5
TV studio	10^3
Shop windows	10^3
Drafting office	500
Business office	250
Good street lighting	20
Poor street lighting	10^{-1}

2.4. SOURCES

As shown in Figure 2-3 (page 22), the peak wavelength shifts toward the blue end of the spectrum as the color temperature increases. This puts more energy in the visible region. There is approximately 70% more luminous flux available from a lamp operating at 3200 K than a light operating at 3000 K based on blackbody curves. This is the basis for using a higher color temperature lamp for higher illumination levels. However, the lifetime of tungsten halogen bulbs decreases dramatically with increasing color temperature. It is far better to use more lamps to increase the luminous flux than to increase the color temperature and sacrifice lifetime.

2.4.1. CALIBRATION SOURCES

The CIE (Commission Internationale de l'Eclairage or International Commission on Illumination) recommended four illuminants that could be used for calibration of cameras sensitive in the visible region (Table 2-7). Commercially available sources simulate these illuminants only over the visible range (0.38 to 0.75 μm). The source intensity is not specified; only the effective color temperature is specified. Illuminants A and D_{65} are used routinely whereas illuminants B and C are of historical interest.

Table 2-7
CIE RECOMMENDED ILLUMINANTS

CIE ILLUMINANT	EFFECTIVE COLOR TEMPERATURE	DESCRIPTION
A	2856 K	Light from an incandescent source
B	4870 K	Average noon sunlight
C	6770 K	Average daylight (sun + sky)
D_{65} or D_{6500}	6500 K	Daylight with a corrected color temperature

Illuminant A is a tungsten filament whose output follows Planck's blackbody radiation law. The other illuminants are created by placing specific filters in front of illuminant A. These filters have spectral characteristics that, when combined with illuminant A, provide the relative outputs illustrated in Figure 2-8. On a relative basis, the curves can be approximated by blackbodies whose color temperatures are 4870, 6770, and 6500 K for illuminant B, C, D_{65}, respectively.

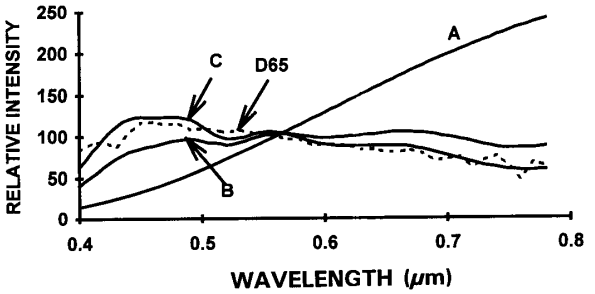


Figure 2-8. Relative output of the standard illuminants. Compare the shapes to Figure 2-3, page 22.

2.4.2. REAL SOURCES

A tungsten filament bulb emits light that is closely matched to an ideal blackbody radiator. This is not so for discharge lamps such as fluorescent lamps. For these lamps, a blackbody curve that approximates the output is fit to the spectral radiant exitance (Table 2-8). This approximation is for convenience and should not be used for scientific calculations. The actual spectral photon sterance must be used for nonideal sources[2] (discussed in the next three sections). Although many discharge lamps seem uniform white in color, they have peaks in the emission spectra (Figure 2-9). The apparent solar color temperature depends upon latitude, time of day, and a variety of environmental conditions (e.g., aerosols, pollutants, and cloud cover). The standardization of illuminant C at 6770 K was arbitrary. It was considered representative of "average" daylight.

Table 2-8
APPROXIMATE COLOR TEMPERATURE

SOURCE	APPROXIMATE COLOR TEMPERATURE
Northern sky light	7500 K
Average daylight	6500 K
Xenon (arc or flash)	6000 K
Cool fluorescent lamps	4300 K
Studio tungsten lamps	3200 K
Warm fluorescent lamps	3000 K
Floodlights	3000 K
Domestic tungsten lamps	2800 - 2900 K
Sunlight at sunset	2000 K
Candle flame	1800 K

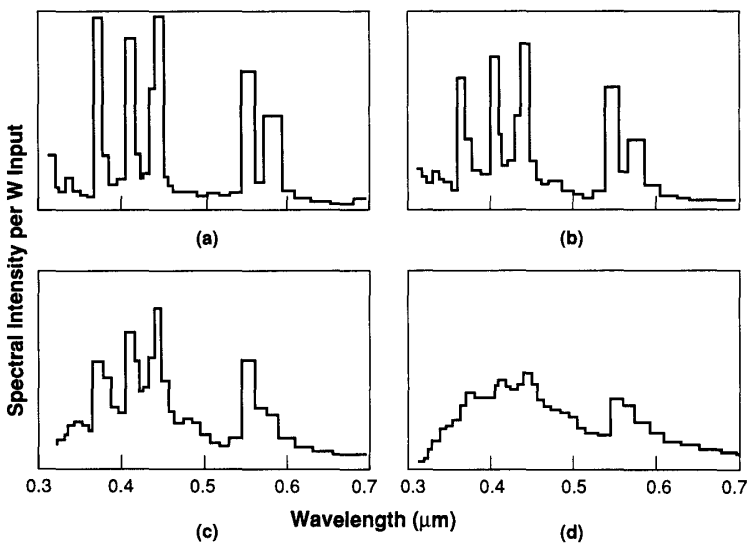


Figure 2-9. Relative output of a mercury arc lamp as a function of pressure.[3] (a) 21 atm, (b) 75 atm, (c) 165 atm, and (d) 285 atm.

Usually there is no relationship between the color temperature and contrast of a scene and what is seen on a display. This is not the fault of the camera and display manufacturers; they strive for compatibility. The observer, who usually has no knowledge of the original scene, will adjust the display for maximum visibility and aesthetics.

The video signal is simply a voltage with no color temperature associated with it. The display gain and offset can be adjusted such that a white target appears as if it was illuminated by a source whose color temperature is between 3200 K and 10,000 K. Most displays are preset to either 6500 K or 9300 K (Figure 2-10). As the color temperature increases, whites appear to change from a yellow tinge to a blue tinge. The perceived color depends on the adapting illumination (e.g., room lighting) Setting the color temperature to 9300 K provides aesthetically pleasing imagery and this setting is unrelated to the actual scene color temperature.

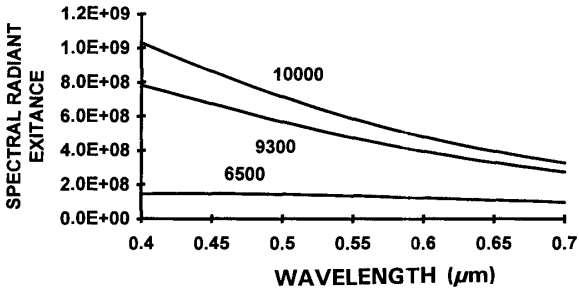


Figure 2-10. Display color temperature is typically preset at either 6500 or 9300 K. Some displays allow tuning from 3200 to 10,000 K. The displayed color temperature is independent of the scene color temperature. A scene illuminated with a 2854 K source (yellow tinge) can appear bluish when displayed at 9300 K.

2.5. CAMERA FORMULA

If an imaging system is at a distance R_1 from a source (Figure 2-11), the number of photons incident onto the optical system of area A_o is during time t_{INT} is

$$n_{LENS} = L_q \frac{A_o}{R_1^2} A_S T_{ATM} t_{INT} \ , \tag{2-8}$$

where the small angle approximation was used (valid when $R_1^2 \gg A_o$). T_{ATM} is the intervening atmospheric transmittance. The number of on-axis photons reaching the image plane is:

$$n_{IMAGE} = L_q \frac{A_o}{R_2^2} A_S T_{OPTICS} T_{ATM} t_{INT} \ , \tag{2-9}$$

where T_{OPTICS} is the system's optical transmittance.

When the image size is much larger than the detector area ($A_I \gg A_D$), the source is said to be resolved or the system is viewing an extended source. Equivalently, the detector is flood-illuminated. In most solid state camera applications, the source is resolved. The value A_D is the effective sensitive detector area. Microlenses (discussed in Section 3.11.1., *Microlenses*) and optical anti-alias filters (discussed in Section 10.3.4., *Optical Anti-alias Filter*) increase the effective area.

The number of photons incident onto the detector is simply the ratio of the areas:

$$n_{DETECTOR} = n_{IMAGE} \frac{A_D}{A_I} \ . \tag{2-10}$$

Using the small angle approximation for paraxial rays,

$$\frac{A_S}{R_1^2} = \frac{A_I}{R_2^2} \ . \tag{2-11}$$

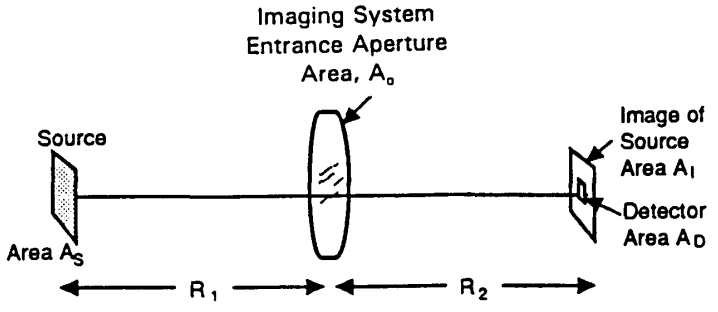


Figure 2-11. An imaging system directly viewing a source.

The number of photons reaching the detector becomes:

$$n_{DETECTOR} = \frac{L_q A_o A_D}{fl^2 (1 + M_{OPTICS})^2} T_{OPTICS} T_{ATM} t_{INT} \quad . \tag{2-12}$$

The optical magnification is $M_{OPTICS} = R_2/R_1$. Here, R_1 and R_2 are related to the system's effective focal length, fl, by

$$\frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{fl} \quad . \tag{2-13}$$

Assuming a circular aperture ($A_o = \pi D^2/4$) and defining the f-number as F = fl/D (see Appendix):

$$n_{DETECTOR} = \frac{\pi}{4} \frac{L_q A_D}{F^2 (1 + M_{OPTICS})^2} T_{OPTICS} T_{ATM} t_{INT} \quad . \tag{2-14}$$

An off-axis image will have reduced incidance compared to an on-axis image by cosine$^4\theta$. As required, cosine$^4\theta$ can be added to all the equations.

Many of the variables are a function of wavelength. The number of photoelectrons generated in a solid state detector is

$$n_{pe} = \int_{\lambda_1}^{\lambda_2} R_q(\lambda) n_{DETECTOR}(\lambda)\, d\lambda \qquad (2\text{-}15)$$

or

$$n_{pe} = \int_{\lambda_1}^{\lambda_2} R_q(\lambda) \frac{\pi}{4} \frac{L_q(\lambda)\, t_{INT}\, A_D}{F^2 (1 + M_{OPTICS})^2} T_{OPTICS}(\lambda)\, T_{ATM}(\lambda)\, d\lambda \ , \qquad (2\text{-}16)$$

where $R_q(\lambda)$ has units of electrons per photon and is simply the detector's quantum efficiency and t_{INT} becomes is the integration time. As the source moves to infinity, M_{OPTICS} approaches zero.

In photography, shutter speeds (exposure times) vary approximately by a factor of two (e.g., 1/30, 1/60, 1/125, 1/250, etc.). Thus, changing the shutter speed by one setting changes n_{pe} approximately by a factor of 2. f-stops have been standardized to 1, 1.4, 2, 2.8, 4, 5.6, 8..... The ratio of adjacent f-stops is $\sqrt{2}$. Changing the lens speed by one f-stop changes the f-number by a factor of $\sqrt{2}$. Here, also, the n_{pe} changes by a factor of 2.

In an actual application, what is of interest is the signal difference produced by a target and its immediate background. Here, both the target and the background are assumed to be illuminated by the same source (artificial lighting, sun, moon, night glow, star light). Let $\Delta\rho = \rho_T(\lambda) - \rho_B(\lambda)$ where $\rho_T(\lambda)$ and $\rho_B(\lambda)$ are the spectral reflectances of the target and background, respectively. Although the number of electrons is used for solid state array calculations, the camera output is a voltage:

$$\Delta V_{CAMERA} = G_{CAMERA} \Delta n_{pe} \ , \qquad (2\text{-}17)$$

where

$$\Delta n_{pe} = \int_{\lambda_1}^{\lambda_2} R_q(\lambda) \frac{\pi}{4} \frac{\Delta\rho\, E_{q\text{-}SCENE}(\lambda)\, t_{INT}\, A_D}{F^2 (1 + M_{OPTICS})^2} T_{OPTICS}(\lambda)\, T_{ATM}(\lambda)\, d\lambda \ . \qquad (2\text{-}18)$$

The value G_{CAMERA} contains both the array output conversion gain (units of volts/electron) and the subsequent amplifier gain. Tables 2-5 and 2-6 provide nominal values for $E_{e\text{-}SCENE}$.

It would appear that by simply decreasing the f-number, n_{pe} or Δn_{pe} would increase. Often, lower f-number lens systems have more optical elements and therefore may have lower transmittance. The value T_{OPTICS}/F^2 must increase to increase n_{pe} or Δn_{pe}. Lens systems may be described by the T-number:

$$T/\# = \frac{F}{\sqrt{T_{OPTICS}}} . \qquad (2\text{-}19)$$

If T_{OPTICS} is a function of wavelength, then T/# will also be a function of wavelength. Here T_{OPTICS}/F^2 in Equation 2-16 or 2-18 is replaced by $1/(T/\#)^2$.

For back-of-the-envelope calculations, the atmospheric transmittance is assumed to have no spectral features $T_{ATM}(\lambda) \approx T_{ATM}$ (discussed in Section 12.6.1., *Contrast Transmittance*). For detailed calculations, the exact spectral transmittance[4] must be used.

☞

Example 2-1
VISUAL THRESHOLD

An object is heated to incandescence. What is the approximate color temperature for it to be just perceived? Let the object area be 1 cm^2. The observer is 1 m from the object.

After adapting to the dark at least 60 min, the eye can perceive a few photons per second (absolute threshold). Here, the pupil dilates to about 8 mm ($A_o = 5 \times 10^{-5}\ m^2$). When the eye is dark adapted, only the rods are functioning and the eye's spectral response is approximately 0.38 μm to 0.66 μm. The number of photons per second reaching the eye from a Lambertian source is

$$n_{EYE} = \int_{0.38\,\mu m}^{0.66\,\mu m} \frac{A_o}{R_1^2} \frac{M_q(\lambda,T)}{\pi} A_S d\lambda \approx \frac{A_o}{R_1^2} \frac{M_q(\lambda_o,T)}{\pi} A_S \Delta\lambda , \qquad (2\text{-}20)$$

where $\Delta\lambda = 0.28\,\mu m$. The peak scotopic eye response occurs at $\lambda_o = 0.515\,\mu m$. The observer can just perceive this object at 640 K when he is in a completely blackened room. As the ambient illumination increases, the pupil constricts (A_o decreases). Furthermore, the eye's detection capability depends on the ambient lighting. As the light level increases, the target flux must also increase to be perceptible.

The eye's ability to just discern intensity differences depends on the intensity of the surrounding illuminance. At 0.01 lux the object must provide 10 times more flux than the background to be discerned by the eye. If the surround is at 700 K, than the target must be at 750 K to be perceptible. At 1000 lux, the object must provide 1000 times more flux then the surround. If the surround color temperature is 1000 K, then the target must be heated to 1350 K to be perceptible. Photometric units are not linearly related to color temperature.

These back-of-the-envelope calculations do not include the eye's spectral response (Table 2-2, page 24) and therefore only illustrate required intensity differences. The eye changes its ability to perceive differences with the surround illumination. This explains, in part, why television receiver imagery appears so bright in a darkened room and cannot be seen in bright sunlight.

2.6. NORMALIZATION

"Normalization .... is the process of reducing measurement results as nearly as possible to a common scale"[5]. Normalization is essential to ensure that appropriate comparisons are made. Figure 2-12 illustrates the relationship between the spectral response of a system to two different sources. The output of a system depends on the spectral features of the input and the spectral response of the imaging system. Simply stated, the camera output depends on the source used.

Variations in output can also occur if "identical" systems have different spectral responses. Equation 2-16 (page 35) is integrated over the wavelength interval of interest. Since arrays may have different spectral responses, an imaging system whose spectral response is 0.4 to 0.7 μm may have a different responsivity than a system that operates 0.38 to 0.75 μm although both systems are labeled as visible systems. The spectral response of solid state arrays varies from manufacturer to manufacturer (discussed in Section 4.1.1., *Spectral Response*). Systems can be made to appear as equivalent or one can be made to provide better performance by simply selecting an appropriate source.

While the sensor may be calibrated with a standard illuminant, it may not be used with the same illumination. Calibration with a standard illuminant is useful for comparing camera responsivities under controlled conditions. However, if the source characteristics in a particular application are significantly different from the standard illuminant, the selection of one sensor over another must be done with care. For example, street lighting from an incandescent bulb is different from that of mercury-arc lights (compare Figure 2-8 with Figure 2-9, pages 30 and 31, respectively). The only way to determine the relative responsivities is to perform the calculation indicated in Equation 2-16.

The effect of ambient lighting is particularly noticeable when using low light level televisions at nighttime. The spectral output of the sky (no moon) and moon (clear night) are significantly different. Figure 2-13 illustrates the natural night sky luminous incidence. Since an abundance of photons exists in the near infrared (0.7 to 1.1 μm), most night vision systems (e.g., image intensifiers, starlight scopes, snooper scopes, etc.) are sensitive in this region.

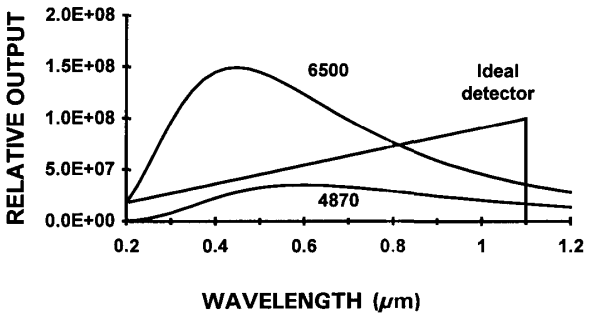


Figure 2-12. Sources with different spectral outputs can produce different system outputs. The 6500-K source provides more in-band radiant flux than the 4870-K source. The system output will be higher when viewing the 6500-K source.

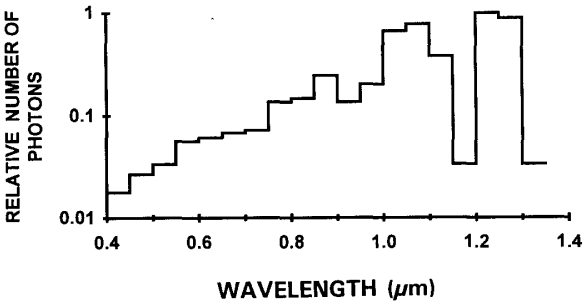


Figure 2-13. Night sky luminous incidance (From Reference 6).

2.7. NORMALIZATION ISSUES

A camera output depends on both the system spectral response and the color temperature of the illuminating source (Equation 2-16, page 35). When the camera spectral response is within the eye's response, then photometric units are reasonable. When the camera's response is outside the eye's response curve, photons contribute to signal but the photometry remains constant. Even though two sources may provide the same luminous incidance, the camera output can be quite different. Sometimes an infrared filter is used to restrict the wavelength response. When the filter is added, the average responsivity becomes

$$R_v = \frac{\int_{\lambda_1}^{\lambda_2} T_{IR\text{-}FILTER} R_e(\lambda) M_e(\lambda)\, d\lambda}{683 \int_{0.38}^{0.75} V(\lambda) M_e(\lambda,T)\, d\lambda} \quad \frac{A}{lumens} . \tag{2-21}$$

This average type responsivity is useful for comparing the performance of cameras with similar spectral responses. It is appropriate if the source employed during actual usage is similar to the one used for the calibration.

For ideal photon detectors the spectral responsivity (expressed in A/W) is:

$$R_e(\lambda) = \frac{\lambda}{\lambda_P} R_P \quad when \quad \lambda \le \lambda_p$$
$$= 0 \quad elsewhere \quad , \tag{2-22}$$

where R_P is the peak responsivity and λ_P is the wavelength at which R_P occurs. Since R_q is the quantum efficiency, $R_e = (q\lambda/hc)R_q$. When the wavelength is expressed in micrometers, $R_e = (\lambda/1.24)R_q$ and $R_P = R_q/E_G$. For silicon, the band gap is $E_G = 1.12$ eV resulting in a cutoff wavelength of $\lambda_P \approx 1.24/E_G \approx 1.1$ μm. The quantity q is the electron charge ($q = 1.6 \times 10^{-19}$ coul).

The ideal silicon detector response is used to estimate the relative outputs of two systems: one with an ideal IR filter (transmittance is unity from 0.38 to 0.70 μm and zero elsewhere) and one without. The relative outputs are normalized to that expected when illuminated with a CIE A illuminant (Table 2-9). The sources are considered ideal blackbodies with exitances that follow Equation 2-4 (page 21). As the source temperature increases, the output increases but not linearly with color temperature.

Table 2-9
RELATIVE OUTPUT
This is assuming ideal blackbodies and ideal Si spectral response.
A specific sensor may deviate significantly from these values.

SOURCE TEMPERATURE	RELATIVE OUTPUT (0.38 to 1.1 μm)	RELATIVE OUTPUT (0.38 to 0.70 μm)
2856	1.00	0.196
3200	1.96	0.482
4870	15.7	6.95
6500	48.9	27.4
6770	56.5	32.4

Only photons within the visible region affect the number of available lumens. For the system without the IR filter, the detector is sensing photons whose wavelengths are greater than 0.7 μm even though the number of lumens is not affected by these photons (Figure 2-14). As the color temperature increases, the flux available to the observer (lumens) increases faster than the output voltage. As a result, responsivity, expressed in A/lumen, decreases with increasing color temperature (Table 2-10). This decrease affects the wide spectral response systems to a greater extent. Specifying the output in lumens is of little value for an electro-optical sensor whose response extends beyond the eye's response. Radiometric units are advised here.

The values in Tables 2-9 and 2-10 should be considered as illustrative. Solid state detectors do not follow the ideal response (discussed in Section 4.1.1., *Spectral Response*). The IR filter cutoff also varies by manufacturer. Only a detailed spectral response evaluation permits correct comparisons among detectors. Since there is considerable variation in spectral response and type of IR filter used, it is extremely difficult to compare systems based on responsivity values only.

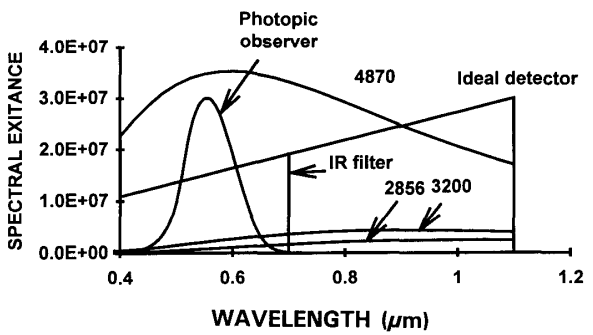


Figure 2-14. Spectral relationship between source characteristic, detector response, and the photopic observer. As the color temperature increases, the energy available to the observer increases rapidly. Therefore, the number of lumens increases dramatically.

Table 2-10
RELATIVE RESPONSIVITY (A/lumen)
A specific sensor may deviate significantly from these values

SOURCE TEMPERATURE	RELATIVE RESPONSIVITY (0.38 to 1.1 μm)	RELATIVE RESPONSIVITY (0.38 to 0.70 μm)
2856	1.00	0.196
3200	0.754	0.185
4870	0.387	0.171
6500	0.314	0.176
6770	0.308	0.177

Example 2-2
SOURCE SELECTION

A camera's signal-to-noise ratio is too small. The engineer can either change the light source to one with a higher color temperature or add more lamps. What are the differences?

With an ideal blackbody, the color temperature completely specifies the luminous exitance. Increasing the color temperature (by increasing the voltage on the bulb) will increase the luminous flux. For real sources, the color temperature is used only to denote the relative spectral content. A 50-W bulb whose output approximates a 3200-K source will have a lower photometric output than a 1000-W bulb operating at 2856 K. Adding more lamps does not change the relative spectral content but increases the luminous flux.

Example 2-3
RELATIVE VERSUS ABSOLUTE OUTPUT

Should a detector be calibrated with a 5- or 1000-W bulb?

If the relative spectral output of the two bulbs is the same (e.g., CIE A), the responsivity is independent of the source flux. The bulb intensity should be sufficiently high to produce a good signal-to-noise ratio. It should not be so high that the detector saturates.

System output does not infer anything about the source other than that an equivalent blackbody of a certain color temperature would provide the same output. This is true no matter what output units are used (volts, amps, or any other arbitrary unit). These units, by themselves, are not very meaningful for system-to-system comparison. For example, ΔV_{CAMERA} (and equivalently the camera responsivity) can be increased by increasing the system gain. As such, it is dangerous to compare system response based on only a few numbers.

Arrays may be specified by their radiometric responsivity ($V/(J\text{-}m^{-2})$) or by their photometric responsivity (V/lux). The relationship between the two depends on the spectral content of the source and the spectral response of the array (further discussed in Section 4.1.2., *Responsivity*). In principle, the source can be standardized (e.g., selection of the CIE A illuminant), but the array spectral response varies among manufacturers and may vary within one manufacturing environment if different processes are used. Therefore, there is no simple (universal) relationship between radiometric and photometric responsivities.

The correct analysis method is to perform a radiometric calculation using the source spectral output (which may not be an ideal blackbody) and using the spectral response of the camera system. Using standard sources for quality control on a production line is valid since the spectral response of the detectors should not change very much from unit to unit. The camera spectral response must include the optical spectral transmittance of all filters present.

A 3200-K color temperature source is not well matched to the eye's response. The eye peaks at 0.555 μm but the source peaks at 0.90 μm. With silicon detectors, the spectral response tends to match the output of a 3200-K source. Thus CCD detectors will perform rather well with ambient lighting. For machine vision systems, the infrared-blocking filter is not typically used.

However, since the human observer is not familiar with infrared images, the imagery may appear slightly different compared to that obtained with the IR-blocking filter.

Low light level televisions can have a variety of spectral responses (discussed in Section 5.4.2., *Intensified Solid State Cameras*). The methodology presented in this chapter can be used to calculate responsivity and output voltage. The results given in Tables 2-9 and 2-10 do not apply to intensified systems.

2.8. REFERENCES

1. C. L. Wyatt, *Radiometric System Design*, Chapter 3, Macmillan Publishing Co., New York, NY (1987).
2. D. Kryskowski and G. H. Suits, "Natural Sources," in *Sources of Radiation*, G. J. Zissis, ed., pp. 151-209, Volume 1, *The Infrared and Electro-Optical Systems Handbook*, J. S. Accetta and D. L. Shumaker, eds., copublished by Environmental Research Institute of Michigan, Ann Arbor, MI, and SPIE Press, Bellingham, WA (1993).
3. E. B. Noel, "Radiation from High Pressure Mercury Arcs," *Illumination Engineering*, Vol. 36, pg. 243 (1941).
4. The ONTAR Corporation, 129 University Road, Brookline, MA 02146-4532, offers a variety of atmospheric transmittance codes.
5. F. E. Nicodemus, "Normalization in Radiometry," *Applied Optics*, Vol. 12(12), pp. 2960-2973 (1973).
6. *Burle Electro-Optics Handbook*, TP-135, pg. 73, Burle Industries, Inc., Lancaster, PA (1974). Formerly known as the *RCA Electro-Optics Handbook*.

3

SOLID STATE ARRAYS

CCD refers to a semiconductor architecture in which charge is transferred through storage areas. The CCD architecture has three basic functions: (a) charge collection, (b) charge transfer, and (c) the conversion of charge into a measurable voltage. A CCD can be attached to a variety of detectors. The basic building block of the CCD is the metal-insulator-semiconductor (MIS) capacitor. It is also called a gate. The most important MIS is the metal-oxide-semiconductor (MOS). Because the oxide of silicon is an insulator, it is a natural choice.

For monolithic devices, charge generation is often considered as the initial function of the CCD. With silicon photodetectors, each absorbed photon creates an electron-hole pair. Either the electrons or holes can be stored and transferred. For frame transfer devices, charge generation occurs under a MOS capacitor (also called a photogate). For some devices (notably interline transfer devices) photodiodes create the charge. The charge created at a pixel site is in proportion to the incident light level present. The aggregate effect of all the pixels is to produce a spatially sampled representation of the continuous scene.

Because most sensors operating in the visible region use a CCD architecture to move a charge packet, they are popularly called CCD arrays. However, the charge injection device (CID) does not use a CCD for charge transfer. Rather, two overlapping silicon MOS capacitors share the same row and column electrode. Pixel readout occurs by sensing the charge transfer between the capacitors at the pixel site.

Recently, active pixels have been introduced. These devices are fabricated with complementary-metal-oxide-semiconductor (CMOS) technology. The advantage is that one or more active transistors can be integrated into the pixel. As such, they become fully addressable (can read selected pixels) and can perform on-chip image processing.

Considerable literature[1-8] has been written on the physics, fabrication, and operation of CCDs. The charge transfer physics is essentially the same for all CCD arrays. However, the number of phases and number and location of the serial shift readout registers vary by manufacturer. The description that follows should be considered as illustrative. A particular design may vary significantly

from the simplified diagrams shown in this chapter. A specific device may not have all the features listed or may have additional features not listed.

Although CCD arrays are commonplace, the fabrication is quite complex. Theuwissen[9] provides an excellent step-by-step procedure for fabricating CCDs. He provides a 29-step procedure supplemented with detailed diagrams. Actual CCDs may require over 150 different operations. The complexity depends upon the array architecture.

Devices may be described functionally according to their architecture (frame transfer, interline transfer, etc.) or by application. To minimize cost, array complexity, and electronic processing, the architecture is typically designed for a specific application. For example, astronomical cameras typically use full frame arrays whereas video systems generally use interline transfer devices. The separation between general imagery, machine vision, scientific, and military devices becomes fuzzy as technology advances.

The symbols used in this book are summarized in the *Symbol List* (page xviii) which appears after the *Table of Contents*.

3.1. PHOTODETECTION

When an absorbed photon creates an electron-hole pair, the photodetection process has occurred. To be a useful detector in a solid state array, the photogenerated carrier must be collected in a storage site. The absorption coefficient is wavelength specific and decreases with increasing wavelength. Some of the incident photons will be reflected at the array's surface. The remainder may pass through layers of electrodes and insulators before reaching the photoactive material. Thus, the absorption characteristics of the overlying layers, thickness of the photoactive material, and the location of the storage site determine the spectral quantum efficiency (discussed in Section 4.1.1., *Spectral Response*).

After photodetection, the stored charge is read out. The array architecture often is optimized for a specific application. General imagery systems are designed for interlace operation consistent with a standard video (e.g., EIA 170). Research devices may operate in the progressive scan mode. This reduces some image distortion when the object is moving. For scientific applications the output may be in the slow scan mode, which simply means that the data read out rate is lower than a standard frame rate (e.g., EIA 170).

3.1.1. PHOTOGATE

The photogate is a photoactive MOS capacitor where the photogenerated carriers are stored in a depletion region. The storage capacity depends upon substrate doping, gate voltage, and oxide thickness. Very large capacities are possible. Because photogates have overlaying electrodes, the quantum efficiency in the blue region of the spectrum is reduced.

3.1.2. PHOTODIODE

With photodiodes, a depletion region is formed between the n- and p-type regions. The photogenerated carriers are usually stored at the junction until readout. The width of the junction limits the storage capability. The junction is initially reverse biased by a MOS gate. As the electrons, n_e, are collected, the voltage drops by

$$\Delta V = \frac{n_e q}{C} . \tag{3-1}$$

After readout, the voltage is reset to its original value. If the number of electrons is large, the voltage is reduced to zero and saturation occurs. Excess photoelectrons will diffuse to neighboring wells and create blooming. Photodiodes do not have an overlaying structure and therefore have higher quantum efficiency than comparable photogates.

With a photodiode, a transfer gate is required to move the stored electrons into the CCD. Devices with photodiodes are often labeled as CCPD for charge-coupled photodiodes. While transfer gates are essential for CCPDs, they may also exist in devices with photogates.

The hole accumulation diode (HAD) is a pinned photodiode that offers the best features of the photogate and photodiode. It offers the high blue response of the photodiode and large well capacity of the photogate.

3.2. CCD ARRAY OPERATION

Applying a positive voltage to the CCD gate causes the mobile positive holes in the p-type silicon to migrate toward the ground electrode because like charges repel. This region, which is void of positive charge, is the depletion region (Figure 3-1). If a photon whose energy is greater than the energy gap is

absorbed in the depletion region, it produces an electron-hole pair. The electron stays within the depletion region whereas the hole moves to the ground electrode. The amount of negative charge (electrons) that can be collected is proportional to the applied voltage, oxide thickness, and gate electrode area. The total number of electrons that can be stored is called the well capacity.

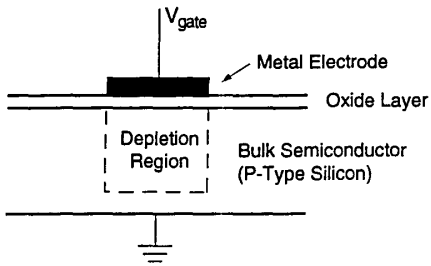


Figure 3-1. Metal-oxide-semiconductor (MOS) gate for p-type silicon. Although the depletion region (charge well) is shown to have an abrupt transition, its actual shape is gradual. That is, the well is formed by a two-dimensional voltage gradient.

The CCD register consists of a series of gates. Manipulation of the gate voltage in a systematic and sequential manner transfers the electrons from one gate to the next in a conveyor-belt-like fashion. For charge transfer, the depletion regions must overlap. The depletion regions are actually gradients and the gradients must overlap for charge transfer to occur. Efficient overlap occurs when the gate electrodes overlap (Figure 3-2). The multilevel polysilicon structure minimizes the gap between adjacent electrodes. Without this structure, charge transfer would be poor. For clarity, throughout the remainder of the text, the structure will be illustrated as adjacent, non-overlapping gates.

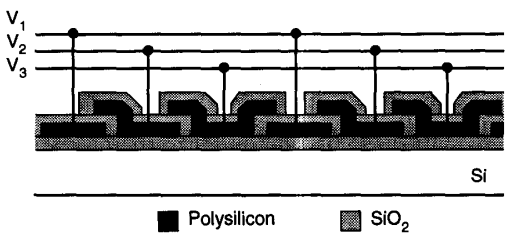


Figure 3-2. Three-phase CCD. The gates must overlap for efficient charge transfer.

Each gate has its own control voltage that is varied as a function of time. The voltage is called the clock or clocking signal. When the gate voltage is low, it acts as a barrier whereas when the voltage is high, charge can be stored. Initially, a voltage is applied to gate 1 and photoelectrons are collected in well 1 (Figure 3-3b). When a voltage is applied to gate 2, electrons move to well 2 in a waterfall manner (Figure 3-3c). This process is rapid and the charge quickly equilibrates in the two wells (Figure 3-3d). As the voltage is reduced on gate 1, the well potential decreases and electrons again flow in a waterfall manner into well 2 (Figure 3-3e). Finally, when gate 1 voltage reaches zero, all the electrons are in well 2 (Figure 3-3f). This process is repeated many times until the charge is transferred through the shift register.

The CCD array is a series of column registers (Figure 3-4). The charge is kept within rows or columns by channel stops or channel blocks and the depletion regions overlap in one direction only. At the end of each column is a horizontal register of pixels. This register collects a line at a time and then transports the charge packets in a serial fashion to an output amplifier. The entire horizontal serial register must be clocked out to the sense node before the next line enters the serial register. Therefore, separate vertical and horizontal clocks are required for all CCD arrays. This process creates a serial data stream that represents the two-dimensional image.

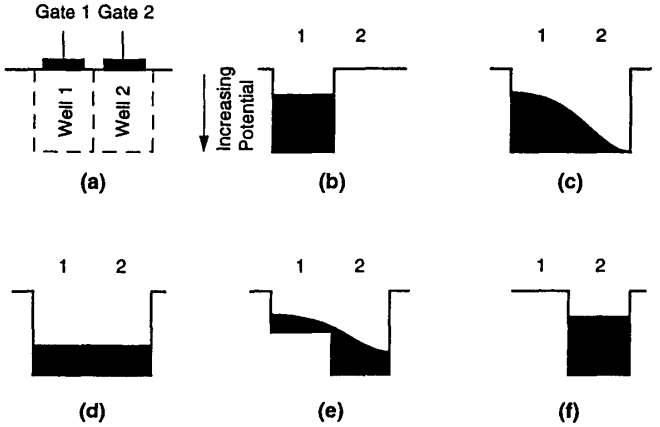


Figure 3-3. Charge transfer between two wells. (a) Adjacent wells. (b) Charge in well 1. (c) After a voltage is applied to gate 2, electrons flow into well 2. (d) Equilibration of charge. (e) Reduction of gate 1 voltage causes the electrons in well 1 to transfer into well 2. (f) All electrons have been transferred to well 2.

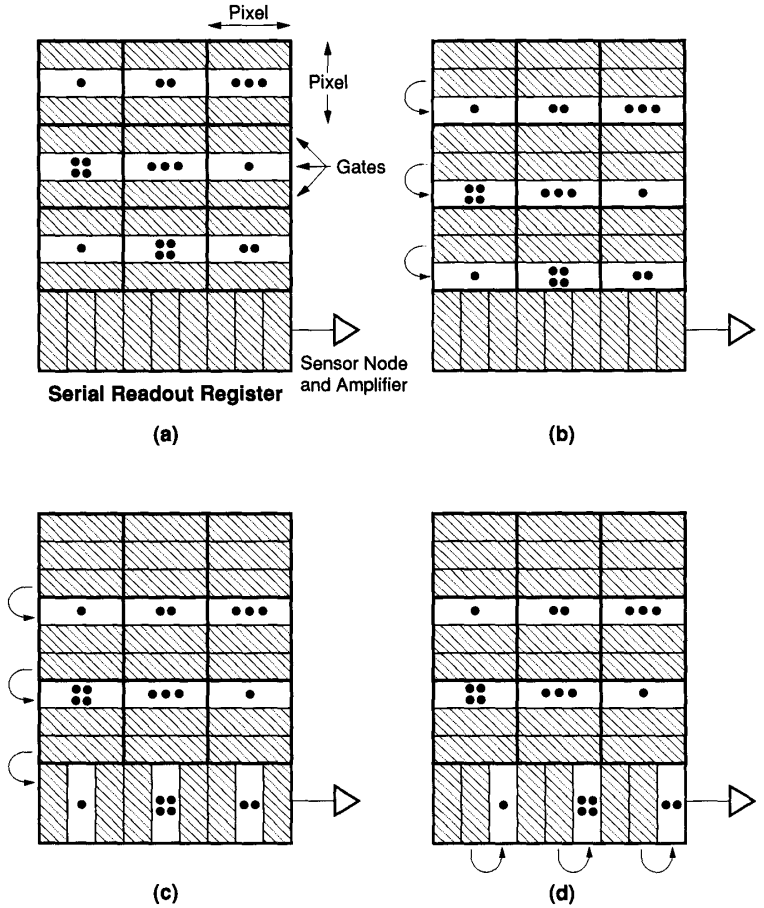


Figure 3-4. Representative CCD operation of a 3 × 3 array with three gates per pixel. Nearly all the photoelectrons generated within a pixel are stored in the well beneath that pixel. (a) The CCD is illuminated and an electronic image is created. (b) The columns are shifted down in parallel, one gate at a time. (c) Once in the serial row register, the pixels are shifted right to the output node. (d) The entire serial register must be clocked out before the next packet can be transferred into the serial readout register.

Although any number of transfer sites (gates) per detector area can be used, it generally varies from two to four. With a four-phase device, the charge is stored under two or three wells depending upon the clock cycle. Figure 3-5 represents the steady state condition after the gates have switched and the charge has equilibrated within the wells. It takes a finite time for this to happen. Figure 3-6 illustrates the voltage levels that created the wells shown in Figure 3-5. The time that the charge has equilibrated is also shown. The clock rate is $f_{CLOCK} = 1/t_{CLOCK}$ and is the same for all phases. The output video is valid once per clock pulse (selected as T_2) and occurs after the charge has equilibrated (e.g., Figure 3-3b or Figure 3-3f). Only one master clock is required to drive the array. Its phase must be varied for the gates to operate sequentially. An anti-parallel clocking scheme is also possible. Here V_3 is the inverse of V_1 and V_4 is the inverse of V_2. That is, when V_1 is high, V_3 is low. The anti-parallel clock system is easier to implement because phasing is not required. With the anti-parallel scheme, charge is always stored in two wells. For equal pixel sizes, four-phase devices offer the largest well capacity compared to the two- or three-phase systems. With the four-phase system, 50% of the pixel area is available for the well.

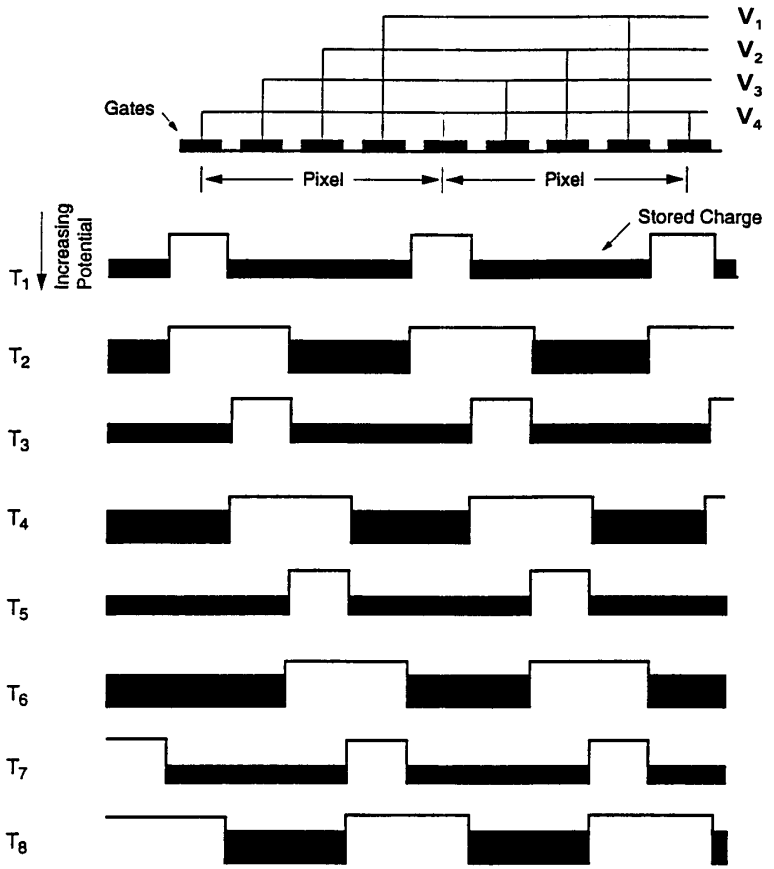


Figure 3-5. Charge transfer in a four-phase device after equilibration. This represents one column. Rows go into the paper. Charge is moved through storage sites by changing the gate voltages as illustrated in Figure 3-6. A four-phase array requires four separate clock signals to move a charge packet to the next pixel. A four-phase device requires eight steps to move the charge from one pixel to the next. (After Reference 10).

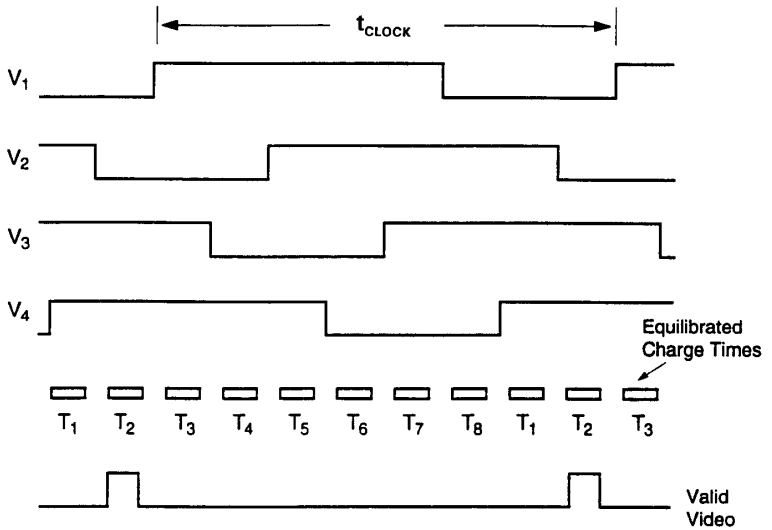


Figure 3-6. Voltage levels for a four-phase system. As the voltage is applied, the charge packet moves to that well. By sequentially varying the gate voltage, the charge moves off the horizontal shift register and onto the sense capacitor. The clock signals are identical for all four phases but offset in time (phase). The valid video is available once per clock pulse (selected as T_2). (After Reference 10).

The three-phase system (Figure 3-7 and Figure 3-8) stores the charge under one or two gates. Only 33% of the pixel area is available for the well capacity. With equal potential wells, a minimum number of three phases are required to clock out charge packets efficiently.

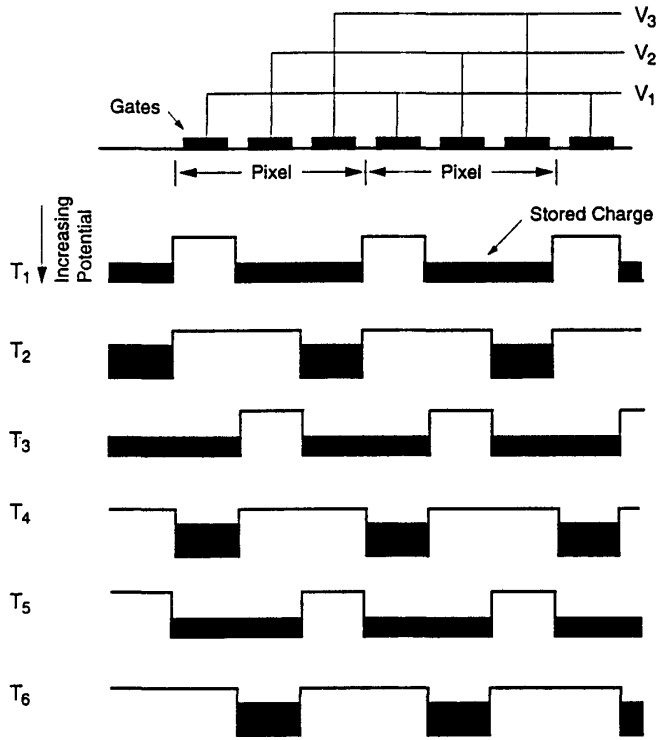


Figure 3-7. Charge transfer in a three-phase device. This represents one column. Rows go into the paper. Six steps are required to move the charge one pixel. (After Reference 10).

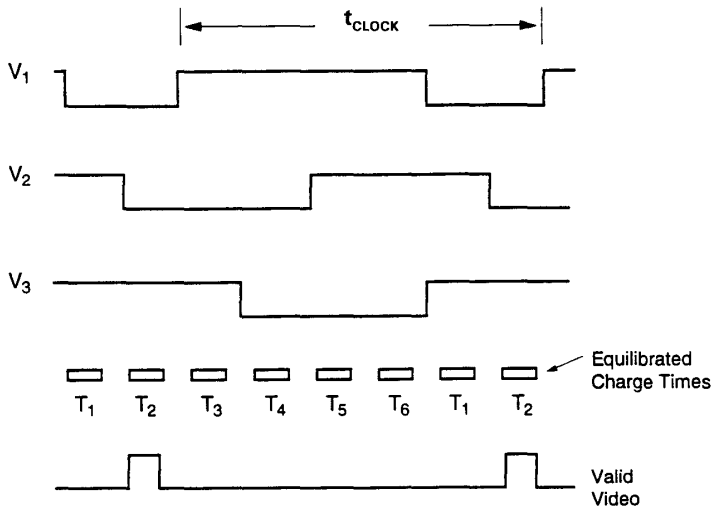


Figure 3-8. Voltage levels for a three-phase system. The clock signals are identical for all three phases but offset in time. T_2 was selected as the valid video time. Valid video exists only after the charge has equilibrated. (After Reference 10).

The well potential depends on the oxide thickness underneath the gate. In Figure 3-9, the alternating gates have different oxide thicknesses and therefore will create different well potentials. With different potentials it is possible to create a two-phase device. Charge will preferentially move to the right-hand side of the pixel where the oxide layer is thinner. Assume initially that the wells controlled by V_2 are empty. This will occur after a few clock pulses. At time T_1, both clocks are low. When V_2 is raised, the potential increases as shown at time T_2. Because the effective potential increases across the gates, the charge cascades down to the highest potential. Then V_2 is dropped to zero and the charge is contained under the V_2 gate at time T_3. This process is repeated until the charge is clocked off the array. The voltage timing is shown in Figure 3-10. While Figure 3-9 illustrates variations in the oxide layer, the potential wells can also be manipulated though ion implantation. As with the four-phase device, the two-phase device can also be operated in the anti-parallel mode: V_2 is a mirror image of V_1. That is, when V_1 is high, V_2 is low.

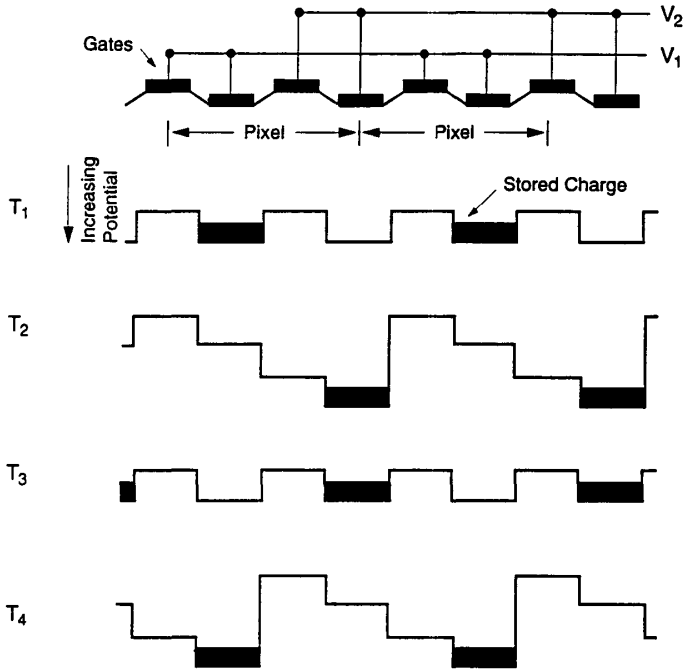


Figure 3-9. Charge transfer in a two-phase device. Well potential depends on the oxide thickness. (After Reference 10).

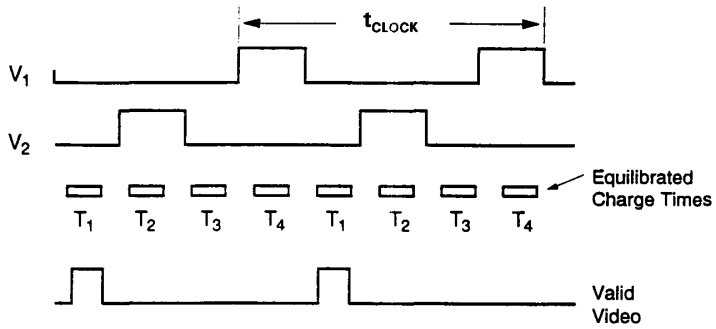


Figure 3-10. Voltage levels for a two-phase system. (After Reference 10).

The virtual phase device requires only one clock (Figure 3-11). Additional charge wells are created by internally biased gates. Charge is stored either beneath the active gate or the virtual gate. When V_1 is low, the charge will cascade down to the highest potential (which is beneath the virtual well). When V_1 is applied at T_2, the active gate potential increases and the charge moves to the active gate well.

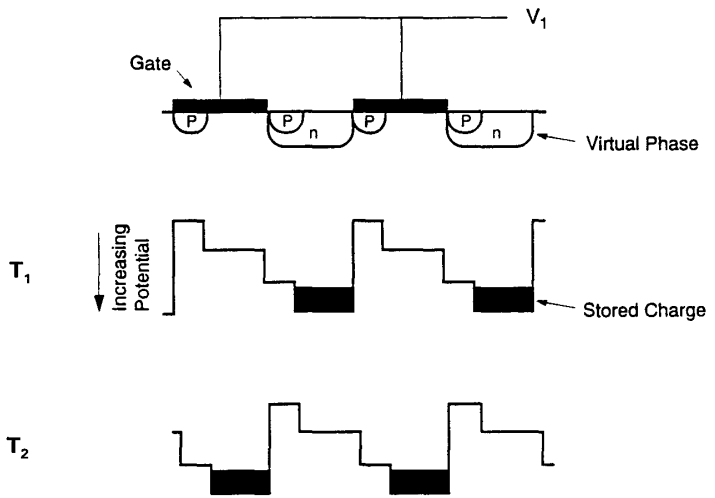


Figure 3-11. Charge transfer in a device with a virtual phase. The virtual well is created by p- and n-material implants. These ions create a fixed bias and therefore a well with fixed potential. By changing V_1, the active gate potential can be lower than the virtual well (T_1) or higher (T_2). This represents one column. Rows go into the paper.

The final operating step is to convert the charge packet to a measurable voltage. This is accomplished by a floating diode or floating diffusion. The diode, acting as a capacitor, creates a voltage proportional to the number of electrons [discussed in Section 3.5., *Charge Conversion (Output Structure)*]. With many arrays, it is possible to shift more than one row of charge into the serial register. Similarly, it is possible to shift more than one serial register element into a summing gate just before the output node. This is called charge grouping, binning, super pixeling, or charge aggregation. Binning increases

signal output and dynamic range at the expense of spatial resolution. Because it increases the signal-to-noise ratio, binning is useful for low light level applications for those cases where resolution is less important. Serial registers and the output node require larger capacity charge wells for binning operation. If the output capacitor is not reset after every pixel, then it can accumulate charge.

3.3. CCD ARRAY ARCHITECTURE

Array architecture is driven by the application. Full frame and frame transfer devices tend to be used for scientific applications. Interline transfer devices are used in consumer camcorders and professional television systems. Linear arrays, progressive scan, and time delay and integration (TDI) are used for industrial applications. Despite an ever increasing demand for color cameras, black and white cameras are widely used for many scientific and industrial applications.

3.3.1. LINEAR ARRAYS

The simplest arrangement is the linear array or single line of detectors. While these could be photogates, photodiodes are used more often because they have higher quantum efficiency. Linear arrays are used in applications where either the camera or object is moving in a direction perpendicular to the row of sensors. They are used where rigid control is maintained on object motion such as in document scanning.

Located next to each sensor is the transfer gate and then the CCD shift register (Figure 3-12a). The shift register is also light sensitive and is covered with a metal light shield. Overall pixel size is limited by the photodiode size. For example, with a three-phase system, the pixel width is three times the width of a single gate. For a fixed gate size, the active pixel width can be reduced with a bilinear readout (Figure 3-12b). Because resolution is inversely related to the detector-to-detector spacing (pixel pitch), the bilinear readout has twice the resolution. For fixed pixel size, the bilinear readout increases the effective charge transfer efficiency because the number of transfers is reduced by a factor of 2 (discussed in Section 3.4., *Charge Transfer Efficiency*).

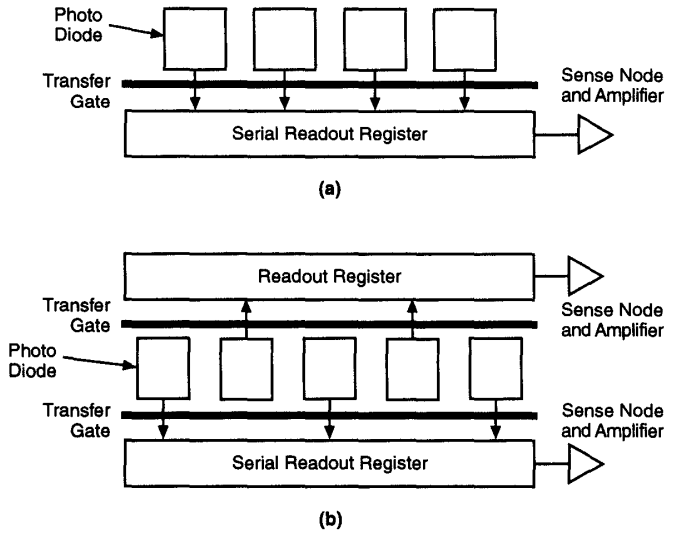


Figure 3-12. Linear array. (a) Simple structure and (b) bilinear readout. The transfer gate clock pulse allows charge to move from the photodiode into the CCD transfer register.

3.3.2. FULL FRAME ARRAYS

Figure 3-13 illustrates a full frame transfer (FFT) array. After integration, the image pixels are read out line-by-line through a serial register that then clocks its contents onto the output sense node (see Figure 3-4, page 50). All charge must be clocked out of the serial register before the next line can be transferred. In full frame arrays, the number of pixels is often based upon powers of 2 (e.g., 512 × 512, 1024 × 1024) to simplify memory mapping. Scientific arrays have square pixels and this simplifies image processing algorithms.

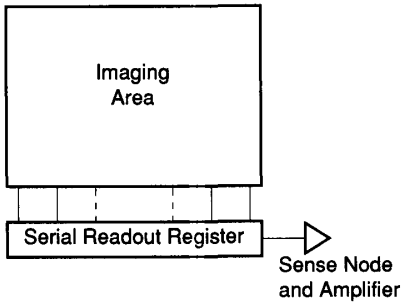


Figure 3-13. Full frame architecture. Because photogates are used, transfer gates are not essential.

During read out, the photosites are continually illuminated resulting in a smeared image. The smear will be in the direction of the charge transport in the imaging part of the array. A mechanical or external electronic shutter can be used to shield the array during readout to avoid smear. When using strobe lights to illuminate the image, no shutter is necessary if the transfer is between strobe flashes. If the image integration time is much longer than the readout time, then the smear may be considered insignificant. This situation often occurs with astronomical observations.

Data rates are limited by the amplifier bandwidth and, if present, the conversion capability of the analog-to-digital converter. To increase the effective readout rate, the array can be divided into subarrays that are read out simultaneously. In Figure 3-14, the array is divided into four subarrays. Because they are all read out simultaneously, the effective clock rate increases by a factor of 4. Software then reconstructs the original image. This is done in a video processor external to the CCD device where the serial data are decoded and reformatted.

Large area devices often allow the user to select one subarray for readout. Thus, the user can trade frame rate against image size. This allows the user to obtain high frame rates over the area of interest (discussed in Section 4.5., *Frame Rates*).

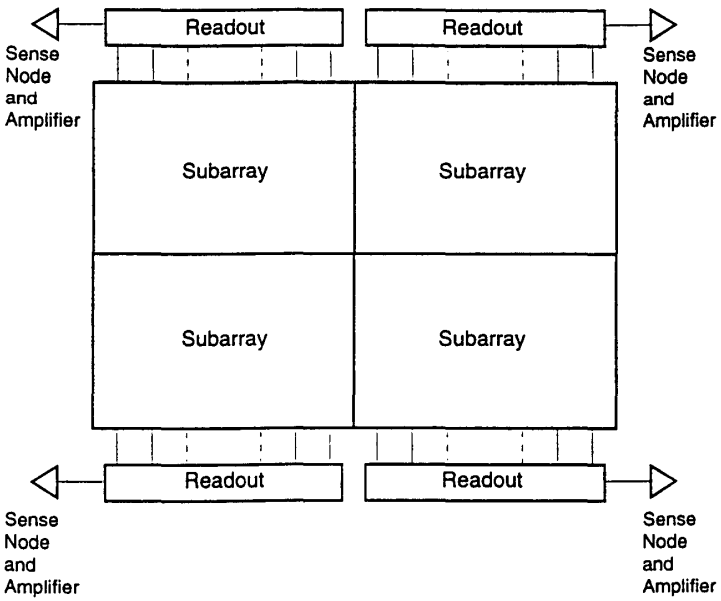


Figure 3-14. A large array divided into four subarrays. Each subarray is read out simultaneously to increase the effective data rate. Very large arrays may have up to 32 parallel outputs. The subimages are reassembled in electronics external to the device to create the full image.

3.3.3. FRAME TRANSFER

A frame transfer (FT) imager consists of two almost identical arrays, one devoted to image pixels and one for storage (Figure 3-15). The storage cells are identical in structure to the light sensitive cells but are covered with a metal light shield to prevent any light exposure. After the integration cycle, charge is transferred quickly from the light sensitive pixels to the storage cells. Transfer time to the shielded area depends upon the array size but is typically less than 500 μs. Smear is limited to the time it takes to transfer the image to the storage area. This is much less time than that required by a full frame device.

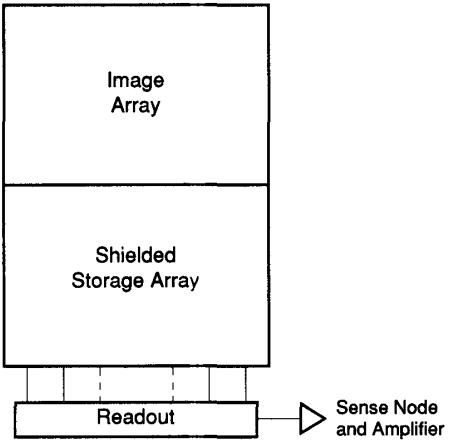


Figure 3-15. Basic architecture for a full frame CCD. Frame transfer devices can also have a split architecture similar to that shown in Figure 3-14.

The array may have a few dummy charge wells between the active and storage areas. These wells collect any anticipated light leakage and their data are ignored. Some arrays do not have the integral light shield. These arrays can either be operated in the full frame mode (e.g., 512 × 512) or in the frame transfer mode (e.g., 256 × 512). Here, it becomes the user's responsibility to design the light shield for the frame transfer mode.

☞

Example 3-1
SHIELD DESIGN

A camera manufacturer buys a full frame transfer array (512 × 512) but wants to use it in a frame transfer mode (256 × 512). An engineer designs an opaque shield that is placed 1 mm above the array. He is using f/5 optics. What is the usable size of the array if the pixels are 20 μm square and the shield edge is centered on the array?

When placed near the focal plane, the shield edge will create a shadow that is approximately the distance from the focal plane (d = 1 mm) divided by the f-number (F = 5). The full shadow is 200 μm or 10 pixels. One-half of the shadow falls on the active pixels and one-half illuminates the pixels under the shield. This results in a usable array size of 251 × 512. An alignment error of 0.5 mm adds an additional 25 pixels. If the alignment error covers active pixels, the active area is reduced to 226 × 512. If the error covers shielded pixels, the storage area is reduced to 226 × 512. Either way, the usable array is 226 × 512.

Example 3-2
USEFUL ARRAY SIZE

A consumer wants to operate the camera designed in Example 3-1 under relatively low illumination levels. He changes the lens to one with f-number of 1.25. Will he be satisfied with the imagery?

The lower f-number increases the shadow to 40 pixels. Assuming an alignment error of 0.5 mm, the minimum usable area size has dropped to 211 × 512 with a "soft edge" of 40 pixels. Acceptability of this image size depends upon the application.

3.3.4. INTERLINE TRANSFER

The interline transfer (IL) array consists of photodiodes separated by vertical transfer registers that are covered by an opaque metal shield (Figure 3-16). Although photogates could be used, photodiodes offer higher quantum efficiency (discussed in Section 4.1.1., *Spectral Response*). After integration, the charge generated by the photodiodes is transferred to the vertical CCD registers in about 1 μs and smear is minimized. The main advantage of interline transfer is that the transfer from the active sensors to the shielded storage is quick. There is no need to shutter the incoming light. The electronic shutter open time can be variable if an antibloom drain is present (discussed in Section 3.8., *Antibloom Drain*). The shields act like a Venetian blind that obscures half the information that is available in the scene. The area fill-factor may be as low as 20%. Because the detector area is only 20% of the pixel area, the output voltage is only 20% of a detector that would completely fill the pixel area. Microlenses increase the optical fill-factor (discussed in Section 3.11.1., *Microlenses*).

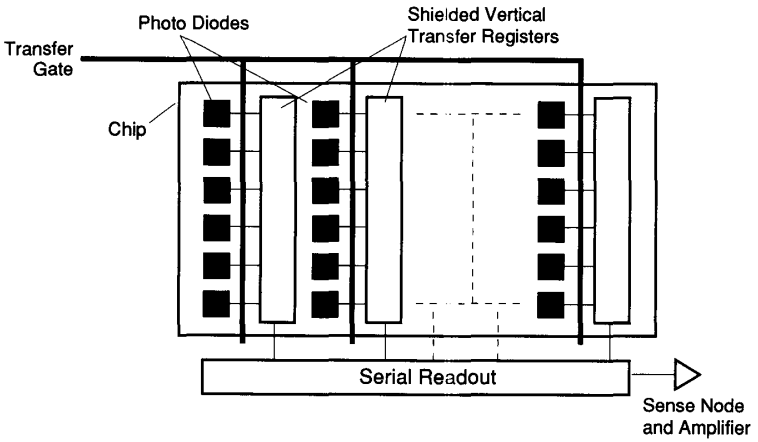


Figure 3-16. Interline transfer architecture. The charge is rapidly transferred to the interline transfer registers via the transfer gate. The registers may have three or four gates. Interline transfer devices can also have a split architecture similar to that shown in Figure 3-14 (page 61).

The combination of the capacitances associated with the photodiode and transfer gate provide a time constant that limits the speed of transfer to the vertical CCD register. As a result not all the charge is read out and some is left behind. This residual charge is added to the next frame and results in image lag. That is, with photodiodes, the remaining charge is drained in successive readouts causing an after-image. But hole accumulation devices can be completely emptied and no image lag occurs.

A fraction of the light can leak into the vertical registers. This effect is most pronounced when viewing an ambient scene that has a bright light in it. For professional television applications, the frame interline transfer (FIT) array was developed to achieve even lower smear values (Figure 3-17).

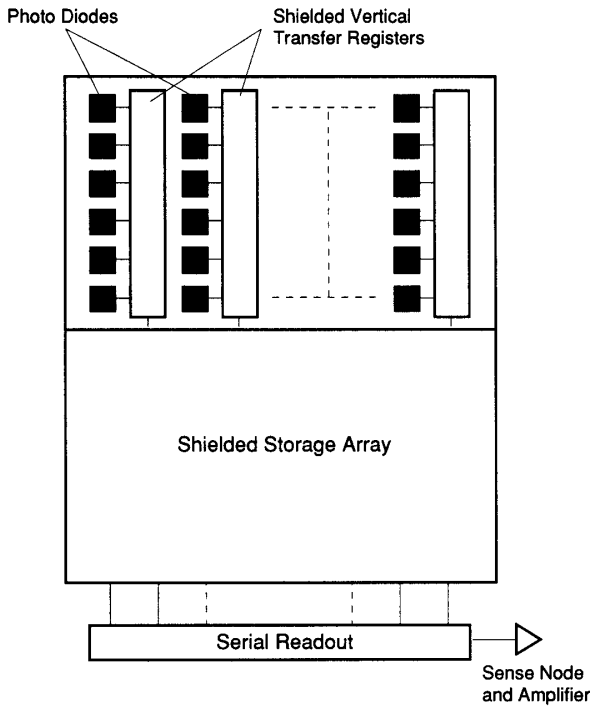


Figure 3-17. Frame interline transfer architecture. Both the vertical transfer registers and storage area are covered with an opaque mask to prevent light exposure. The transfer gates are not shown.

A variety of transfer register architectures are available. A register can operate with virtual, two, three, or four phases. The number of phases selected depends upon the application. Because interline devices are most often found in general imagery products, most transfer register designs are based upon standard video timing (discussed in Section 5.2.1., *Video Timing*). Figure 3-18 illustrates a four-phase transfer register that stores charge under two gates. With 2:1 interlace, both fields are collected simultaneously but are read out alternately. This is called frame integration. With EIA 170 (formerly called RS 170), each field is read every 1/60 s. Because the fields alternate, the maximum integration time is 1/30 s for each field.

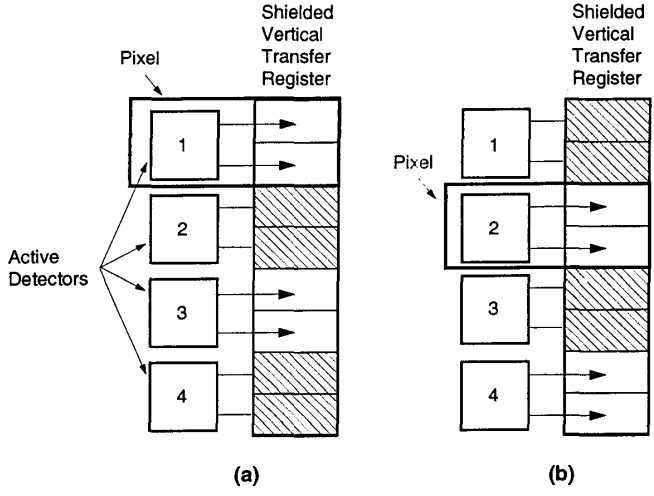


Figure 3-18. Detailed layout of the 2:1 interlaced array. (a) The odd field is clocked into the vertical transfer register and (b) the even field is transferred. The vertical transfer register has four gates and charge is stored under two wells. The pixel is defined by the detector center-to-center spacing and it includes the shielded vertical register area. The transfer gate is not shown. (After Reference 11).

Pseudo-interlacing (sometimes called field integration) is shown in Figure 3-19. Changing the gate voltage shifts the image centroid by one-half pixel in the vertical direction. This creates 50% overlap between the two fields. The pixels have twice the vertical extent of standard interline transfer devices and therefore have twice the sensitivity. An array that appears to have 240 elements in the vertical direction is clocked so that it creates 480 lines. However, this reduces the vertical MTF (discussed in Section 8.3.2., *Detector Array Nyquist Frequency*). With some devices, the pseudo-interlace device can also operate in a standard interlace mode.

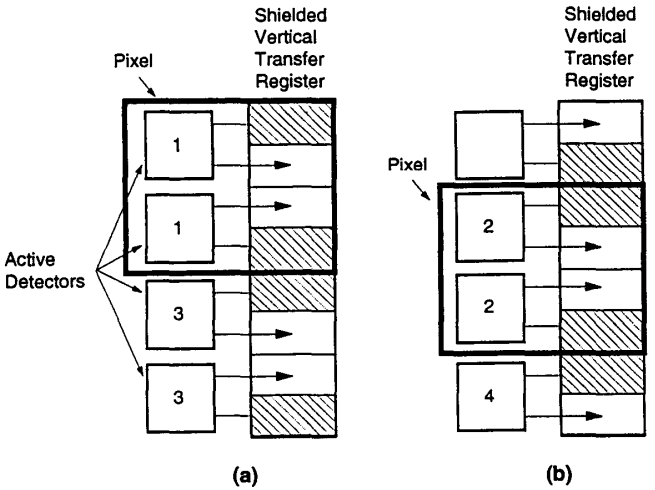


Figure 3-19. Pseudo-interlace. By collecting charge from alternating active detector sites, the pixel centroid is shifted by one-half pixel. (a) Odd field and (b) even field. The detector numbers are related to the "scan" pattern shown in Figure 3-20b. The transfer gate is not shown. (After Reference 11).

While any architecture can be used to read the charge, the final video format generally limits the selection. Although the charge is collected in rows and columns, the camera reformats the data into a serial stream consistent with the monitor requirements. This serial stream is related to a fictitious "scan" pattern. Figure 3-20 illustrates the video lines and field-of-view of a standard format imaging system. Figure 3-20a represents the "scan" pattern created by the interlace device shown in Figure 3-18 whereas Figure 3-20b represents the pseudo-interlace "scan" pattern of the device shown in Figure 3-19. Although the dissected scene areas overlap in Figure 3-20b, the video timing is consistent with the video format.

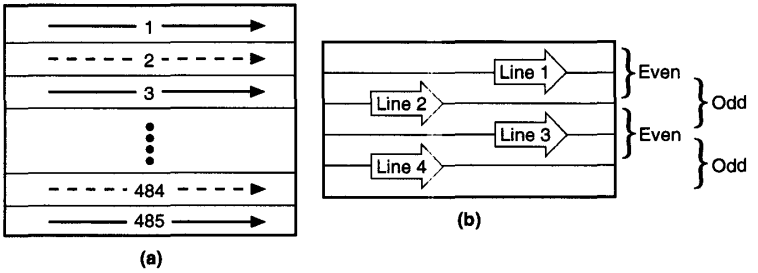


Figure 3-20. Fictitious "scan" patterns. The scan patterns help to visualize how the scene is dissected by the array architecture. (a) Interlace "scanning" achieved with the array shown in Figure 3-18. The solid lines are the odd field and the dashed lines are the even field. (b) Pseudo-interlace "scanning" provided by the array shown in Figure 3-19.

3.3.5. PROGRESSIVE SCAN

CCD arrays do not scan the image, but it is convenient to represent the output as if it were scanned. Progressive scan simply means the noninterlaced or sequential line-by-line scanning of the image. This is important to machine vision because it supplies accurate timing and has a simple format. Any application that requires digitization and a computer interface will probably perform better with progressively scanned imagery.

The primary advantage of progressive scan is that the entire image is captured at one instant of time. In contrast, interlaced systems collect fields sequentially in time. Any vertical image movement from field-to-field smears the interlaced image in a complex fashion. Horizontal movement serrates vertical lines. With excessive movement, the two fields will have images that are displaced one from the other (Figure 3-21).

If a strobe light is used to stop motion, the image will appear only in the field that was active during the light pulse. Because of image motion effects, only one field of data can be used for image processing. This reduces the vertical resolution by 50% and will increase vertical image aliasing.

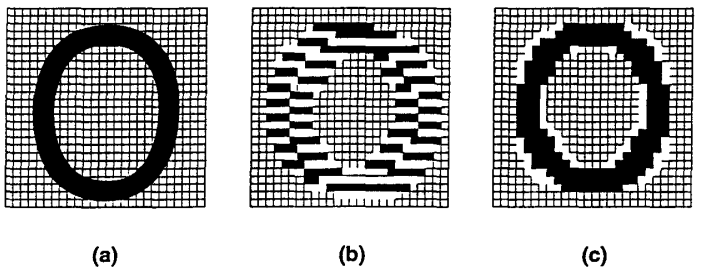


Figure 3-21. Effects of movement. (a) Character moving horizontally, (b) serration that occurs with an interlace system, and (c) output of a progressively scanned array. An image can still be smeared with progressive scanning but it won't be serrated.

Because an entire frame is captured, progressive scan devices do not suffer from same image motion effects seen with interlaced systems. Thus, progressive scan cameras are said to have improved vertical resolution over interlaced systems. Full frame and frame transfer devices inherently provide progressive scan readout. Although early interline transfer architecture was designed for interlace scanning, some current devices also provide progressive scan output.

Although the imagery is collected in a sequential method, camera electronics convert progressively scanned data into a standard video format (discussed in Section 5.2.2., *Broadcast/Non-Broadcast Design*). This allows the progressively scanned camera output to be viewed on any standard interlaced display. For scientific and industrial applications, the progressive output is captured by a frame grabber. Once in the computer, algorithms automatically reformat the data into the format required by the monitor. The camera pixel layout and array timing do not need to fit a specified format for these applications.

3.3.6. TIME DELAY and INTEGRATION

Time delay and integration (TDI) is analogous to taking multiple exposures of the same object and adding them[12]. The addition takes place automatically in the charge well and the array timing produces the multiple images. Figure 3-22 illustrates a typical TDI application. A simple camera lens inverts the image so the image moves in the opposite direction from the object. As the image is swept across the array, the charge packets are clocked at the same rate. The relative motion between the image and the target can be achieved in many ways. In airborne reconnaissance, the forward motion of the aircraft provides the motion with respect to the ground. With objects on a conveyor belt, the camera is stationary and the objects move. In a document flat bed scanner, the document is stationary but either the camera moves or a scan mirror moves.

Figure 3-23 illustrates four detectors operating in TDI mode. The charge increases linearly with the number of detectors in TDI. Shot noise also increases, but as the square root of the number of TDI elements, N_{TDI}. Theoretically, this results in an SNR improvement of $\sqrt{N_{TDI}}$. The well capacity limits the maximum number of TDI elements that can be used.

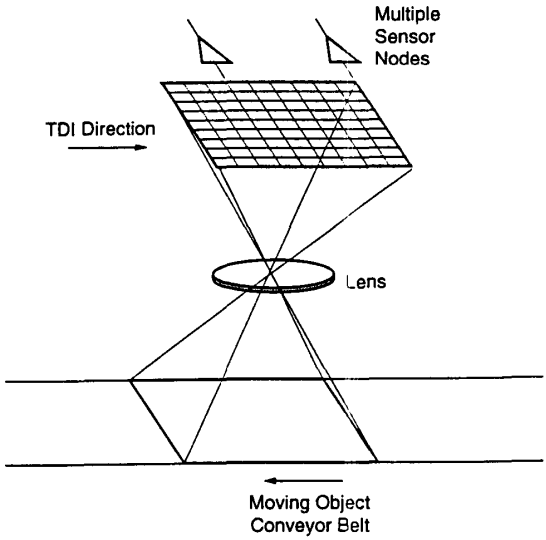


Figure 3-22. Typical TDI operation. The pixel clock rate must be matched to the image velocity.

For this concept to work, the charge packet must always be in synch with the moving image. If there is a mismatch between the image scan velocity and pixel clock rate, the output is smeared and this adversely affects the in-scan MTF (discussed in Section 10.3.3., *TDI*). The accuracy to which the image velocity is known limits the number of useful TDI stages.

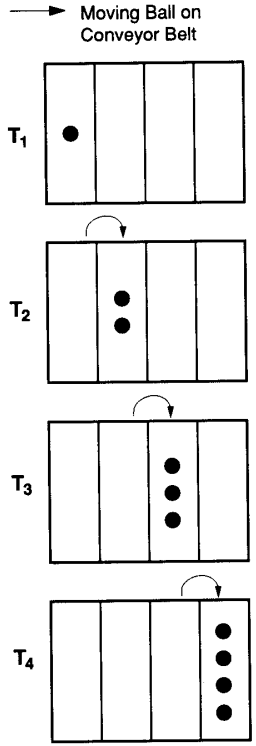


Figure 3-23. TDI concept. At time T_1, the image is focused onto the first detector element and creates a charge packet. At T_2 the image has moved to the next detector. Simultaneously, the charge packet is clocked down to the next pixel site. Here, the image creates additional charge that is added to the stored packet (created by the first detector). After four stages, the signal increases fourfold.

Although Figure 3-22 illustrates two readout registers, an array may have multiple readout registers. Array sizes may be as large as 2048 by 32 TDI elements. Additional readouts are required to avoid charge well saturation. The multiple readout outputs may be summed in electronics external to the TDI device or may be placed in a horizontal serial register for summing (Figure 3-24).

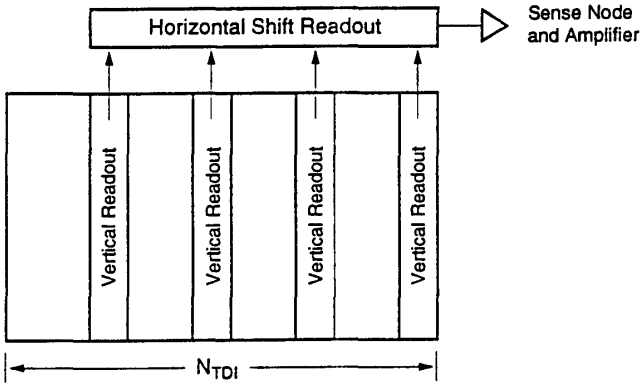


Figure 3-24. The array height must match the image height. To avoid saturation, multiple vertical register readouts may be used.

The vertical readout register must clock out the data before the next line is loaded into the register. The amplifier bandwidth or analog-to-digital conversion time limits the pixel rate. For example, if the amplifier can support 10^7 pixels/s and the register serves 500 vertical pixels, then the maximum (in-scan) line rate is 20,000 lines/s. If the amplifier can only handle 2×10^6 pixels/s and the line rate is maintained at 20,000 lines/s, then the vertical transfer register must be divided into five registers - each connected to 100 pixels. These registers and associated outputs operate in parallel (Figure 3-25).

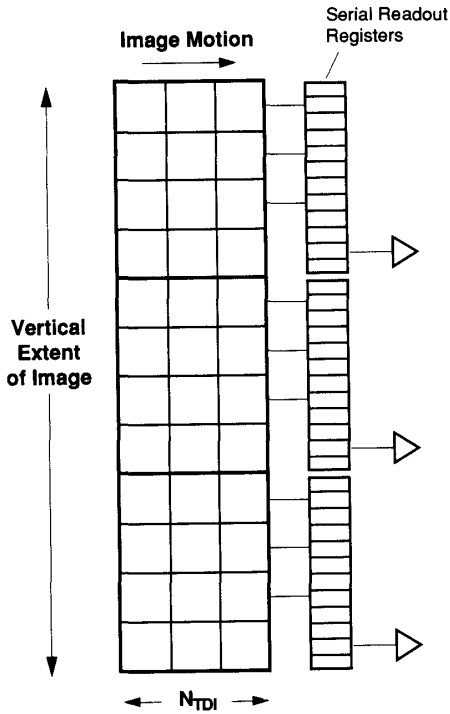


Figure 3-25. The vertical direction may be divided into subarrays to increase the line rate. The transfer register typically is a three-phase device.

TDI improves responsivity nonuniformity (discussed in Section 4.2.6., *Pattern Noise*). Variations in individual detector responsivities in the TDI direction are averaged by $1/N_{TDI}$. Multiple outputs are used to increase the effective pixel rate. Each output is amplified by a separate amplifier that has unique gain characteristics. If the gains are poorly matched, the image serviced by one amplifier will have a higher contrast than the remaining image. In TDI devices, streaking occurs when the average responsivity changes from column-to-column.[13] Careful selection of arrays will minimize these nonuniformities. They can also be minimized though processing techniques employed external to the device.

Example 3-3
TDI LINE RATE

The pixels in a TDI are 15 μm square. The object is moving at 2 m/s and the object detail of interest is 150 μm. What is the lens magnification and what is the pixel clock rate?

Each pixel must view 150 μm on the target. Because the pixel size is 15 μm, the lens magnification is $M_{OPTICS} = 15/150 = 1/10$. The TDI pixel clock rate is (2 m/s)/150 μm = 133 Khz.

3.4. CHARGE TRANSFER EFFICIENCY

The early CCDs transferred electrons in channels that were at the surface. Surface defects interfere with the transfer and create poor charge transfer efficiency. These were SCCD or surface channel CCDs. By burying the channel (BCCD), the charge was no longer affected by surface defects and this improved the CTE significantly. Nearly all CCDs use buried channel technology today.

As the charge packet moves from storage site to storage site, a few electrons are left behind. The ability to transfer all the charge is given by the CTE. As the number of transfers increases, the charge transfer efficiency must also increase to maintain signal strength.

The number of transfers across the full array is the number of pixels multiplied by the number of gates. For example, for an array that is 1000 × 1000, the charge packet farthest from the sense node must travel 2000 pixels. If four gates (four-phase device) are used to minimize cross talk, the farthest charge packet must pass through 8000 wells (neglecting isolation pixels). The net efficiency varies with the target location. If the target is in the center of the array, then it only experiences one-half of the maximum number of transfers on the array. If the target is at the leading edge, it is read out immediately with virtually no loss of information. The loss of information is described by the CTE MTF (discussed in Section 10.3.2., *Charge Transfer Efficiency*).

If many transfers exist (e.g., 8000) the transfer efficiency must be high (e.g., 0.999995) so that the net efficiency is reasonable ($0.999995^{8000} = 0.96$). Arrays for consumer applications typically have charge transfer efficiencies greater than 0.9999 and CTEs approach 0.999999 for scientific grade devices.

A fraction of the charge, as specified by the charge *in*efficiency, is left behind. After the first transfer, ϵ (ϵ is the CTE) remains in the leading well and $(1-\epsilon)$ enters the first trailing well. After the second transfer, the leading well transfers ϵ of the previous transfer and loses $(1-\epsilon)$ to the trailing well. The trailing well gains the loss from the first well and adds it to the charge transferred into the trailing well in the last go-around. Charge is not lost, just rearranged. The fractional values follow a binomial probability distribution:

$$P_R = \binom{N}{R} \epsilon^R (1 - \epsilon)^{N-R} , \tag{3-2}$$

where N is the number of transfers. For the first well, R = N. For the second well, R = N - 1 and so on.

Figure 3-26 illustrates the output when a large number of pixels are illuminated. The leading edge loses charge and decreases in value. For pixels near the end of the pulse, the charge lost to trailing wells is replaced by the charge acquired from leading wells. Here, the signal strength is maintained. Finally, charge is lost to wells after the signal and the pulse width is stretched.

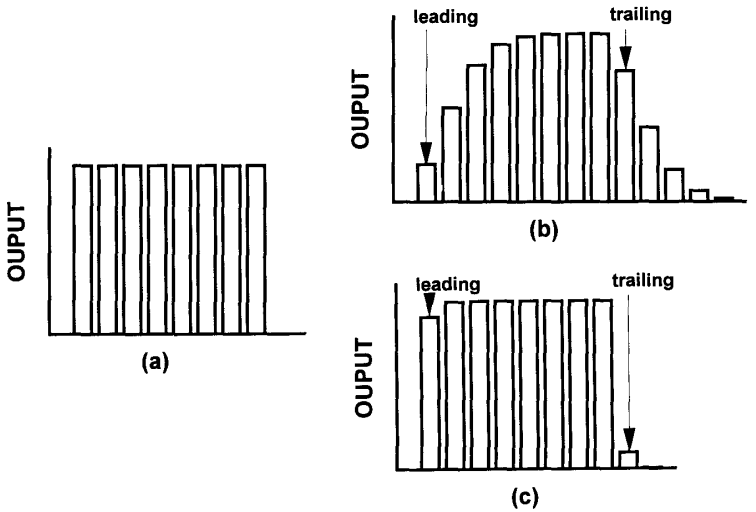


Figure 3-26. Input and output signals after many transfers. (a) Initially generated charge, (b) charge distribution with low CTE, and (c) charge distribution with high CTE.

Consider a very small spot of light that illuminates just one pixel and creates n_o electrons. A well, N transfers away, will contain n_N electrons. The fractional amount of charge left is

$$\frac{n_N}{n_o} = \epsilon^N . \tag{3-3}$$

The fractional amount of charge appearing in the first trailing well after N transfers is

$$\frac{n_{(first\ trailing\ well)-N}}{n_o} = N\epsilon^{N-1}(1-\epsilon) . \tag{3-4}$$

The fractional signal strength, n_N/n_o, is plotted in Figure 3-27. The first trailing well fractional strength, $n_{first\ trailing\ well\text{-}N}/n_o$, is plotted in Figure 3-28. Most of the electrons lost from the first well appear in the second and third wells. For high CTE arrays, a negligible amount appears in successive wells.

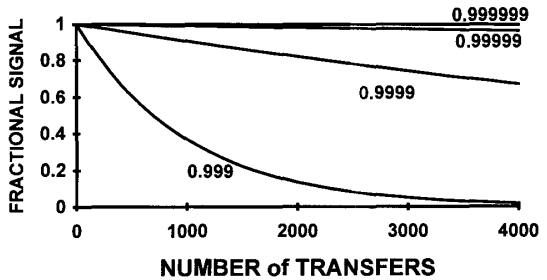


Figure 3-27. Fractional signal as a function of transfer efficiency and number of transfers.

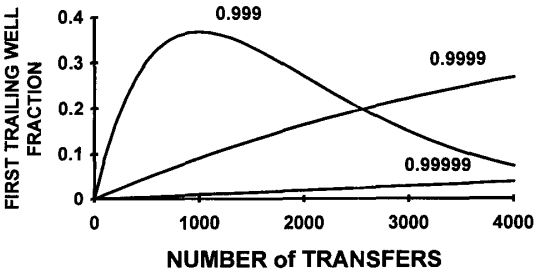


Figure 3-28. Fractional signal for the first trailing charge well. With low CTEs, as the number of transfers increases, more charge spills into the second trailing well. The charge in the first trailing well decreases.

CTE is not a constant value but depends upon the charge packet size. For very low charge packets, the CTE decreases due to surface state interactions. This is particularly bothersome when the signal-to-noise ratio is low and the array is very large (typical of astronomical applications). The CTE also decreases near saturation due to charge spill effects.

The fraction of charge left behind also depends upon the clocking frequency. At high frequencies, efficiency is limited by the electron mobility rate. That is, it takes time for the electrons to move from one storage site to the next. Thus, there is a tradeoff between frame rate (dictated by clock frequency) and image quality (affected by CTE). Using smaller pixels partially overcomes the limitation imposed by the electron mobility rate. However, a smaller pixel has a smaller charge well capacity.

If the surface states significantly affect CTE, then the fat zero technique helps. Fat zero puts a known amount of electrons into the well. These electrons will fill the interface states and thereby improve the CTE for the photoelectrons. Because the fat zero electrons partially fill the charge well, the remaining well capacity for the photoelectrons is reduced.

CTE effects in large linear arrays can be minimized by using alternate readouts. In Figure 3-29, the alternate readouts reduce the number of transfers by a factor of 2. Adjacent pixels experience the same number of transfers so that the signal is preserved at any particular location.

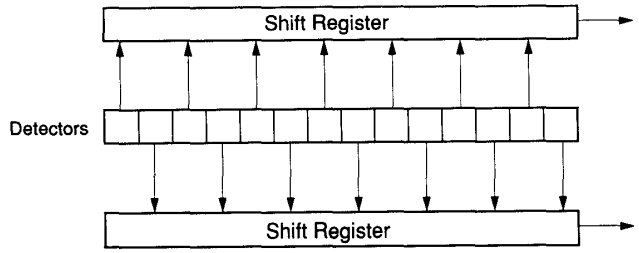


Figure 3-29. For long arrays, the MTF can be increased by alternating the readouts or by reading out subarrays (see Figure 3-14, page 61).

There are three methods to measure CTE: current, optical, and x-ray injection. Because most devices do not have current injection terminals, the optical and x-ray injection methods are used more often. With the optical technique, either a row of detectors or a single detector is illuminated. The ratio of the output of the first illuminated pixel to the average value of trailing pixels is measured. The CTE is found using Equation 3-3. The optical technique requires a spot[14] whose diameter is less than a pixel width. This can only be achieved with carefully designed optics that are critically aligned. But x-rays[1] generate a known number of electrons in a localized area that is much smaller than a pixel. The x-ray technique probably provides the highest accuracy.

3.5. CHARGE CONVERSION (OUTPUT STRUCTURE)

Charge is converted to a voltage by a floating diode or floating diffusion. Figure 3-30 illustrates the last two gates of a four-phase system. The low voltage gate allows charge transfer only when gate 4 goes to zero. The diode, acting as a capacitor, is precharged at a reference level. The capacitance, or sense node, is partially discharged by the amount of charge transferred. The difference in voltage between the final status of the diode and its precharged value is linearly proportional to the number of electrons, n_e. The signal voltage after the source follower is

$$V_{SIGNAL} = V_{RESET} - V_{OUT} = n_e \frac{Gq}{C} . \qquad (3\text{-}5)$$

The gain, G, of a source follower amplifier is approximately 1 and q is the electronic charge (1.6×10^{-19} coul). The charge conversion is q/C. The output gain conversion or OGC is Gq/C. It typically ranges from $0.1\,\mu V/e^-$ to $10\,\mu V/e^-$. The signal is then amplified and processed by electronics external to the CCD sensor.

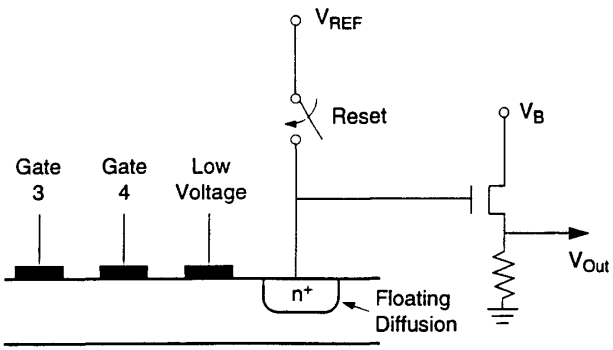


Figure 3-30. Output structure. The charge on the floating diffusion is converted to a voltage by the diffusion capacitance. The low voltage gate, which operates in a manner similar to a virtual well, allows charge transfer when the voltage on gate 4 is zero (see Figure 3-11, page 57).

3.6. DARK CURRENT

The CCD output is proportional to the exposure, $L_q(\lambda)t_{INT}$. The output can be increased by increasing the integration time and long integration times are generally used for low light level operation. However, this approach is ultimately limited by dark current leakage that is integrated along with the photocurrent. With a large pixel (24 μm square), a dark current density of 1000 pA/cm^2 produces 36,000 electrons/pixel/s. If the device has a well capacity of 360,000 electrons, the well fills in 10 s. The dark current is only appreciable when t_{INT} is long. It is not usually a problem for consumer applications but is a concern for scientific applications.

A critical design parameter is dark current noise reduction. There are three main sources of dark current: (1) thermal generation in the depletion region, (2) thermal generation and diffusion in the neutral bulk material, and (3) thermal generation due to surface states. Dark current densities vary significantly among manufacturers with values ranging from 0.1 nA/cm^2 to 10 nA/cm^2 in silicon CCDs. Dark current due to thermally generated electrons can be reduced by cooling the device. Surface state dark current is minimized with multi-phase pinning[15,16] (MPP).

The number of thermally generated dark current electrons is

$$n_{DARK} = \frac{J_D A_D t_{INT}}{q} , \tag{3-6}$$

where J_D is the dark current density. For thermal generation in the depletion region, It is proportional to

$$J_D \propto T^2 e^{-\frac{(E_G - E_T)}{kT}} , \tag{3-7}$$

where T is the absolute temperature, E_G is the band gap, and E_T is the impurity energy gap. With TDI, the pixels in a TDI column are summed and the dark current increases:

$$n_{DARK} = \frac{J_D A_D t_{INT} N_{TDI}}{q} . \tag{3-8}$$

In principle, dark current density can be made negligible with sufficient cooling. The dark current density decreases approximately twofold for every 8 to 9°C drop in array temperature from ambient. Conversely, the dark current density increases by a factor of 2 for every 8 to 9°C increase in array temperature. Because the temperature is in the exponent of Equation 3-7, the doubling temperature decreases as the array temperature decreases. At low temperatures, a 1°C change in temperature changes the dark current by a significant amount. This means that if cooling is desired, the temperature must be maintained with high precision. Table 3-1 provides the dark current for Texas Instruments chip TC271 using experimentally[17] obtained values of $E_G - E_T \approx 0.58$ eV and $J_D = 100\,pA/cm^2$ at 27°C.

The most common cooler is a thermoelectric cooler (TEC or TE cooler). Thermoelectric coolers are Peltier devices driven by an electric current that pumps heat from the CCD to a heat sink. The heat sink is cooled by passive air, forced air, or forced liquid (usually water). The final array temperature depends upon the amount of heat generated in the array, the cooling capacity of the thermoelectric device, and the temperature of the TEC heat sink. Peltier devices can cool to about -60°C. Thermoelectric coolers can be integrated into the same package with the array (Figure 3-31). Cooling probably is only worthwhile for scientific applications. However, many devices contain TE coolers to stabilize the temperature so that the dark current does not vary with ambient temperature changes.

Table 3-1
DARK CURRENT VALUES for TI TC271
(Normalized to 100 pA/cm^2 at 27°C)

TEMPERATURE (°C)	DARK CURRENT (pA/cm^2)	PERCENT CHANGE for 1° C	DOUBLING TEMPERATURE
60	1136	6.9	10.8
40	276	7.7	9.5
20	55.8	8.9	8.4
0	9.02	10.2	7.2
-20	1.11	11.9	6.2
-40	9.58×10^{-2}	14.1	5.3
-60	5.33×10^{-3}	17.0	4.4
-80	1.66×10^{-4}	20.9	3.6

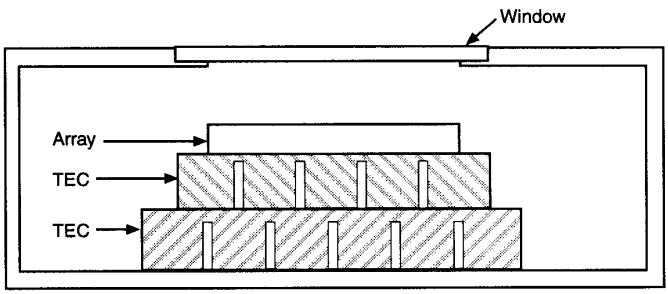


Figure 3-31. Cross section of a two stage TEC integrated into a package.

Although liquid nitrogen (LN, LN2, or LN_2) is at -200°C, the typical operating temperature is -120°C to -60°C because the CTE and quantum efficiency drop at very low temperatures. Condensation is a problem and the arrays must be placed in an evacuated chamber or a chamber that is filled with a dry atmosphere. Dark current can approach[18] 3.5 electrons/pixel/s at -60°C and 0.02 electrons/pixel/h at -120°C. Various cooling techniques and fabrication concepts are provided in Reference 19.

Multi-pinned phasing technology CCDs can operate at room temperature with reduced dark current.[15,16] MPP devices have dark current density levels as low as 2 pA/cm^2 with a high value of 1 nA/cm^2. Typically, MPP devices have charge wells that are two to three times smaller than conventional devices and these devices are more likely to saturate. Some devices operate in either mode. MPP is the most common architecture used in scientific and medical applications.

The output amplifier constantly dissipates power. This results in local heating of the silicon chip. Because dark current is dependent upon temperature, a small variation in the device temperature profile produces a corresponding profile in the dark current. To minimize this effect, the output amplifier is often separated by several isolation pixels to move the amplifier away from the active sensors.

Example 3-4
DARK CURRENT

A CCD array has a well size of 150,000 electrons, dark current density of 10 pA/cm^2, and detector size 8 μm square. The advertised dynamic range is 80 dB. Is the dark current noise significant? Assume that the integration time is 1/60 s (16.67 ms).

Using Equation 3-6, the number of dark electrons is

$$n_{DARK} = \frac{\left(10\times10^{-9}\frac{A}{cm^2}\right)(0.0167s)(8\times10^{-4}cm)^2}{1.6\times10^{-19}\frac{coul}{electron}} = 668\ e^- . \qquad (3\text{-}9)$$

Assuming Poisson statistics, the dark current shot noise is $\sqrt{668}$ or 26 electrons rms. The dynamic range is the charge well capacity divided by the noise floor. With a dynamic range of 10,000:1, the noise floor is 15 electrons. Because the noise variances add, the total noise is:

$$\langle n_{SYS} \rangle = \sqrt{26^2 + 15^2} = 30\ \ e^-\ rms . \qquad (3\text{-}10)$$

Therefore, the dark current will affect the measured SNR at low levels. The dark current will saturate the well in 3.75 s.

Example 3-5
ARRAY COOLING

The dark current density decreases by a factor of 2 for every 8°C drop in temperature. What should the array temperature be so that the dark current noise is 5 electrons rms for the array described in Example 3-4?

The dark current noise must be reduced by a factor of 5/26 = 0.192 and the dark current must be reduced by $(0.192)^2$ or 0.0369. The doubling factor approximately provides

$$\frac{n_{DARK\text{-}COOL}}{n_{DARK\text{-}AMBIENT}} = 2^{\left(-\frac{T_{AMBIENT} - T_{COOL}}{8}\right)}. \qquad (3\text{-}11)$$

Then, $T_{AMBIENT} - T_{COOL} = 38.1°C$. That is, the array temperature must be 38.1°C below the temperature at which the dark current density was originally measured. Because the dark current is reduced by 0.0369, the time to saturation increased by 1/0/0369 or 27.1 times to 102 s.

Although this varies by manufacturer, many use $T_{AMBIENT} = 20°C$. Note that this is an approximation only and Equation 3-7 should be used to determine the required temperature. In reality, the user does not precisely select the temperature but, if possible, cools the array well below the desired temperature. The lowest possible temperature is limited by the cooling device.

3.7. DARK PIXELS

Many arrays have "extra" photosites at the end of the array. These shielded photosites are used to establish a reference dark current level (or dark signal). The number of dark elements varies with device and manufacturer and ranges from a few to 25. The average value of the dark current pixels is subtracted from the active pixels leaving only photogenerated signal. For example, if the output of the dark pixels is 5 mV, then 5 mV is subtracted from the signal level of each active pixel.

Figure 3-32a illustrates the output from a single charge well. Light leakage may occur at the edge of the shield (see Example 3-1, page 62) and partially fill a few wells. Because of this light leakage, a few pixels (called dummy or

isolation pixels) are added to the array. Figure 3-32b illustrates the output of three dark reference detectors, three isolation detectors, and three active pixels. The average value of the three dark pixels is subtracted from every active pixel value.

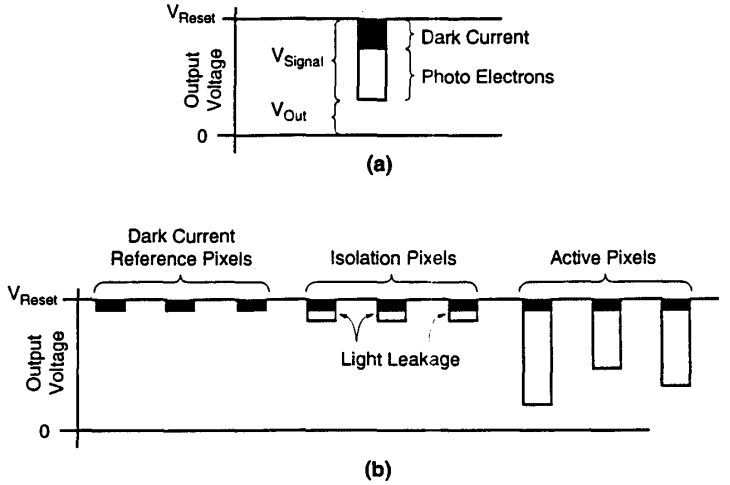


Figure 3-32. Typical output. (a) Single pixel and (b) active line with isolation and dark pixels. The varying voltage on the active pixels is proportional to the scene intensity. The dark pixels are covered with an opaque mask. Light can leak around the mask and partially fill the isolation pixel wells.

While this process is satisfactory for general imagery and machine vision use, it may be unacceptable for scientific applications. Dark pixels have slightly different dark current than the active pixels. Furthermore, the dark current value varies from pixel to pixel. This variation appears as fixed pattern noise (discussed in Section 4.2.6., *Pattern Noise*). Removing the average value does not remove the variability.

In critical scientific applications, only the dark value from a pixel may be removed from that pixel. That is, the entire array is covered (mechanical shutter in place) and the individual pixel values are stored in a matrix. Then the shutter is removed and the stored values are subtracted, pixel by pixel, as a matrix subtraction. This method ensures that the precise dark voltage is subtracted from each pixel. This, of course, increases computational complexity. Removal of dark current values allows maximum use of the analog-to-digital converter

dynamic range. This method cannot remove the dark current shot noise. Dark current shot noise is reduced by cooling.

Additional isolation pixels may exist between the readout register and the sense node amplifier. These pixels reduce any possible interaction between the amplifier temperature and first column pixel dark current. They also reduce amplifier noise interactions.

3.8. ANTIBLOOM DRAIN

When a well fills, charge spills over into adjacent pixels in the same column resulting in an undesirable overload effect called blooming. Channel stops prevent spill over onto adjacent columns. Overflow drains or antibloom drains prevent spill over. The drain can be attached to every pixel or may only operate on a column of pixels. Any photoelectron that is swept into the drain is instantly removed. Thus, the location of the drain is important.

If the drain is buried, the long wavelength quantum efficiency is reduced because long wavelength charge is generated below the wells and near the buried drain. A buried drain is also called a vertical overflow drain (VOD). Here, vertical refers to the *depth* of the device whereas the horizontal and vertical *directions* define the pixel size. If the drain is next to the detector, it is a lateral antibloom drain (Figure 3-33). This takes up real estate and lowers the fill-factor.

Figure 3-34 illustrates the potential well created by a lateral drain. When the charge well fills over the drain barrier, charge flows into the drain. The barrier height can be controlled by an external voltage so that variable integration times are possible. Here, the antibloom drain is used for exposure control. The exposure control can work at any speed. The drain, acting as an electronic shutter, may be used synchronously or asynchronously. Asynchronous shuttering is important for use in processes where the motion of objects does not match the array frame timing or in high light level applications. The charge well is drained, reset, and then photoelectrons are integrated for the desired time.

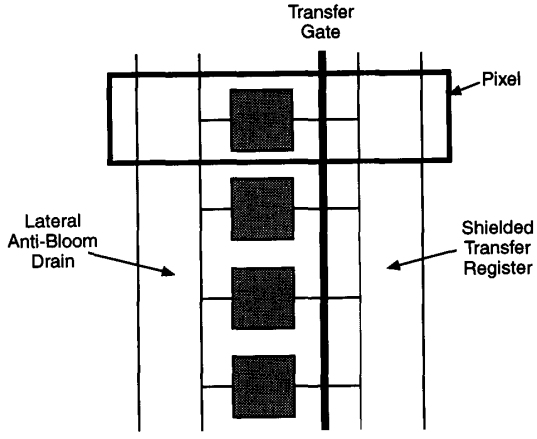


Figure 3-33. Lateral antibloom drain. The transfer gate controls the desired integration time.

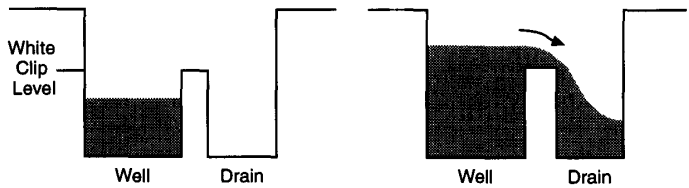


Figure 3-34. Lateral overflow drain. When the number of stored electrons exceeds the saddlepoint, the excess electrons flow into the drain

In a perfect system, the output is linearly related to the input up to the antibloom drain limit. In real arrays, because of imperfect drain operation, a knee is created[20] (Figure 3-35). The response above the knee is called knee slope or knee gain. While scientific cameras operate in a linear fashion, there is no requirement to do so with consumer cameras. The knee is a selling feature in the consumer marketplace. The advantage of the knee is that it provides some

response above the white clipping level. Each CCD manufacturer claims various competitive methods of handling wide scene contrasts. These methods usually involve manipulating the drain potential.

Interline transfer devices may have the drain at the top of a vertical shift register. The excess charge cannot be drained while image information is in the vertical shift registers (Figure 3-36). This limits the range of integration times available. Figure 3-37 illustrates the variable integration afforded by the drain.

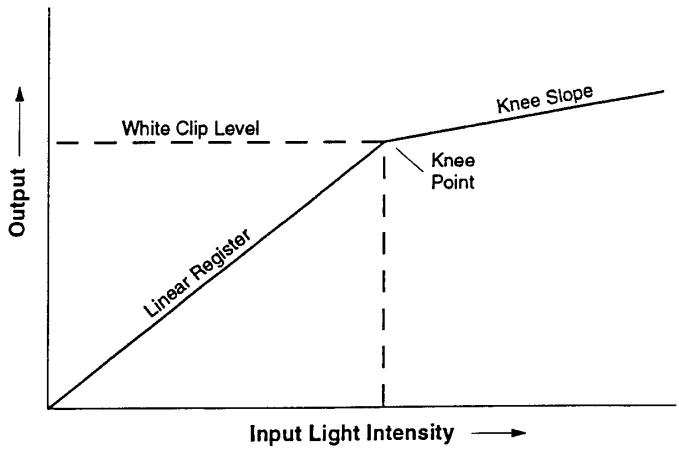


Figure 3-35. Characteristic knee created by an overflow drain.

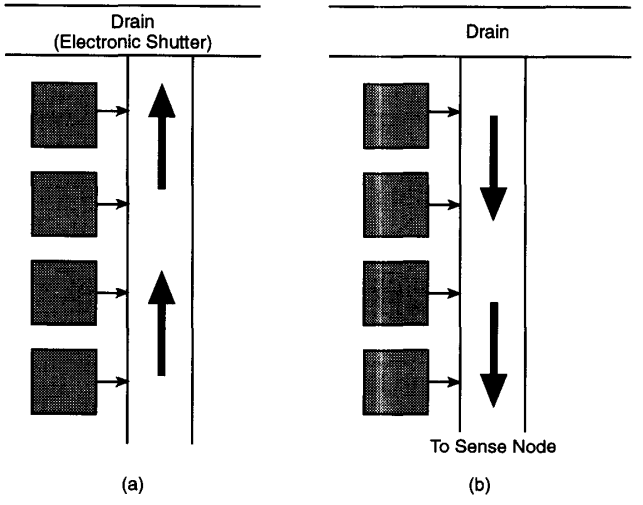


Figure 3-36. Interline transfer device where the drain is at the top of the vertical shift register. (a) Unwanted charge is sent to the drain and (b) desired charge is read out.

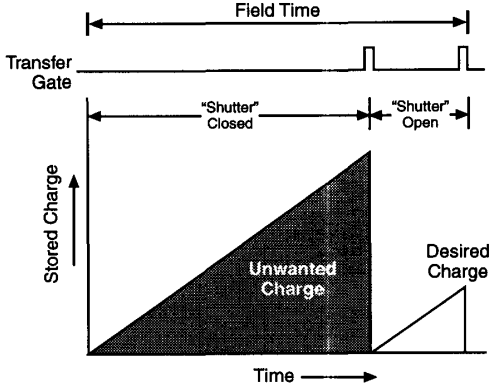


Figure 3-37. By judiciously activating the transfer gate, any integration time can be achieved. A variable integration time is equivalent to a variable electronic shutter.

3.9. CID

A charge injection device consists of two overlapping MOS capacitors sharing the same row and column electrode. Figure 3-38 illustrates the pixel architecture. Physically, the capacitors are orthogonal. Charge is stored under the horizontally or vertically oriented capacitor. The nearly contiguous pixel layout provides a fill-factor of 80% or greater.

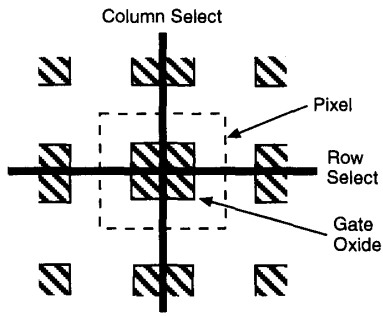


Figure 3-38. CID pixel architecture.

Although the capacitors are physically orthogonal, it is easier to understand their operation by placing them side-by-side. Figure 3-39 illustrates a functional diagram of an array and Figure 3-40 illustrates the pixel operation. In Figure 3-40a, a large voltage is applied to the columns and photogenerated carriers (usually holes) are stored under the column gate. If the column voltage is brought to zero, the charge transfers to the row gate (Figure 3-40b). The change in charge causes a change in the row gate potential which is then amplified and outputted. If V_1 is reapplied to the columns, the charge transfers back to the column gate. This is non-destructive readout and no charge is lost. Now additional charge integration is possible. Reset occurs by momentarily setting the row and column electrodes to ground. This *injects* the charge into the substrate (Figure 3-40c).

The conversion of charge into voltage depends upon the pixel, readout line, and amplifier capacitance. The capacitance is high because all the pixels on a given row are tied in parallel. Therefore, when compared to a CCD, the charge conversion is small, yielding a small signal-to-noise ratio. Because the readout is nondestructive, it can be repeated numerous times. The multiple reads are averaged together to improve the signal-to-noise ratio.

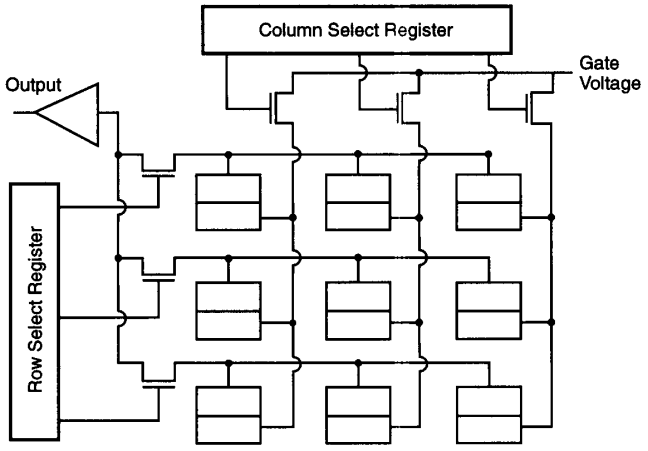


Figure 3-39. Functional layout of a 3 × 3 CID array. The row and column select registers are also called decoders. The select registers and readout electronics can be fabricated with CMOS technology.

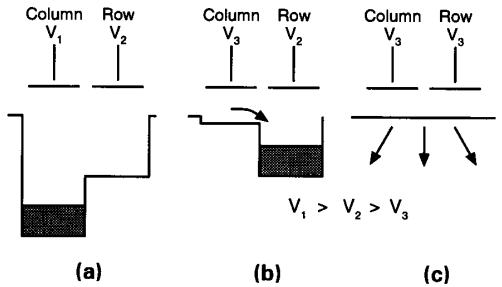


Figure 3-40. CID pixel operation. (a) Integration, (b) readout, and (c) injection.

Because each pixel sees a different capacitance, CIDs tend to have higher pattern noise compared to CCDs. However, off-chip algorithms can reduce the amount of pattern noise. With no charge transfer, CIDs are not sensitive to CTE effects. Without multiple gates, CIDs have larger well capacities than comparably sized CCDs. CIDs inherently have antibloom capability. Because charge is limited to a single pixel, it cannot overflow into neighboring pixels. Perhaps the greatest advantage of CIDs is random access to any pixel or pixel cluster. Subframes and binned pixels can be read out at high frame rates.

3.10. CMOS

The main advantage of CMOS imagers is that they are compatible with mainstream silicon chip technology. Because transistors are made by this technology, on-chip processing is possible.[21] An active pixel sensor (APS) has one or more active transistors integrated into the pixel (Figure 3-41).

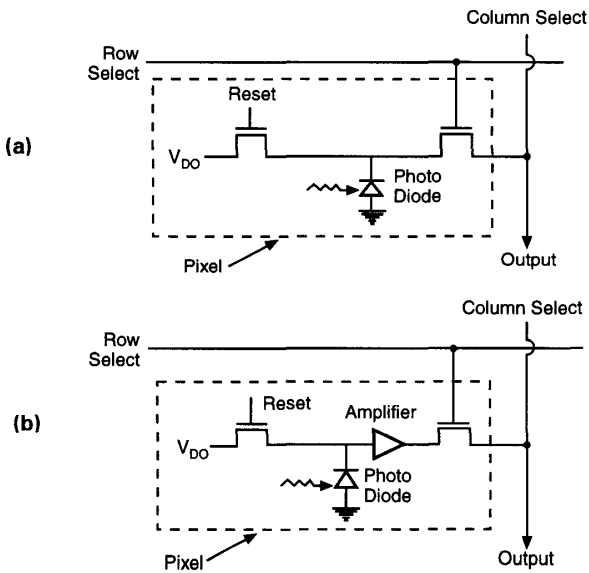


Figure 3-41. (a) Passive pixel device and (b) active pixel sensor. CCDs and CIDs are regarded as passive pixel sensors. CIDs use photogates, CCDs use either photogates or photodiodes, and CMOS devices typically use photodiodes.

With the APS approach, highly integrated image sensors are possible.[22] By placing processing on the chip, a CMOS camera is typically physically smaller than a CCD camera which requires clocks, image reformatting, and signal processing in separate hardware. A sophisticated APS array could create[23] a "camera-on-a-chip." It is possible to build a frame transfer device where pixels can be binned to enhance the SNR and provide variable resolution imaging.[24]

In 1993, Fossum[25] described state-of-the-art active pixel concepts. At that time they were the double-gate floating surface transistor, charge modulation device, bulk charge modulation device, based-stored image sensor, and static induction transistor. That review continued[23] in 1997.

The APS contains a photodiode, row-select transistor, and a reset transistor. As shown in Figure 3-41, by activating a row, the data from the pixels in that row are simultaneously copied into the columns. Each column will have a load transistor, column select switch, and a sampling switch. The chip may also have an analog-to-digital converter and provide correlated double sampling. The pixels are then reset and a new integration is started. The APS does not rely upon charge transfer. Rather the photodiode drives a capacitance line. The total capacitance limits the chip speed. APSs can be fully addressable and subarrays can be read out at high frame rates just like CIDs.

CMOS devices tend to have higher dark currents due to the highly doped silicon used. As such it does not appear that CMOS sensors will replace CCDs in low noise scientific applications. Because the active devices often take up real estate, the area for the photosensor is reduced. This leads to reduced sensitivity that can be partially offset by a microlens. Because each pixel has its own amplifier, pattern noise is larger. However, more logic can be added to each pixel and it seems possible to use on-chip signal processing that suppresses FPN.[26] With charge limited to a single pixel, it cannot overflow into neighboring pixels and create blooming as seen with CCDs.

While more integration can lead to the camera-on-a chip, production yield will probably decrease as the number of transistors increase. Consider an array that contains 640 × 480 pixels (307,200 pixels). With a CCD, there are 307,200 elements on the chip (neglecting dark pixels and the readout structure). With an APS that has a photodiode and three transistors per pixels there are four times more elements (1,228,800 total). If any one of the four elements fails, that pixel is dead.

A further advantage is their low power consumption. APSs operate from 5-volt supply or less (compared to 10 V to 15 V for a CCD). This is a very

attractive feature for man-portable, battery-operated devices. It appears at this time, CMOS will compete with CCDs in the general video marketplace where weight, size, and power consumption are factors.

3.11. PHYSICAL PARAMETERS

All detector arrays are fabricated for specific applications. Microlenses are used to increase the optical fill factor. Color filter arrays employ filters that are placed over the detectors to create specific spectral responses that lend themselves to color imagery. For general video applications, the number of detectors is matched to the video standard to maximize bandwidth and image sharpness. The overall array size follows the convention used with vidicons.

3.11.1. MICROLENSES

Optical fill factor may be less than 100% due to manufacturing constraints in full transfer devices. In interline devices, the shielded vertical transfer register can reduce the fill factor to less than 20%. Microlens assemblies (also called microlenticular arrays or lenslet arrays) increase the effective optical fill-factor (Figure 3-42). But it may not reach 100% due to slight misalignment of the microlens assembly, imperfections in the microlens itself, nonsymmetric shielded areas, and transmission losses. As shown by the camera formula (Equation 2-16, page 35), the output is directly proportional to the detector area. Increasing the optical fill factor with a microlens assembly increases the effective detector size and, therefore, the output voltage.

The photosensitive area is below the gate structure and the ability to collect the light depends upon gate thickness. The cone of light reaching the microlens depends upon the f-number of the primary camera lens. Figure 3-42 illustrates nearly parallel rays falling on the microlens. This case is encountered with high f-number lens systems. Low f-number primary camera lenses increase the cone angle and the effective fill-factor decreases with decreasing f-number.[27] Microlenses are optimized for most practical f-numbers. As the array size grows, off-axis detectors do not obtain the same benefit as on-axis detectors.[28]

Hyper-HAD is a Sony trademark which indicates the presence of a microlens assembly over a hole accumulation diode. While many camera manufacturers use the Hyper-HAD detectors, they do not identify the detector by Sony's trademark. This leads the user to believe that there are many different detectors and hence many different cameras on the market.

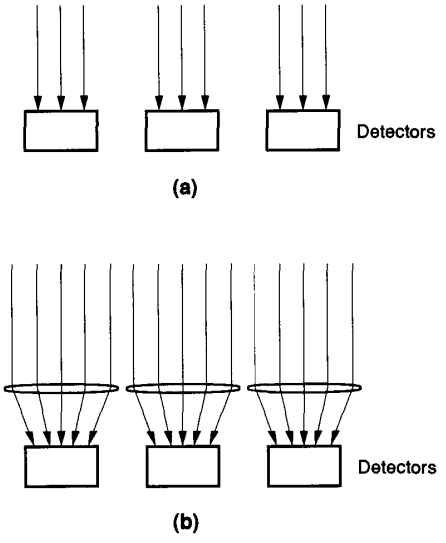


Figure 3-42. Optical effect of a microlens assembly. (a) With no microlens, a significant amount of photon flux is not detected. (b) The microlens assembly can image nearly all the flux onto the detector when a high f-number lens is used. These lenslets can either be grown on the array during the fabrication process or manufactured out of a material such as quartz and placed on the array surface during packaging.

3.11.2. COLOR FILTER ARRAYS

The subjective sensation of color can be created from three primary colors. By adjusting the intensity of each primary (additive mixing), a full gamut (rainbow) of colors is experienced. This approach is used on all color displays. They have red, green, and blue phosphors that, when appropriately excited, produce a wide spectrum of perceived colors. The CIE committee standardized a color perception model for the human observer in 1931. The literature[29] is rich with visual data that form the basis for color camera design.

The "color" signals sent to the display must be generated by three detectors, each sensitive to a primary or its complement. The primary additive colors are red, green, and blue (R, G, B) and their complementary colors are yellow, cyan and magenta (Ye, Cy, Mg). For high quality color imagery, three separate detectors (discussed in Section 5.3.3., *Color Correction*) are used whereas for consumer applications, a single array is used. The detectors are covered with different filters that, with the detector response, can approximate the primaries or their complements (Figure 3-43). A single array with filters is called a color filter array (CFA).

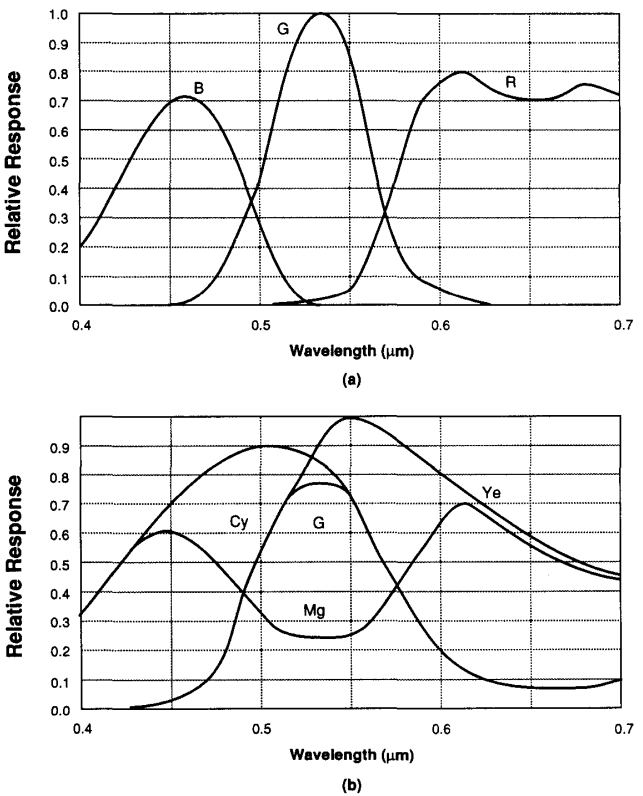


Figure 3-43. Desired spectral response for the three primaries and their complements. (a) The primaries and (b) their complements.

While CFAs can have filters with the appropriate spectral transmittance to create the primaries, it is more efficient to create the complementary colors. Complementary filters have higher transmittances than the primary filters. The luminance channel is identical to the green channel and is often labeled Y. "White" (no color filter) is represented by W = R + G + B.

A linear set of equations relates the primaries to their complements. These equations are employed (called matrixing) in cameras to provide either or both output formats. Matrixing can convert RGB into other color coordinates:

$$\begin{aligned} Ye &= R + G = W - B \\ Mg &= R + B = W - G \\ Cy &= G + B = W - R \,. \end{aligned} \tag{3-12}$$

The arrangement of the color filters for a single array system is either a stripe or a mosaic pattern (Figure 3-44). The precise layout of the mosaic varies by manufacturer. One basic CFA patent[30] was granted to Bryce E. Bayer at Eastman Kodak in 1976.

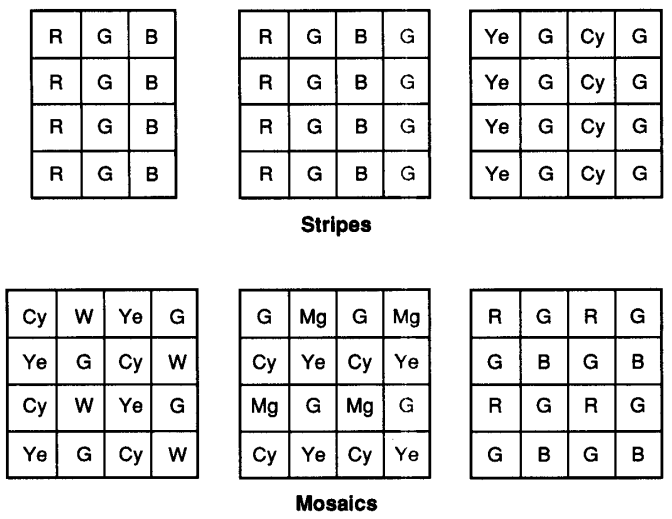


Figure 3-44. Representative stripe and mosaic arrays. Although shown as a full frame device for clarity, they typically are interline transfer devices. The layout depends upon the manufacturer's philosophy and cleverness in reducing color aliasing.

In many sensors, the number of detectors devoted to each color is different. The basic reason is that the human visual system (HVS) derives its detail information primarily from the green portion of the spectrum. That is, luminance differences are associated with green whereas color perception is associated with red and blue. The HVS requires only moderate amounts of red and blue to perceive color. Thus, many sensors have twice as many green as either red or blue detector elements. An array that has 768 horizontal elements may devote 384 to green, 192 to red, and 192 to blue. This results in an unequal sampling of the colors. A birefringent crystal (discussed in Section 10.3.4., *Optical Anti-alias Filter*), inserted between the lens and the array, accommodates the different sampling rates (discussed in Section 8.3.2., *Detector Array Nyquist Frequency*). The video signal from the CFA is embedded in the architecture of the CFA. The color video from a single chip CFA must be decoded (unscrambled) to produce usable R, G, and B signals.

3.11.3. NUMBER OF DETECTORS

Each detector array is designed for a specific application (discussed in Section 5.2.2., *Broadcast/Non-Broadcast Design*) The EIA 170 video standard supports 485 lines (discussed in Section 5.2.1., *Video Timing*). However, one line is split between the even and odd fields so that there are only 484 continuous lines. Thus detector arrays designed for EIA 170 compatibility tend to have 484 detectors in the vertical direction. For convenience, this has been reduced to 480 detectors. Image processing algorithms are more efficient with square pixels. With a 4:3 aspect ratio, the desired number of detectors is 640 × 480. Less expensive arrays will have a submultiple (for easy interpolation) such as 320 × 240.

For video applications, several rows or columns are devoted to dark current and for light leakage. Therefore an array may be 650 × 492 but the light sensitive part may be 640 × 480. It is manufacturer dependent whether to specify the array size by the number of active pixels or the total number of pixels (which includes the dark pixels).

Beat frequencies are very obvious in color images. To avoid beat frequencies, the color pixel rate should be a multiple of the chrominance subcarrier frequency (3.579545 MHz). This results in arrays that contain 384, 576, or 768 horizontal detectors. While CFAs are designed for NTSC operation, the same chip can be used for monochrome video without the CFA. Therefore, many monochrome arrays also contain 384, 576, or 768 horizontal detectors. These arrays do not have square pixels. But, from a manufacturing point-of-view, this reduces the inventory of arrays offered.

Scientific array size tends to a power of 2 (e.g., 512 × 512, 1024 × 1024) for easy image processing. There is a perception that "bigger is better" both in terms of array size and dynamic range. Arrays may reach 8192 × 8192 with a dynamic range of 16 bits. This array requires (8192)(8192)(16) or 1.07 Gbits of storage for each image. Image compression schemes may be required if storage space is limited. The user of these arrays must decide which images are significant and through data reduction algorithms, store only those that have value. Otherwise, he will be overwhelmed with mountains of data.

While large format arrays offer the highest resolution, their use is hampered by readout rate limitations. For example, consumer camcorder systems operating at 30 frames/s have a data rate of about 10 Mpixels/s. An array with 5120 × 5120 elements operating at 30 frames/s has a data rate of about 768 Mpixels/s. Large arrays can reduce readout rates by having multiple parallel ports servicing subarrays. Each subarray requires separate vertical and horizontal clock signals. The tradeoff is frame rate (speed) versus number of parallel ports (complexity of CCD design) and interfacing with downstream electronics. Because each subarray is serviced by different on-chip and off-chip amplifiers, the displayed image of the subarrays may vary in contrast and level. This is due to differences in amplifier gains and level adjustments.

3.11.4. VIDEO CHIP SIZE

Vidicon vacuum tubes were originally used for professional television applications. These were specified by the tube diameter. To minimize distortion and nonuniformities within the tube, the recommended[31] image size was considerably less than the overall tube diameter. When CCDs replaced the tubes, the CCD industry maintained the image sizes but continued to use the tube format nomenclature (Table 3-2).

Table 3-2
ARRAY SIZE for STANDARD FORMATS

CAMERA FORMAT	STANDARDIZED ARRAY SIZE (H × V)	ARRAY DIAGONAL
1 inch	12.8 mm × 9.6 mm	16 mm
2/3 inch	8.8 mm × 6.6 mm	11 mm
1/2 inch	6.4 mm × 4.8 mm	8 mm
1/3 inch	4.8 mm × 3.6 mm	6 mm
1/4 inch	3.2 mm × 2.4 mm	4 mm

Although each manufacturer supplies a slightly different array size and pixel size, nominal sizes for a 768 × 480 array are given in Table 3-3. With interline transfer devices, approximately one-half of the pixel width is devoted to the shielded vertical transfer register. That is, the detector active width is one-half of the pixel width. Thus the active area of the pixel is rectangular in interline transfer devices. This asymmetry does not appear to affect image quality in consumer video products significantly.

Table 3-3
NOMINAL PIXEL SIZE for a 768 × 480 ARRAY
Sizes vary by manufacturer. Detector sizes are smaller.

CAMERA FORMAT	NOMINAL PIXEL SIZE (H × V)
1 inch	16.7 μm × 20 μm
2/3 inch	11.4 μm × 13.8 μm
1/2 inch	8.33 μm × 10 μm
1/3 inch	6.25 μm × 7.5 μm
1/4 inch	4.17 μm × 5 μm

The decrease in optical format is related to cost. The price of CCD arrays is mainly determined by the cost of processing semiconductor wafers. As the chip size decreases, more devices can be put on a single wafer and this lowers the price of each individual device. The trend of going to smaller devices will probably continue as long as the optical and electrical performance of the imagers does not change. However, smaller pixels reduce the charge well size. For a fixed flux level and lens f-number, the smaller arrays have reduced sensitivity.

Smaller chips make for smaller cameras. However, to maintain resolution, pixels can only be made so small. Here, the tradeoff is among pixel size, optical focal length, and overall chip size. Further unless the lens f-number is reduced as the chip size is reduced, the system will move from being detector-limited to optically-limited (discussed in Section 10.4., *Optical-Detector Subsystem*). As this happens, the system MTF will change and smaller detectors will provide a less sharp image.

3.12. REFERENCES

1. J. Janesick, T. Elliott, R. Winzenread, J. Pinter, and R. Dyck, "Sandbox CCDs," in *Charge-Coupled Devices and Solid State Optical Sensors V*, M M. Blouke, ed., SPIE Proceedings Vol. 2415, pp. 2-42 (1995).
2. J. Janesick and T. Elliott, "History and Advancement of Large Area Array Scientific CCD Imagers," in *Astronomical Society of the Pacific Conference Series, Vol. 23, Astronomical CCD Observing and Reduction*, pp. 1-39, BookCrafters (1992).
3. M. J. Howes and D. V. Morgan, eds., *Charge-Coupled Devices and Systems*, John Wiley and Sons, New York, NY (1979).
4. C. H. Sequin and M. F. Tompsett, *Charge Transfer Devices*, Academic Press, New York, NY (1975).
5. E. S. Yang, *Microelectronic Devices*, McGraw-Hill, NY (1988).
6. E. L. Dereniak and D. G. Crowe, *Optical Radiation Detectors*, pp. 186-269, John Wiley and Sons, New York, NY (1984).
7. A. J. P. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices*, Kluwer Academic Publishers, Dordrecht, The Netherlands (1995).
8. M. Kimata and N. Tubouchi, "Charge Transfer Devices," in *Infrared Photon Detectors*, A. Rogalski, ed., pp. 99-144, SPIE Press, Bellingham, WA (1995).
9. A. J. P. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices*, pp. 317-348, Kluwer Academic Publishers, Dordrecht, The Netherlands (1995).
10. A. J. P. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices*, pp. 54-66, Kluwer Academic Publishers, Dordrecht, The Netherlands (1995).
11. A. J. P. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices*, pp. 161-165, Kluwer Academic Publishers, Dordrecht, The Netherlands (1995).
12. H.-S. Wong, Y. L. Yao, and E. S. Schlig, "TDI Charge-Coupled Devices: Design and Applications," *IBM Journal Research Development*, Vol. 36(1), pp. 83-106 (1992).
13. T. S. Lomheim and L. S. Kalman, "Analytical Modeling and Digital Simulation of Scanning Charge-Coupled Device Imaging Systems," in *Electro-Optical Displays*, M. A. Karim, ed., pp. 551-560, Marcel Dekker, New York (1992).
14. G. Wan, X. Gong, and Z. Luo, "Studies on the Measurement of Charge Transfer Efficiency and Photoresponse Nonuniformity of Linear Charge-coupled Devices," *Optical Engineering*, Vol. 34 (11), pp.3254-3260 (1995).
15. J. Janesick, "Open Pinned-Phase CCD Technology," in *EUV, X-ray, and Gamma Ray Instruments for Astronomy and Atomic Physics*, C. J. Hailey and O. H. Siegmund, eds., SPIE Proceedings Vol. 1159, pp 363-373 (1989).
16. J. Janesick, T. Elliot, G. Fraschetti, S. Collins, M. Blouke, and B. Corey, "CCD Pinning Technologies," in *Optical Sensors and Electronic Photography*, M. M. Blouke, ed., SPIE Proceedings Vol. 1071, pp. 153-169 (1989).
17. J. S. Campbell, "TC271 Characterization Report" in *Area Array Image Sensor Products*, Texas Instruments Data Catalog, pp. B-43 to B-46 (1996).
18. J. R. Janesick, T. Elliott, S. Collins, M. M. Blouke, and J. Freeman, "Scientific Charge-coupled Devices," *Optical Engineering*, Vol. 26(8), pp. 692-714 (1987).
19. C. Buil, *CCD Astronomy*, pp. 125-152, Willmann-Bell, Richmond, VA (1991).
20. S. Kawai, M. Morimoto, N. Mutoh, and N. Teranishi, "Photo Response Analysis in CCD Image Sensors with a VOD Structure," *IEEE Transactions on Electron Devices*, Vol. 40(4), pp. 652-655 (1995).
21. M. Schanz, W. Brockherde, R. Hauschild, B. Hosticka, and M. Schwarz, "Smart CMOS Image Sensor Arrays," *IEEE Transactions on Electron Devices*, Vol. 44(10), pp. 1699-1704 (1997).
22. S. K. Mendis, S. E. Kemeny, and E. R. Fossum, "A 128 × 128 CMOS Active Pixel Sensor for Highly Integrated Imaging Systems," *IEEE IEDM Technical Digest*, pp. 583-586 (1993).

23. E. R. Fossum, "CMOS Image Sensors: Electronic Camera-on-a-chip," *IEEE Transactions on Electron Devices*, Vol. 44(10), pp. 1689-1698 (1997).
24. Z. Zhou, B. Pain, and E. R. Fossum, "Frame-transfer CMOS Active Pixel Sensor with Pixel Binning," *IEEE Transactions on Electron Devices*, Vol. 44(10), pp. 1764-1768 (1997).
25. E. R. Fossum, "Active Pixel Sensors: Are CCD's Dinosaurs?, in *Charge-coupled Devices and Solid State Optical Sensors III*, M. M. Blouke, ed., SPIE Proceedings Vol. 1900, pp 2-14 (1993).
26. J. Nakamura, T. Nomoto, T. Nakamura, and E. R. Fossum, "On-focal-plane Signal Processing for Current-mode Active Pixel Sensors," *IEEE Transactions on Electron Devices*, Vol. 44(10), pp. 1747-1757 (1997).
27. J. Furukawa, I. Hiroto, Y. Takamura, T. Walda, Y. Keigo, A. Izumi, K. Nishibori, T. Tatebe, S. Kitayama, M. Shimura, and H. Matsui, "A ⅓-inch 380k Pixel (Effective) IT-CCD Image Sensor," *IEEE Transactions on Consumer Electronics*, Vol. CE-38(3), pp. 595-600 (1992).
28. M. Deguchi, T. Maruyama, F. Yamasaki, T. Hamamoto, and A. Izumi, "Microlens Design Using Simulation Program for CCD Image Sensor," *IEEE Transactions on Consumer Electronics*, Vol. CE-38(3), pp. 583-589 (1992).
29. See, for example, J. Peddie, *High-Resolution Graphics Display Systems*, Chapter 2, Windcrest/McGraw-Hill, New York, NY (1993) or A. R. Robertson and J. F. Fisher "Color Vision, Representation, and Reproduction," in *Television Engineering Handbook*, K. B. Benson, ed., Chapter 2, McGraw Hill, New York, NY (1985).
30. B. E. Bayer, U.S. patent #3,971,065 (1976).
31. R. G. Neuhauser, "Photosensitive Camera Tubes and Devices," in *Television Engineering Handbook*, K. B. Benson, ed., page 11.34, McGraw-Hill, New York, NY (1986).

4

ARRAY PERFORMANCE

The most common array performance measures are read noise, charge well capacity, and responsivity. From these the minimum signal, maximum signal, signal-to-noise ratio, and dynamic range can be calculated. Full characterization includes quantifying the various noise sources, charge transfer efficiency, spectral quantum efficiency, linearity, and pixel nonuniformity. These additional metrics are necessary for the most critical scientific applications.

Because the array is the basic building block of the camera, camera terminology is often used for array specifications. Additional performance measures exist at the system level. These include the noise equivalent irradiance (NEI) and noise equivalent differential reflectance (NE$\Delta\rho$) and are discussed in Chapter 6, *Camera Performance.*

The magnitude of each noise component must be quantified and its effect on system performance must be understood. Noise sources may be a function of the detector temperature. Predicted system performance may deviate significantly from actual performance if significant 1/f noise or other noise is present. There are a myriad of factors involved in system optimization. It is essential to understand what limits the system performance so that intelligent improvements can be made.

Janesick[1] uses three methods to characterize CCD performance: (1) photon transfer, (2) x-ray transfer, and (3) photon standard. The photon transfer technique (described in this chapter) appears to be the most valuable in calibrating, characterizing, and optimizing arrays for general video and industrial applications. Additional information, necessary for characterizing scientific arrays, is obtained from the x-ray transfer and photon standard methods.

Soft x-ray photons produce a known number of electron-hole pairs in a localized area of the CCD. Because the number and location of electron hole-pairs are known precisely, it is easy to determine the charge transfer efficiency. With the photon standard technique, the CCD array views a source with a known spectral distribution. Using calibrated spectral filters or a monochrometer, the spectral quantum efficiency is determined.

The symbols used in this book are summarized in the *Symbol List* (page xviii) which appears after the *Table of Contents*.

4.1. SIGNAL

Device specifications depend, in part, upon the application. Arrays for general video applications may have responsivity expressed in units of V/lux. For scientific applications, the units may be in V/(J-cm^{-2}) or, if a digital output is available, DN/(J-cm^{-2}) where DN refers to a digital number. For example, in an 8-bit system the digital numbers range from zero to 255. These units are incomplete descriptors unless the device spectral response and source spectral characteristics are furnished.

The maximum output occurs when the charge wells are filled. The exposure that produces this value is the saturation equivalent exposure (SEE). With this definition, it is assumed that the dark current produces an insignificant number of electrons so that only photoelectrons fill the well.

Sensitivity suggests something about the lowest signal that can be detected. It is usually defined as the input signal that produces a signal-to-noise ratio of one. This exposure is the noise equivalent exposure (NEE). The minimum signal is typically given as equivalent electrons rms. It is only one of many performance parameters used to describe system noise performance.

Figure 4-1 illustrates a signal transfer diagram. The shutter represents the integration time, t_{INT}. The detector converts the incident photons, $n_{DETECTOR}$, into electrons at a rate determined by the quantum efficiency. A voltage is created when the electrons are transferred to the sense node capacitance. The signal after the source follower amplifier is

$$V_{SIGNAL} = n_e \frac{G\,q}{C} , \qquad (4\text{-}1)$$

where n_e is the total number of electrons in the charge packet. It includes both photoelectrons and dark current electrons. The output gain conversion, Gq/C, is typically $0.1\,\mu V/e^-$ to $10\,\mu V/e^-$. The source follower amplifier gain, G, is near unity and, therefore, it is sometimes omitted from radiometric equations. Amplifiers, external to the CCD device, amplify the signal to a useful voltage. These amplifiers are said to be off-sensor or off-chip.

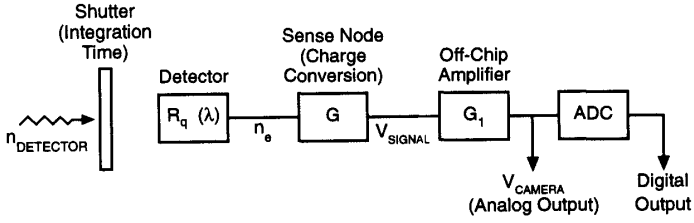


Figure 4-1. Representative signal transfer diagram. Many systems do not have a shutter, but it represents t_{INT}. The output may be specified as the number of electrons, output voltage of the source follower amplifier, or output of the off-chip amplifier. Some devices provide a digital output. V_{SIGNAL}, V_{CAMERA}, and DN can be measured whereas n_e is calculated.

4.1.1. SPECTRAL RESPONSE

For an ideal material, when the photon energy is greater than the semiconductor band gap energy, each photon produces one electron-hole pair (quantum efficiency is one). However, the absorption coefficient in silicon is wavelength dependent (Figure 4-2). Long wavelength photons are absorbed deeper into the substrate than short wavelengths. Very long wavelength photons may pass through the CCD and not be absorbed. Beyond 1.1 μm, the absorption is essentially zero because the photon energy is less than the band gap energy. For a doping concentration of $10^{17}cm^{-3}$, the depth at which 90% of the incident photons are absorbed is given in Table 4-1. Because of reflections, the photon incidance is higher. The values in Table 4-1 apply only to those photons that enter the material.

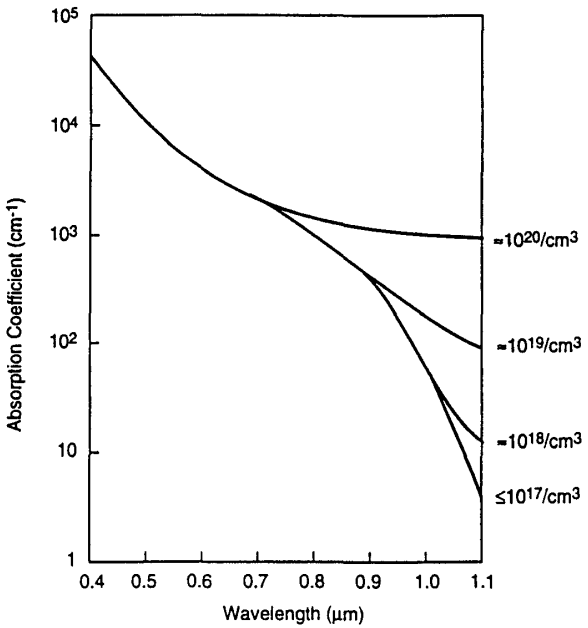


Figure 4-2. Silicon absorption coefficient as a function of wavelength and doping concentration at 25°C. Detectors typically have doping concentrations less than $10^{17}/cm^3$. (From Reference 2).

Any photon absorbed within the depletion will yield a quantum efficiency near unity. However, the depletion region size is finite and long wavelength photons will be absorbed within the bulk material. An electron generated in the substrate will experience a three-dimensional random walk until it recombines or reaches the edge of a depletion region where the electric field exists. If the diffusion length is zero, all electrons created within the bulk material will recombine immediately and the quantum efficiency approaches zero for these photons. As the diffusion length approaches infinity, the electrons eventually reach a charge well and are stored. Here, the quantum efficiency approaches one.

Table 4-1
REPRESENTATIVE 90% ABSORPTION DEPTH

Wavelength (μm)	Depth (μm)
0.40	0.19
0.45	1.0
0.50	2.3
0.55	3.3
0.60	5.0
0.65	7.6
0.70	8.5
0.75	16
0.80	23
0.85	46
0.90	62
0.95	150
1.00	470
1.05	1500
1.10	7600

Doping controls the diffusion length and the quantum efficiency is somewhere between these two extremes (Figure 4-3). The quantum efficiency is dependent upon the gate voltage (low voltages produce small depletion regions) and the material thickness (long wavelength photons will pass through thin substrates). Diffusion creates a responsivity that overlaps pixels (Figure 4-4). The maximum signal and other performance parameters defined in this chapter are for large area illumination. That is, a large number of detectors are illuminated and the average output is used. As the image size decreases, diffusion effects becomes more apparent. Diffusion effects are described by an MTF (discussed in Section 10.3.1., *Diffusion MTF*).

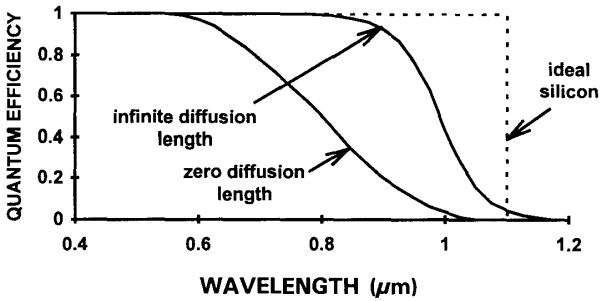


Figure 4-3. Theoretical internal quantum efficiency for silicon photosensors 250 μm thick. The long wavelength quantum efficiency depends upon the thickness of the substrate and the diffusion length. Most arrays have a relatively large diffusion length so that the quantum efficiency tends to be near the infinite diffusion length curve. (From Reference 3).

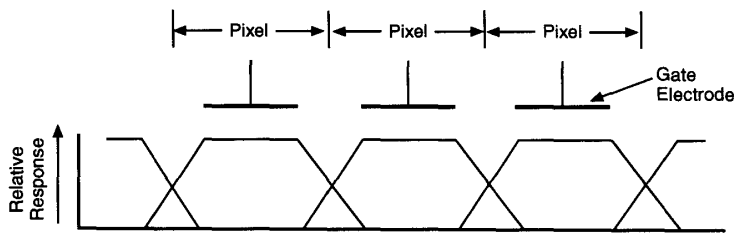


Figure 4-4. Idealized relative spatial response to long wavelength photons.

For front illuminated devices with photogates, the spectral responsivity deviates from the ideal spectral response due to the polysilicon gate electrodes (see Figure 3-2, page 48). The transmittance of polysilicon starts to decrease below 0.6 μm and becomes opaque at 0.4 μm (Figure 4-5). The transmittance depends upon the gate thickness and material.[4] The gate structure is a thin film and therefore interference effects will cause variation in quantum efficiency that is wavelength dependent. Interference effects can be minimized through choice of material and film thickness. Because different manufacturing techniques are used for each array type, the spectral response varies with manufacturer. If a manufacturer uses different processes, then the spectral response will vary within his product line.

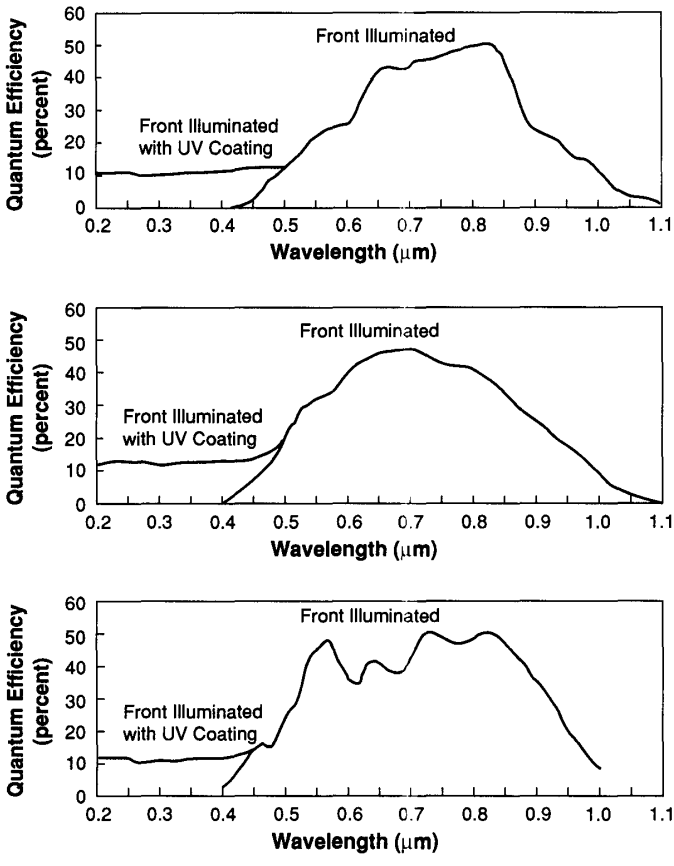


Figure 4-5. Typical response curves provided by different manufacturers. Above 0.8 μm, the quantum efficiency drops because the electron-hole pairs, generated deep in the array, may recombine before reaching a storage site. Below 0.6 μm, the polysilicon overcoat starts to become opaque. The variation in spectral response may not be a concern for many applications. Organic phosphors convert UV photons into visible photons and thereby extend the device's spectral response. The ripples in the responsivity are caused by interference effects.

Photodiodes, as used in most interline transfer devices, do not have polysilicon gates. Therefore these devices tend to have a response that approaches the ideal silicon spectral response. That is, there are no interference effects and the blue region response is restored. With virtual phase devices, one polysilicon electrode is replaced by a shallow highly doped layer. Because there is no electrode over this layer, higher quantum efficiency occurs in the blue region (see Figure 3-11, page 57). CIDs consist of two overlapping MOS capacitors sharing the same row and column electrode. Because the electrode only covers a part of the pixel, the area-averaged spectral response is higher the CCDs (see Figure 3-38, page 89).

Devices with vertical antibloom drains have reduced red response. The buried drain is in close proximity to where long photon absorption occurs. Any electron that enters the drain is instantly removed.

Even with no polysilicon overcoat, UV photons are absorbed at the surface. These electrons recombine before they reach a storage site. Hence the quantum efficiency is very low in the UV. The UV response can be enhanced with a UV fluorescent phosphor. These phosphors are deposited directly onto the array and they emit light at approximately 0.54 to 058 μm when excited by 0.12 to 0.45 μm light. Phosphors radiate in all directions and only the light fluorescing toward the array is absorbed. Lumogen, which is one of many available organic phosphors, has an effective quantum efficiency of about 15%. It does not degrade the quantum efficiency in the visible region because it is transparent at these wavelengths. The coating does not change the spectral response of the detector elements. Rather, it converts out-of-band photons into in-band photons. Therefore, the detector with the coating appears to have a spectral response from 0.12 to 1.1 μm.

Illuminating the array from the back side avoids the polysilicon problem and increases the quantum efficiency below 0.6 μm (Figure 4-6). Photons entering the back side are absorbed in the silicon and diffuse to the depletion region. However, short wavelength photons are absorbed near the surface and these electron-hole pairs recombine before reaching a storage site in a thick wafer (see Table 4-1). Therefore, the wafer is thinned to about 10 μm to maintain good spectral responsivity (Figure 4-7). In back-side thinned devices, the incident photon flux does not have to penetrate the polysilicon gate sandwich structure and interference effects are much easier to control. With a proper anti-reflection coating, a quantum efficiency of approximately 85% is possible. However, silicon has an index of refraction that varies with wavelength making it somewhat difficult to manufacture an anti-reflection coating that is effective across the entire spectrum. Coatings that optimize the response in the near IR

have reduced effectiveness in the visible (Figure 4-8). Lumogen reduces the effectiveness of the anti-reflection coating. New anti-reflection coatings overcome this problem and some offer UV response without the use of lumogen.[6]

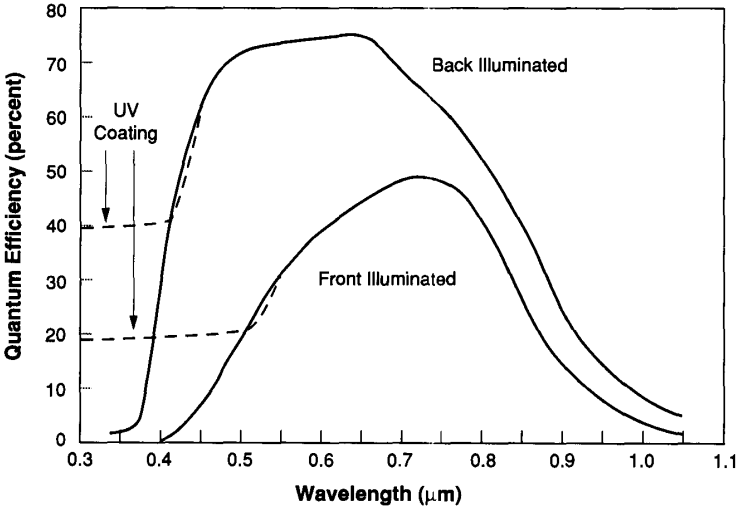


Figure 4-6. Representative spectral response of front illuminated and back illuminated CCD arrays. The spectral response depends upon the process used to manufacture each device.

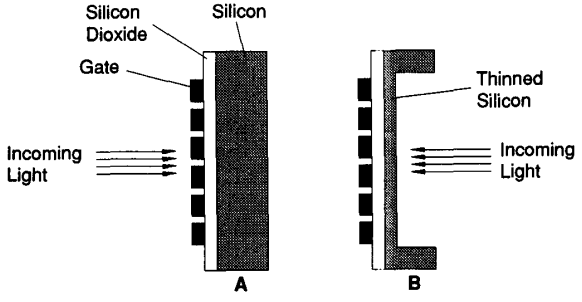


Figure 4-7. (a) Front illuminated and (b) back illuminated arrays. Owing to their extremely complex and fragile design, back illuminated devices are usually limited to scientific applications that require high quantum efficiency.

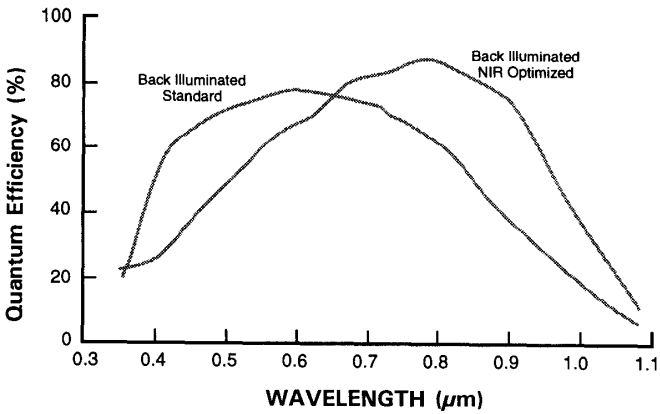


Figure 4-8. Effect of different anti-reflection coatings on back illuminated CCDs. (From Reference 5).

As the wavelength increases, the silicon becomes more transparent. With thin back illuminated devices, the rear surface and silicon oxide layer create an optical etalon. This creates constructive and destructive interference that appear to modulate the spectral responsivity.[5] At 0.8 μm, the modulation occurs approximately every 5 nm. This is bothersome in spectroscopic applications. For other applications, the wave band is sufficiently wide to cover several etalon fringe cycles and this averages out the effect.

CCD arrays for general and industrial applications are within a sealed environment for protection. Light must pass though a window (may be glass or quartz) to reach the array. In addition to reflection losses at all wavelengths, the glass transmittance decreases for wavelengths below 0.4 μm. For scientific applications requiring high sensitivity, the glass can be coated with a broad band anti-reflection coating. For UV applications, a quartz UV transmitting window can be used.

4.1.2. RESPONSIVITY

The spectral quantum efficiency is important to the scientific and military communities. When the array is placed into a general video or industrial camera, it is convenient to specify the output as a function of incident flux density or energy density averaged over the spectral response of the array.

Assume that the array is at a distance of R_1 from an ideal source whose area is A_S. This radiometric setup is similar to that described in Section 2.5., *Camera Formula*, page 33, except no lens is present. For convenience, radiant quantities are used rather than photon flux. Assuming a Lambertian source,

$$n_{pe} = \frac{1}{q}\frac{A_S A_D}{R_1^2}\int_{\lambda_1}^{\lambda_2}\frac{M_e(\lambda)}{\pi}R_e(\lambda)t_{INT}\,d\lambda\ , \tag{4-2}$$

where $R_e(\lambda)$ is the spectral response expressed in units of A/W. Note that the detector incidance is $E_e(\lambda) = M_e(\lambda)A_S/R_1^2$. Then

$$n_{pe} = \frac{A_D}{q}\int_{\lambda_1}^{\lambda_2}\frac{E_e(\lambda)}{\pi}R_e(\lambda)t_{INT}\,d\lambda\ . \tag{4-3}$$

The array output voltage (after the source follow amplifier) is $(G\,q\,n_{pe})/C$:

$$V_{SIGNAL} = \frac{G}{C}\frac{A_S A_D}{R_1^2}\int_{\lambda_1}^{\lambda_2}\frac{M_e(\lambda)}{\pi}R_e(\lambda)t_{INT}\,d\lambda \tag{4-4}$$

or

$$V_{SIGNAL} = \frac{G}{C}A_D\int_{\lambda_1}^{\lambda_2}\frac{E_e(\lambda)}{\pi}R_e(\lambda)t_{INT}\,d\lambda\ . \tag{4-5}$$

It is desirable to express the responsivity in the form

$$V_{SIGNAL} = R_{AVE}\left(\frac{1}{A_D}\frac{A_S A_D}{R_1^2}\int_{\lambda_1}^{\lambda_2}\frac{M_e(\lambda)}{\pi}t_{INT}\,d\lambda\right)\ . \tag{4-6}$$

The value R_{AVE} is an average response that has units of $V/(J\text{-}cm^{-2})$ and the quantity in the brackets has units of J/cm^2.

Combining the two equations provides

$$R_{AVE} = \frac{G}{C} A_D \frac{\int_{\lambda_1}^{\lambda_2} M_e(\lambda)\, R_e(\lambda) d\lambda}{\int_{\lambda_1}^{\lambda_2} M_e(\lambda) d\lambda} \quad \frac{V}{J/cm^2} . \tag{4-7}$$

Note that R_q and R_{AVE} are the for array package. R_q or R_{AVE} is the detector responsivity multiplied by the window transmittance. Usually the manufacturer includes the window transmittance in the quoted responsivity.

Experimentally, the most popular approach to measure R_{AVE} is to vary the exposure (e.g., the integration time). The value R_{AVE} is the slope of the output/input transformation (Figure 4-9). The bracketed term in Equation 4-6 is the input. By placing a calibrated radiometer next to the array (at distance R_1 from the source), the incidance, E_e, can be measured, Multiplying by the integration time provides J/cm^2. These equations are valid over the region that the array output/input transformation is linear. The incidance can also be changed by moving the source (varying R_1).

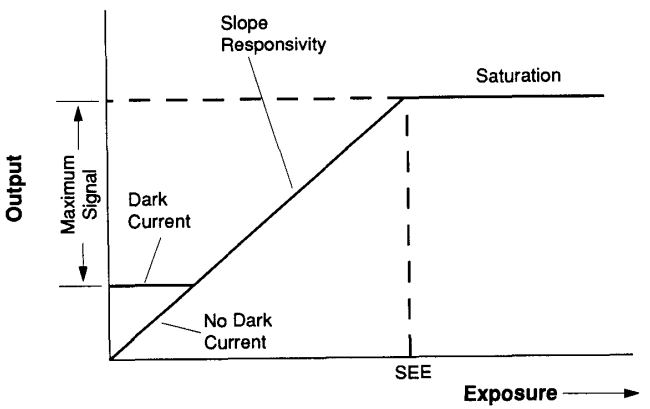


Figure 4-9. The average responsivity is the slope of the output-input transformation. The maximum input or the saturation equivalent exposure (SEE) is the input that fills the charge wells. SEE is used to define the dynamic range. Dark current limits the available signal strength. Cooling can reduce the dark current to a negligible level.

It is imperative that the source color temperature remains constant. Assuming an incandescent bulb is used, reducing the bulb voltage will reduce the output but this changes the color temperature. Changing the source temperature changes the integrals and this leads to incompatible results (see Section 2.7., *Normalization Issues*, page 39). The incidance can only be varied by inserting neutral density filters.

The value R_{AVE} is an average type responsivity that depends upon the source characteristics and the spectral quantum efficiency. While the source can be standardized (e.g., CIE illuminant A or illuminant D_{6500}), the spectral quantum efficiency varies by device (see Figure 4-5, page 108). Therefore extreme care must be exercised when comparing devices solely by the average responsivity.

Both $M_e(\lambda)$ and $R_e(\lambda)$ are functions of wavelength. If a very small wavelength increment is selected, $M_e(\lambda)$ and $R_e(\lambda)$ may be considered as constants. Equation 4-7 can be approximated as

$$R_{AVE} \approx \frac{G}{C} A_D \frac{M_e(\lambda_o)\, R_e(\lambda_o)\, \Delta\lambda}{M_e(\lambda_o)\, \Delta\lambda} \approx \frac{G}{C} A_D R_e(\lambda_o) \,. \qquad (4\text{-}8)$$

The responsivity (with units of A/W) is related to the quantum efficiency R_q by $R_e = (q\lambda/hc)\, R_q$. If the wavelength is measured in micrometers, then $R_e = (\lambda/1.24)\, R_q$. At a specific wavelength

$$\begin{aligned} R_{SIGNAL}(\lambda_o) &= \frac{Gq}{C} \frac{\lambda_o}{hc} A_D R_q(\lambda_o) \quad \frac{v}{J/cm^2} \\ &= \frac{G}{C} A_D R_e(\lambda_o) \end{aligned} \qquad (4\text{-}9)$$

or

$$R_{SIGNAL}(\lambda_o) = \frac{G}{C} \frac{\lambda_o}{1.24} A_D R_q(\lambda_o) \quad \frac{v}{J/cm^2} \,. \qquad (4\text{-}10)$$

Here, the detector area is measured in square centimeters. Typically $R_q(\lambda_o)$ is evaluated at that wavelength which provides the peak quantum efficiency: $R_q(\lambda_P)$. The symbol η is often used for the quantum efficiency.

If the device has a digital output, the number of counts (integer value) is

$$DN = int\left[V_{CAMERA}\frac{2^N}{V_{MAX}}\right], \tag{4-11}$$

where N is the number of bits provided by the analog-to-digital converter and *int* represents the integer value of the bracketed term. The value N is usually 8 for general video applications but may be as high as 16 for scientific cameras. The value V_{MAX} is the maximum output that corresponds to a full well, N_{WELL}:

$$V_{MAX} = G_1\frac{G\,q}{C}N_{WELL}\,. \tag{4-12}$$

The responsivity at the peak wavelength is approximated by

$$R_{AVE} \approx \left[\frac{2^N}{N_{WELL}}\frac{\lambda_P}{1.24}A_D R_q(\lambda_P)\right]\ \frac{DN}{J/cm^2}\,. \tag{4-13}$$

This assumes that the analog-to-digital converter input dynamic range exactly matches V_{MAX}. This may not be the case in real systems. If an antibloom drain is present, then the responsivity is defined up to the knee point (see Figure 3-35, page 87) and Equation 4-11 is modified accordingly.

When using arrays for general video, the output is normalized to photometric units (see Section 2.3.1., *Units*, page 25). It is convenient to express the array average response as

$$V_{SIGNAL} = R_{PHOTOMETRIC}\left(683\frac{1}{A_D}\int_{0.38}^{0.75\,\mu m} M_p(\lambda)\,V(\lambda)\,d\lambda\right), \tag{4-14}$$

where $R_{PHOTOMETRIC}$ has units of V/lux and the bracketed term has units of lux. Combining Equations 4-4 and 4-14 yields

$$R_{PHOTOMETRIC} = \frac{\frac{G}{C}\frac{A_S A_D}{R_1^2}\int_{\lambda_1}^{\lambda_2} L_e(\lambda)\,R_e(\lambda)\,t_{INT}\,d\lambda}{683\frac{1}{A_D}\int_{0.38}^{0.75\,\mu m} M_p(\lambda)\,V(\lambda)\,d\lambda}\ \frac{V}{lux}\,. \tag{4-15}$$

Because $M_P(\lambda)/A_D = E_e(\lambda)$, for a Lambertian source,

$$R_{PHOTOMETRIC} = \frac{\frac{G}{C} A_D \int_{\lambda_1}^{\lambda_2} E_e(\lambda)\, R_e(\lambda)\, t_{INT}\, d\lambda}{683 \int_{0.38}^{0.75\,\mu m} E_e(\lambda)\, V(\lambda)\, d\lambda} \quad \frac{V}{lux} . \qquad (4\text{-}16)$$

$R_{PHOTOMETRIC}$ depends on the integration time and is either 1/60 or 1/30 s depending on the architecture for EIA 170 or NTSC video standards (see Section 3.3.4., *Interline Transfer*, page 63). Experimentally, the photometric illuminance is measured with a calibrated sensor. The input intensity levels are changed either by moving the source (changing R_1) or by inserting neutral density filters between the source and the array.

The responsivity depends on the spectral response and the spectral content of the illumination. The apparent variation in output with different light sources was discussed in Section 2.7., *Normalization Issues*, page 39. Selecting an array based on the responsivity is appropriate if the anticipated scene illumination has the same color temperature as the calibration temperature. That is, if the photometric responsivity is measured with a CIE illuminant A and the scene color temperature is near 2856 K, then selecting an array with the highest photometric responsivity is appropriate. Otherwise, the average photometric responsivity is used for informational purposes only.

Five numerical examples follow. Additional examples can be found in Chapter 6.

Example 4-1
MAXIMUM OUTPUT

A device provides a output gain conversion of 6 μV/e$^-$. The well size is 70,000 electrons. What is the maximum output?

The maximum output is the OGC multiplied by the well capacity or 420 mV. If the dark current creates 5,000 electrons, then the maximum signal voltage is reduced to (70,000 − 5,000)(0.006 mV) = 390 mV. This signal may be amplified by an off-chip amplifier.

Example 4-2
DIGITAL RESPONSIVITY

A device with a well capacity of 50,000 electrons has its output digitized with a 12-bit analog-to-digital converter (ADC). What is the digital step-size (digital responsivity)?

The step size is the well capacity divided by the number of digital steps. $50{,}000/2^{12} = 12.2$ electrons/DN. Equation 4-2 (page 112) provides the number of electrons as a function of input exposure. The relationship between the ADC maximum value and charge well capacity assumes an exact match between V_{MAX} and the ADC input range. If these do not match, an additional amplifier must be inserted just before the ADC.

Example 4-3
RESPONSIVITY

A 2/3-inch format interline transfer device provides an output gain conversion of 6 μV/e$^-$. If the quantum efficiency is 0.8 at 0.6 μm, what is the responsivity in units of V/μJ-cm^{-2} ?

From Table 3-3 (page 99), a 2/3-inch format pixel is approximately 11.4 μm $\times$ 13.8 μm. If the interline transfer device has a 50% fill factor, then $A_D = 7.87 \times 10^{-7}$ cm^2. The constant hc is 1.99×10^{-25} J-m/photon. Using Equation 4-9,

$$R_{SIGNAL}(\lambda_o) = (6 \times 10^{-6})\left(\frac{0.6 \times 10^{-6}}{1.99 \times 10^{-25}}\right)(7.87 \times 10^{-7})\,(0.8) = 11.4 \times 10^{6} \quad (4\text{-}17)$$

$$\frac{V}{J\text{-}cm^{-2}} = \left(\frac{V}{e^-}\right)\left(\frac{m}{(J\text{-}m)/photon}\right)(cm^2)\left(\frac{electrons}{photons}\right)$$

or $R_{SIGNAL}(\lambda_o) = 11.4$ V/μJ-cm^{-2}.

Example 4-4
PHOTOMETRIC OUTPUT

An array designed for video applications has a sensitivity of 250 μV/lux when measured with CIE illuminant A. What is the expected output if the input is 1000 lux?

If the color temperature does not change and only the intensity changes, then the output is (250 μV)(1000) = 250 mV. However, if the color temperature changes, the output changes in a nonlinear manner. Similarly, two sources may produce the same illuminance but V_{SIGNAL} may be quite different. The color temperature must always be specified when $R_{PHOTOMETRIC}$ is quoted.

Implicit in this calculation is that the integration time remains constant. Because photometric units are used for general video applications, it is reasonable to assume that t_{INT} is either 1/60 or 1/30 s for EIA 170 compatibility.

Example 4-5
PHOTOMETRIC CONVERSION

An array designed for video applications has a sensitivity of 20 V/μJ-cm^{-2}. What is the estimated photometric response? The relative spectral response of $R_{SIGNAL}(\lambda)$ is illustrated in Figure 4-10.

Substituting Equation 4-9 into Equation 4-16 provides

$$R_{PHOTOMETRIC} = t_{INT} R_{SIGNAL-MAX} \left[\frac{\int_{\lambda_1}^{\lambda_2} E_e(\lambda) R_{SIGNAL-RELATIVE}(\lambda)\, d\lambda}{683 \int_{0.38}^{0.75\,\mu m} E_e(\lambda) V(\lambda)\, d\lambda} \right] \frac{V}{lux} \quad (4\text{-}18)$$

Two different spectral responsivities are considered: one as shown in Figure 4-10 and the other has an additional infrared blocking filter that cuts off at 0.7 μm. Numerical integration provides the value of the bracketed term (Table 4-2). As the color temperature increases, the bracketed value may either increase or decrease depending upon the spectral response. With $R_{SIGNAL-MAX}$ = 2000 V/J-m^{-2} (note change in units), t_{INT} = 1/60 s, $R_{PHOTOMETRIC}$ is given in Table 4-3.

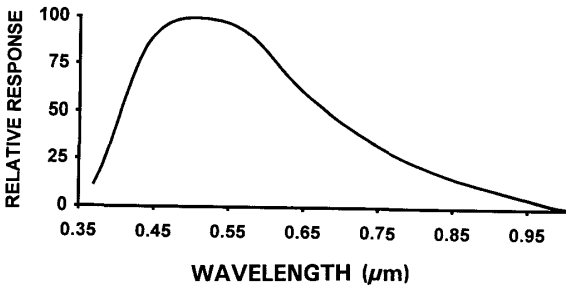


Figure 4-10. Relative spectral responsivity of $R_{SIGNAL}(\lambda)$.

Table 4-2
EQUATION 4-18 BRACKETED TERM

COLOR TEMPERATURE	NO INFRARED FILTER	WITH INFRARED FILTER
2856 K	0.00473	0.00319
6500 K	0.00391	0.00353

Table 4-3
PHOTOMETRIC RESPONSIVITY (V/lux)

COLOR TEMPERATURE	NO INFRARED FILTER	WITH INFRARED FILTER
2856 K	0.158	0.106
6500 K	0.130	0.117

4.1.3. MINIMUM SIGNAL

The noise equivalent exposure (NEE) is an excellent diagnostic tool for production testing to verify noise performance. NEE is a poor array-to-array comparison parameter and should be used cautiously when comparing arrays with different architectures. This is so because it depends on array spectral

responsivity and noise. The NEE is the exposure that produces a signal-to-noise ratio of one. If the measured rms noise on the analog output is V_{NOISE}, then the NEE is calculated from the radiometric calibration:

$$NEE = \frac{V_{NOISE}}{R_{AVE}} \quad \frac{J}{cm^2} \; rms. \tag{4-19}$$

When equated to electrons, NEE is simply the noise value in rms electrons. The absolute minimum noise level is the noise floor and this value is used most often for the NEE. Noise sources are discussed in Section 4.2., *Noise*. Although noise is a rms value, the notation *rms* is often omitted. When divided by the average quantum efficiency, the noise equivalent signal (NES) is obtained:

$$NES = \frac{NEE}{R_{q\text{-}AVE}} \quad photons. \tag{4-20}$$

4.1.4. MAXIMUM SIGNAL

The maximum signal is that input signal that saturates the charge well and is called the saturation equivalent exposure (SEE). It is

$$SEE = \frac{V_{MAX}}{R_{AVE}} \quad \frac{J}{cm^2} \; . \tag{4-21}$$

The well size varies with architecture, number of phases, and pixel size. The well size is approximately proportional to pixel area (Table 4-4). Small pixels have small wells. If an antibloom drain is present, the maximum level is taken as the white clip level (see Figure 3-35, page 87). The maximum value of V_{MAX} is

$$V_{MAX} = G_1 \frac{G\,q}{C} (N_{WELL} - n_{DARK}) \; . \tag{4-22}$$

For back-of-the-envelope calculations, the dark current is considered negligible (i.e., the device is cooled or MPP is implemented). Then $n_{DARK} \approx 0$.

Table 4-4
NOMINAL WELL CAPACITY for
TEXAS INSTRUMENTS FRAME TRANSFER DEVICES
(Varies by manufacturer and device type)

FORMAT	WELL CAPACITY (electrons)
2/3 inch	200,000
1/2 inch	80,000
1/3 inch	30,000

4.1.5. DYNAMIC RANGE

Dynamic range defined as the maximum signal (peak) divided by the rms noise. If an antibloom drain is present, the knee value is used as the maximum (See Figure 3-35, page 87). Expressed as a ratio, the dynamic range is

$$DR_{ARRAY} = \frac{SEE}{NEE} = \frac{V_{MAX}}{V_{NOISE}} \tag{4-23}$$

and when expressed in decibels,

$$DR_{ARRAY} = 20\log\left(\frac{SEE}{NEE}\right) \quad dB\ . \tag{4-24}$$

When equated to electrons,

$$DR_{ARRAY} = \frac{N_{WELL} - n_{DARK}}{\langle n_{SYS} \rangle}\ , \tag{4-25}$$

where $<n_{SYS}>$ is the system noise measured in electrons rms. The effects of $<n_{SYS}>$ on dynamic range is further discussed in Section 4.3., *Array Signal-to-Noise Ratio*. Often, only the noise floor is considered in $<n_{SYS}>$. The noise is dependent upon amplifier design and is not related to charge well capacity.

While larger arrays have larger well capacities, this does not necessary mean that the dynamic range increases by the same amount. As the pixel rate increases, the amplifier bandwidth increases. Simultaneously, the amplifier noise increases and the dynamic range decreases.

Sometimes the manufacturer will specify dynamic range by the peak signal divided by the peak-to-peak noise. This peak-to-peak value includes any noise spikes and, as such, may be more useful in determining the lowest detectable signal in some applications. If the noise is purely Gaussian, the peak-to-peak is assumed to be four to six times the rms value. This multiplicative factor is author dependent.

If the system noise is less than the ADC's least significant bit (LSB), then the quantization noise limits the system dynamic range. The rms quantization noise, $<n_{ADC}>$, is $V_{LSB}/\sqrt{12}$ where V_{LSB} is the voltage step corresponding to the least significant bit. The dynamic range is $2^N \sqrt{12}$. For a 12-bit ADC, the dynamic range is 14,189:1 or 83 dB. An 8-bit system cannot have a dynamic range greater than 887:1 or 59 dB.

Example 4-6
ARRAY PARAMETERS

An array has a dynamic range of 3000:1, average responsivity of 2 $V/(\mu J/cm^2)$, and a saturation equivalent exposure of 150 nJ/cm^2. These values apply only at λ_P. What are the remaining design parameters?

Maximum output:

The maximum output is R_{AVE} SEE = 300 mV.

Noise level:

The noise floor rms value is equal to the maximum output divided by the dynamic range or 100 μV rms.

Analog-to-digital Converter:

For a digital output, the ADC must have 12 bits ($2^{12} = 4096$ levels) to span the array's full dynamic range. The responsivity in DN is 2^{12}/SEE = 27.3 $DN/(nJ/cm^2)$ at λ_P.

4.2. NOISE

Many books and articles[7-13] have been written on noise sources. The level of detail used in noise modeling depends on the application. Shot noise is due to the discrete nature of electrons. It occurs when photoelectrons are created and when dark current electrons are present. Additional noise is added when reading the charge (reset noise) and introduced by the amplifier (1/f noise and white noise). If the output is digitized, the inclusion of quantization noise may be necessary. Switching transients that are coupled though the clock signals also appear as noise. It can minimized by good circuit design.

Figure 4-1 (page 104) illustrated the signal transfer diagram. The noise transfer diagram is somewhat different (Figure 4-11). With different transfer functions, both the optical and electronic subsystems must be considered to maximize the system signal-to-noise ratio.

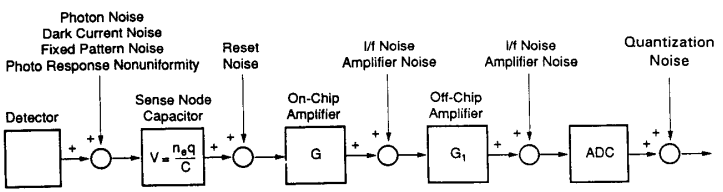


Figure 4-11. The various subsystems are considered as ideal elements with the noise introduced at appropriate locations.

Although the origins of the noise sources are different, they all appear as variations in the image intensity. Figure 4-12a illustrates the ideal output of the on-chip amplifier before correlated double sampling. The array is viewing a uniform source and the output of each pixel is identical. Photon shot noise produces a temporal variation in the output signal that is proportional to the square root of the signal level in electrons. Each pixel output will have a slightly different value (Figure 4-12b).

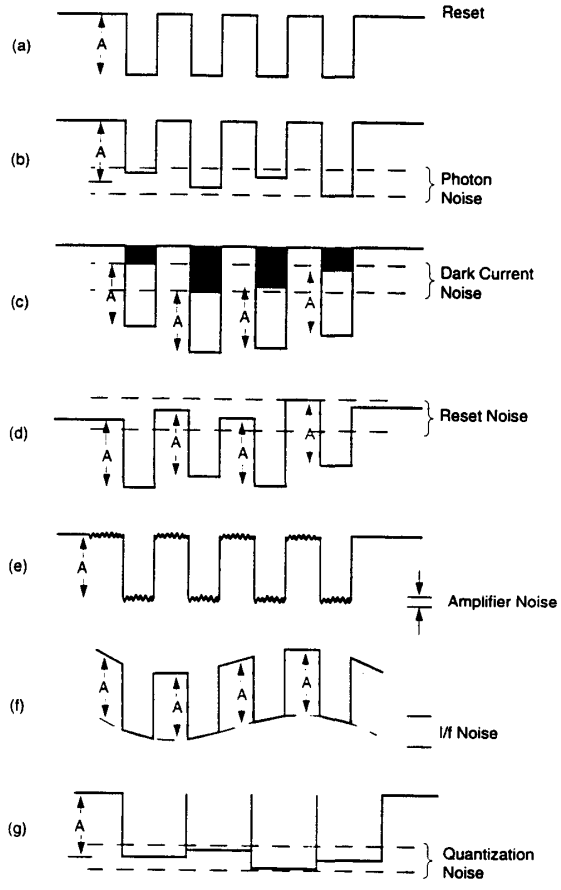


Figure 4-12. Noise sources affect the output differently. (a) though (f) represent the output after the on-chip amplifier. (g) is the output after the sample-and-hold circuitry. The array is viewing a uniform source that produces identical outputs (amplitude = A) for each pixel. (a) ideal output, (b) photon shot noise, (c), dark current shot noise, (d) reset noise, (e) amplifier noise, (f) amplifier 1/f noise, and (g) quantization noise. These values change from frame-to-frame. Pattern noise is not shown but is a variation that does not change significantly each frame. All these processes occur simultaneously.

Under ideal conditions, each pixel would have the same dark current and this value is subtracted from all pixels leaving the desired signal (see Section 3.6., *Dark Current*, page 79). However, dark current also exhibits fluctuations. Even after subtracting the average value, these fluctuations remain and create fixed pattern noise (Figure 4-12c). The output capacitor is reset after each charge is read. Errors in the reset voltage appear as output signal fluctuations (Figure 4-12d). The amplifier adds white noise (Figure 4-12e) and 1/f noise (Figure 4-12f). Amplifier noise affects both the active video and reset levels. The analog-to-digital converter introduces quantization noise (Figure 4-12g). Quantization noise is apparent after image reconstruction. Only photon shot noise and amplifier noise affect the amplitude of the signal. With the other noise sources listed, the signal amplitude, A, remains constant. However, the value of the output fluctuates with all the noise sources. Because this value is presented on the display, the processes appear as displayed noise. These noise patterns change from pixel-to-pixel and on the same pixel from frame-to-frame.

Pattern noise refers to any spatial pattern that does not change significantly from frame-to-frame. Dark current varies from pixel-to-pixel and this variation is called fixed pattern noise (FPN). FPN is due to differences in detector size, doping density, and foreign matter getting trapped during fabrication. Photoresponse nonuniformity (PRNU) is the variation in pixel responsivity and is seen when the device is illuminated. This noise is due to differences in detector size, spectral response, and thickness in coatings. These "noises" are not noise in the usual sense. PRNU occurs when each pixel has a different average value. This variation appears as spatial noise to the observer. Frame averaging will reduce all the noise sources except FPN and PRNU. Although FPN and PRNU are different, they are sometimes collectively called scene noise, pixel noise, pixel nonuniformity, or simply pattern noise.

It is customary to specify all noise sources in units of equivalent electrons at the detector output. Amplifier noise is reduced by the amplifier gain. When quoting noise levels, it is understood that the noise magnitude is the rms of the random process producing the noise. Noise powers are considered additive.

Equivalently, the noise sources are RSSed (root sum of the squares) or added in quadrature:

$$\langle n_{SYS} \rangle = \sqrt{\langle n_{SHOT}^2 \rangle + \langle n_{PATTERN}^2 \rangle + \langle n_{RESET}^2 \rangle + \langle n_{ON\text{-}CHIP}^2 \rangle + \langle n_{OFF\text{-}CHIP}^2 \rangle + \langle n_{ADC}^2 \rangle} \quad (4\text{-}26)$$

where $\langle n_i^2 \rangle$ is the noise variance for source i and $\sqrt{\langle n_i^2 \rangle} = \langle n_i \rangle$ is the standard deviation measured in rms units.

For system analysis, it may be sufficient to consider only the on-chip amplifier noise. The array manufacturer usually provides this value and may call it readout noise, mux noise, noise equivalent electrons, or the noise floor. The value varies by device and manufacturer. Reset noise can be reduced to a negligible level with CDS (discussed in Section 4.4., *Correlated Double Sampling*). CDS can also reduce the source follower 1/f noise. The off-chip amplifier is usually a low noise amplifier such that its noise is small compared to the on-chip amplifier noise. Although always present, quantization noise is reduced by selecting the appropriated sized analog-to-digital converter (e.g., selecting an ADC with many quantization levels). The simplified noise model is

$$\langle n_{SYS} \rangle \approx \sqrt{\langle n_{SHOT}^2 \rangle + \langle n_{FLOOR}^2 \rangle + \langle n_{PATTERN}^2 \rangle} \; . \quad (4\text{-}27)$$

The noise model does not include effects such as banding and streaking which are forms of pattern noise found in linear and TDI arrays. Banding can occur with arrays that have multiple outputs serviced by different nonlinear amplifiers (see Figure 3-14, page 61). Streaking occurs when the average responsivity changes from column to column in TDI devices.[14] These effects are incorporated in the three-dimensional noise model (discussed in Section 12.2., *Three-dimensional Noise Model*) and are used to predict the performance of military systems. Only a complete signal-to-noise ratio analysis can determine which noise source dominates (discussed in Section 4.3., *Array Signal-to-noise Ratio*).

4.2.1. SHOT NOISE

Both photoelectrons and dark current contribute to shot noise. Using Poisson statistics, the variance is equal to the mean:

$$\left\langle n_{SHOT}^2 \right\rangle = \left\langle n_{PE}^2 \right\rangle + \left\langle n_{DARK}^2 \right\rangle = n_{PE} + n_{DARK} . \quad (4\text{-}28)$$

The number of photoelectrons was given by Equation 4-2, page 112, and the number of dark current electrons was given by Equation 3-6, page 80. These values should be modified by the CTE:

$$\left\langle n_{SHOT}^2 \right\rangle = \left\langle n_{pe}^2 \right\rangle + \left\langle n_{DARK}^2 \right\rangle = (CTE)^N n_{pe} + (CTE)^N n_{DARK} . \quad (4\text{-}29)$$

Because the CTE is high ($CTE^N \approx 1$), it is usually omitted from most equations. However, CTE should be included for very large arrays. With TDI, the number of photoelectrons and dark current electrons increases with the number of TDI elements, N_{TDI}:

$$\left\langle n_{SHOT}^2 \right\rangle = N_{TDI} n_{pe} + N_{TDI} n_{DARK} . \quad (4\text{-}30)$$

While the dark current average value can be subtracted from the output to provide only the signal due to photoelectrons, the dark current shot noise cannot. Cooling the array can reduce the dark current to a negligible value and thereby reduce dark current shot noise to a negligible level (see Section 3.6., *Dark Current*, page 79).

4.2.2. RESET NOISE

The noise associated with resetting the sense node capacitor is often called kTC noise. This is due to thermal noise generated by the resistance, R, within the resetting FET. The Johnson noise current variance is

$$\left\langle i_n^2 \right\rangle = \frac{4kT}{R} \Delta f . \quad (4\text{-}31)$$

Because the resistance is in parallel with the sense node capacitor, the noise equivalent bandwidth is $\Delta f = RC/4$. Then

$$\langle i_n^2 \rangle = kTC \ . \tag{4-32}$$

When expressed as equivalent electrons, the rms noise is

$$\langle n_{RESET} \rangle = \frac{\sqrt{kTC}}{q} \quad e^- \ rms \ . \tag{4-33}$$

The value $\langle n_{RESET} \rangle$ represents the uncertainty in the amount of charge remaining on the capacitor following reset.

For a 0.01 pF capacitor at room temperature, kTC noise is about 40 electrons rms and a 0.2 pf capacitor produces about 126 e^- rms. Reducing the capacitance reduces the noise. This has an added benefit that the device output gain conversion, Gq/C, increases. kTC noise may be significantly reduced with correlated double sampling. While cooling reduces this noise source, cooling is used primarily to reduce dark current. That is, cooling reduces dark current noise exponentially whereas reset noise is reduced only by $\sqrt{T}$.

4.2.3. ON-CHIP AMPLIFIER NOISE

Amplifier noise consists of two components: 1/f noise and white noise. If f_{KNEE} is the frequency at which the 1/f noise equals the white noise, then the amplifier noise density is

$$V_{ON\text{-}CHIP} = V_{AMP\text{-}NOISE}\left(1 + \frac{f_{KNEE}}{f}\right) \quad \frac{rms \ volts}{\sqrt{Hz}} \ . \tag{4-34}$$

The 1/f noise can be minimized through correlated double sampling.

When discussing individual noise sources, noise is usually normalized to unit bandwidth. When modified by electronic subsystems, the total system noise power is

$$\sigma^2_{SYS} = \int_0^\infty S(f_{elec})\,|H_{sys}(f_{elec})|^2\,df_{elec} \quad volts\ , \tag{4-35}$$

where $S(f_{elec})$ is the total noise power spectral density from all sources. The subscript *elec* is used to emphasize that this is electrical frequency (Hz). $H_{sys}(f_{elec})$ is the frequency response of the system electronics. The variable $H(f_{elec})$ is used by electronic circuitry designers and its magnitude is identical to the MTF. The noise equivalent bandwidth is that bandwidth with unity value that provides the same total noise power. Assuming the noise is white over the spectral region of interest [$S(f_{elec}) = S_o$], Δf_{elec} is

$$\Delta f_{elec} = \frac{\int_0^\infty S(f_{elec})|H_{sys}(f_{elec})|^2\,df_{elec}}{S_o}\ . \tag{4-36}$$

The bandwidth Δf_{elec} applies only to those noise sources that are white and cannot be applied to 1/f noise. Although common usage has resulted in calling Δf_{elec} the noise *bandwidth*, it is understood that it is a power equivalency.

If t_{CLOCK} is the time between pixels, the bandwidth for an *ideal* sampled-data system is

$$\Delta f_{elec} = \frac{1}{2\,t_{CLOCK}} = \frac{f_{CLOCK}}{2}\ . \tag{4-37}$$

When equated to electrons,

$$\langle n_{ON\text{-}CHIP} \rangle = \frac{C}{G\,q} V_{ON\text{-}CHIP\text{-}AMP\text{-}NOISE} \sqrt{\Delta f_{elec}} \quad e^- rms\ . \tag{4-38}$$

As the clock frequency increases (for faster readout), $< n_{ON\text{-}CHIP} >$ increases.

4.2.4. OFF-CHIP AMPLIFIER NOISE

The off-chip amplifier noise is identical in form to the on-chip amplifier noise (Equation 4-34). The 1/f knee value may be different for the two amplifiers. If f_{KNEE} is small, the equivalent number of noise electrons is

$$\langle n_{OFF\text{-}CHIP} \rangle = \frac{C}{G G_1 q} V_{OFF\text{-}CHIP\text{-}AMP\text{-}NOISE} \sqrt{\Delta f_{elec}} \quad e^- \ rms \ . \quad (4\text{-}39)$$

4.2.5. QUANTIZATION NOISE

The analog-to-digital converter produces discrete output levels. A range of analog inputs can produce the same output. This uncertainty, or error, produces an effective noise given by

$$V_{NOISE} = \frac{V_{LSB}}{\sqrt{12}} \quad V \ rms \ , \quad (4\text{-}40)$$

where V_{LSB} is the voltage associated with the least significant bit. For an ADC with N bits, $V_{LSB} = V_{MAX}/2^N$. When expressed in equivalent electrons,

$$\langle n_{ADC} \rangle = \frac{C}{G G_1 q} \frac{LSB}{\sqrt{12}} \quad e^- \ rms \ . \quad (4\text{-}41)$$

When the ADC is matched to the amplifier output, V_{MAX} corresponds to the full well. Then

$$\langle n_{ADC} \rangle = \frac{N_{WELL}}{2^N \sqrt{12}} \ . \quad (4\text{-}42)$$

Ideally, $\langle n_{ADC} \rangle$ is less than the noise floor. This is achieved by selecting a high resolution ADC (large N).

Example 4-7
TOTAL NOISE

An array has a noise floor of 20 electrons rms, well capacity of 50,000 electrons, and an 8-bit analog-to-digital converter. Neglecting shot noise and pattern noise, what is the total noise?

The LSB represents $50{,}000/2^8 = 195$ electrons. The quantization noise is $195/\sqrt{12}$ or 56.3 electrons rms. The total noise is

$$\langle n_{SYS} \rangle = \sqrt{\langle n_{FLOOR}^2 \rangle + \langle n_{ADC}^2 \rangle} + \sqrt{20^2 + 56.3^2} \approx 59.7 \; e^- rms \; . \quad (4\text{-}43)$$

The array dynamic is range 2500:1 and the ADC only has 256 levels. Here, the ADC noise dominates the array noise. If a 12-bit ADC is used (4096 levels), the LSB represents 12.2 electrons and $\langle n_{ADC} \rangle = 3.5$ electrons rms. The total noise in electrons is

$$\langle n_{SYS} \rangle = \sqrt{\langle n_{FLOOR}^2 \rangle + \langle n_{ADC}^2 \rangle} = \sqrt{20^2 + 3.52^2} \approx 20.3 \; e^- rms \; . \quad (4\text{-}44)$$

4.2.6. PATTERN NOISE

Fixed pattern noise (FPN) refers to the pixel-to-pixel variation[15,16] that occurs when the array is in the dark. It is primarily due to dark current differences. Significant FPN can be caused by synchronous timing effects at high data rates. It is a signal-independent noise and is additive to the other noise powers. PRNU is due to differences in responsivity (when light is applied). It is a signal-dependent noise and is a multiplicative factor of the photoelectron number.

PRNU be either be specified as a peak-to-peak value or a rms value referenced to an average value. This average value may either be full well or one-half full well value. That is, the array is uniformly illuminated and a histogram of responses is created. The PRNU can either be the rms of the histogram divided by the average value or the peak-to-peak value divided by the average value. This definition varies by manufacturer so that the test conditions must be understood when comparing arrays. For this text, the rms value is used.

Because dark current becomes negligible by efficiently cooling the array, PRNU is the dominate pattern component for most arrays. As a multiplicative noise, PRNU is traditionally expressed as a fraction of the total number of charge carriers. This approach assumes that the detectors are operating in a linear region with the only difference being responsivity differences. If U is the fixed pattern ratio or nonuniformity, then

$$\langle n_{PATTERN} \rangle \approx \langle n_{PRNU} \rangle = U n_{pe} . \tag{4-45}$$

As with shot noise, CTE^N should be added to the equation. Because $CTE^N \approx 1$ for modest sized arrays, it is omitted from most equations. CTE should be included for very large arrays. TDI devices, through their inherent averaging, reduces PRNU by $1/N_{TDI}$. PRNU is usually supplied by the manufacturer. Off-chip gain/level correction algorithms can minimize FPN and PRNU. For system analysis, the corrected pattern noise value is used.

In principle, FPN and the noise floor can be reduced so that the system is photon shot noise limited. This provides the theoretical limit:

$$\langle n_{SYS} \rangle_{MIN} = \sqrt{n_{pe}} . \tag{4-46}$$

However, all systems have some pattern noise. The plot of SNR versus nonuniformity suggests a desired maximum level of acceptable nonuniformity.[17] In Figure 4-13, the noise floor is considered negligible:

$$\langle n_{SYS} \rangle = \sqrt{n_{pe} + (U n_{pe})^2} . \tag{4-47}$$

The total noise value $\langle n_{SYS} \rangle$ increases when PRNU is excessive. Typically an array is selected that has small PRNU. This value can be further reduced by off-chip electronics. The final value should be sufficiently small so that it does not contribute to the overall noise. Because cost increases as the PRNU decreases, the optimum PRNU occurs at the knee of the curve. This occurs when the photon shot noise is approximately equal to PRNU or $U \approx 1/\sqrt{n_{pe}}$. For worst case analysis, the charge well capacity should be used for n_{pe}.

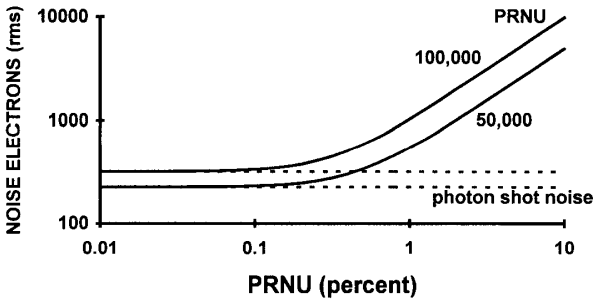


Figure 4-13. System noise as a function of PRNU for two signal levels. The noise floor is considered negligible. The noise floor will increase $\langle n_{SYS} \rangle$ and the location of the knee. Shot noise is independent of PRNU (horizontal dashed lines). The value $1/\sqrt{n_{pe}}$ is 0.4% and 0.3% when n_{pe} is 50,000 and 100,000, respectively. These values represent the "knee" in the curve.

4.2.7. PHOTON TRANSFER

Although various noise sources exist, for many applications it is sufficient to consider photon shot noise, noise floor, and PRNU. Here, the total array noise is

$$\langle n_{SYS} \rangle = \sqrt{\langle n_{SHOT}^2 \rangle + \langle n_{FLOOR}^2 \rangle + \langle n_{PRNU}^2 \rangle} \tag{4-48}$$

or

$$\langle n_{SYS} \rangle = \sqrt{n_{pe} + \langle n_{FLOOR}^2 \rangle + (U n_{pe})^2} \ . \tag{4-49}$$

Either the rms noise or noise variance can be plotted as a function of signal level. The graphs are called the photon transfer curve and the mean-variance curve, respectively. Both graphs convey the same information. They provide[18,19] array noise and saturation level from which the dynamic range can be calculated.

For very low photon fluxes, the noise floor dominates. As the incident flux increases, the photon shot noise dominates. Finally, for very high flux levels, the noise may be dominated by PRNU. As the signal approaches well saturation, the noise plateaus and then drops abruptly at saturation (Figure 4-14). As the charge well reaches saturation, electrons are more likely to spill into adjoining wells (blooming) and, if present, into the overflow drain. As a result, the number of noise electrons starts to decrease. Figure 4-15 illustrates the rms noise as a function of photoelectrons when the dynamic range is 60 dB. Here dynamic range is defined as $N_{WELL}/<n_{FLOOR}>$. With large signals and small PRNU, the total noise is dominated by photon shot noise. When PRNU is large, the array noise is dominated by U at high signal levels.

Dark current shot noise only affects those applications where the signal-to-noise ratio is low (Figure 4-16). At high illumination levels, either photon shot noise or pattern noise dominate. Because many scientific applications operate in a low signal environment, cooling will improve performance. General video and industrial cameras tend to operate in high signal environments and cooling will have little effect on performance. A full SNR analysis is required before selecting a cooled camera.

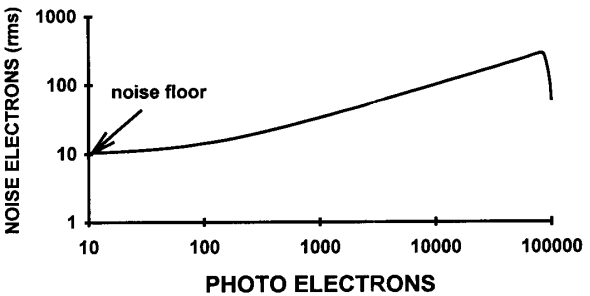
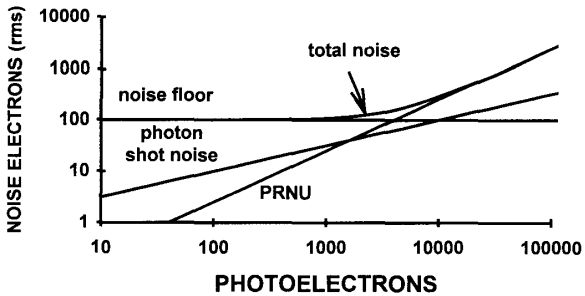
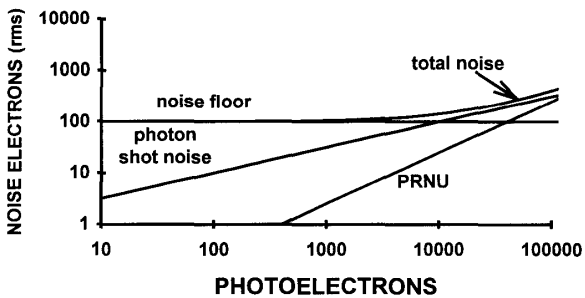


Figure 4-14. Photon transfer curve. At low input signals, the device noise is dominated by the noise floor. As the signal increases, the noise increases due to photon shot noise and pattern noise. As the well reaches saturation, there is a drop in noise electrons due to spill over and well capacity limitation.



(a)



(b)

Figure 4-15. Photon transfer curves for (a) U = 2.5% and (b) U = 0.25%. The charge well capacity is 100,000 electrons and the noise floor is 100 e^- rms to produce a dynamic range of 60 dB. The noise floor, photon shot noise, and PRNU have slopes of 0, 0.5, and 1, respectively. Dark noise is considered negligible. The drop at saturation is not shown (see Figure 4-14).

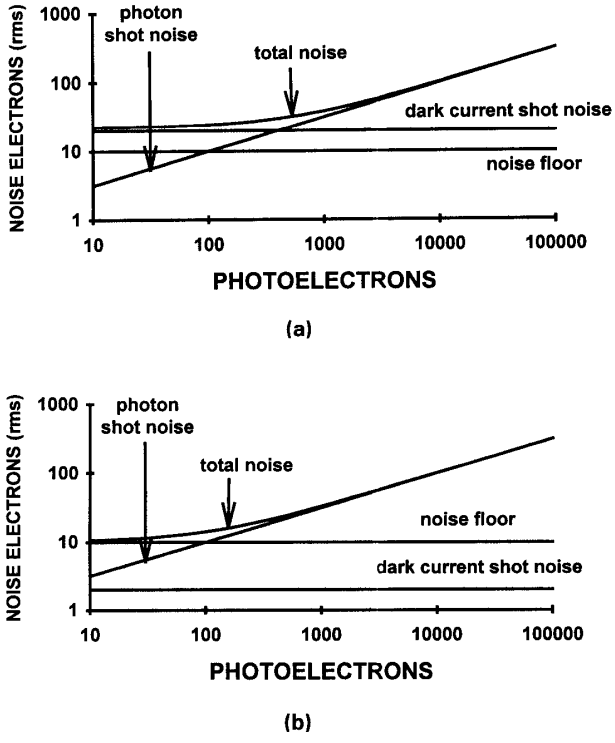


Figure 4-16. Dark current shot noise effects. The noise floor is 10 electrons rms and PRNU is zero. (a) Dark current shot noise is $20e^-$ rms and (b) dark current shot noise is $2e^-$ rms. For large signals, photon shot noise dominates the array noise. The drop at saturation is not shown (Figure 4-14).

Figure 4-17 illustrates the mean-variance technique. In the shot noise limited region,

$$V_{SIGNAL} = \frac{G\,q}{C}\,n_{PE} \tag{4-50}$$

and

$$V_{NOISE}^2 = \left(\frac{G\,q}{C}\right)^2 n_{PE}\ . \tag{4-51}$$

Then the output gain conversion is

$$\frac{G\,q}{C} = \frac{V_{NOISE}^2}{V_{SIGNAL}} \tag{4-52}$$

or the OGC is equal to the noise variance divided by the mean value. It is simply the slope of the shot noise in Figure 4-17.

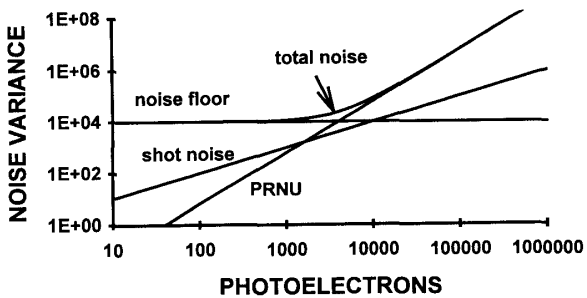


Figure 4-17a. Mean-variance plot which corresponds to Figure 4-15a. The slopes for noise floor, photon shot noise, and PRNU are 0, 1, and 2, respectively.

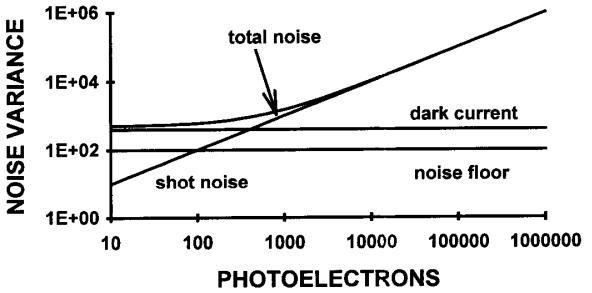


Figure 4-17b. Mean-variance plots which corresponds to Figure 4-16a. The slopes for noise floor, photon shot noise, and PRNU are 0, 1, and 2, respectively.

If both signal and noise are measured after the on-chip amplifier, the measured photogenerated signal is $V_{SIGNAL} = G\,q\,n_{pe}/C$ and the noise is

$$V_{NOISE} = \frac{G\,q}{C}\langle n_{SYS}\rangle \,. \tag{4-53}$$

When $V_{NOISE} = 1$, the signal is C/Gq. The shot noise intercept on the photon transfer curve is 1/OGC (Figure 4-18). The OGC has units of μV/electrons. Similarly, if the values are collected after the ADC, when the shot noise is one, the shot noise signal provides the output gain conversion in units of DN/electrons.

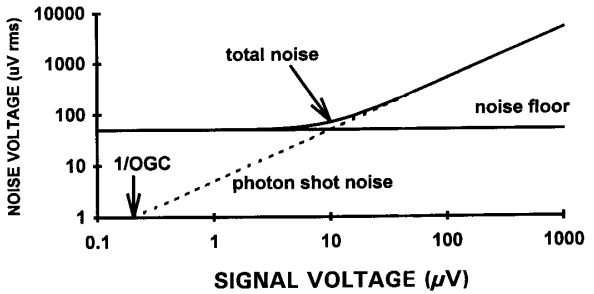


Figure 4-18. Photon transfer curve. The noise floor is 250 e^- and the OGC is 5 μV/e^-. When the photon shot noise is one, the signal is 1/OGC. The drop at saturation is not shown (Figure 4-14).

4.3. ARRAY SIGNAL-TO-NOISE RATIO

The array signal-to-noise ratio is

$$SNR = \frac{n_{pe}}{\sqrt{\left\langle n_{SHOT}^2 \right\rangle + \left\langle n_{FLOOR}^2 \right\rangle + \left\langle n_{PATTERN}^2 \right\rangle}} . \qquad (4\text{-}54)$$

With negligible dark current shot noise,

$$SNR = \frac{n_{pe}}{\sqrt{n_{pe} + \left\langle n_{FLOOR}^2 \right\rangle + (Un_{pe})^2}} . \qquad (4\text{-}55)$$

The charge transfer efficiency factor, CTE^N, should be added to n_{pe} in both the numerator and denominator. For buried channel devices (the most popular) CTE^N is near one and is typically is omitted. Equation 4-55 suggests that the maximum SNR occurs when n_{PE} is a maximum. This occurs when the charge wells are filled. However, as the charge wells fill, electrons can get close to the surface even for buried channel devices. Thus as n_{PE} approaches N_{WELL}, CTE can decrease. The best performance is a compromise between high SNR and high CTE. High CTE is generally only a concern with large devices.

The SNR plots are similar to the photon transfer curves. At low flux levels, the SNR increases with the signal (noise floor is constant). At moderate levels, photon shot noise limits the SNR and at high levels, the SNR approaches 1/U. In Figure 4-19, the SNR is expressed as a power ratio (in dB) which is 20 log(SNR). While the array may have a finite PRNU, it can be minimized off-chip through an appropriate algorithm.

When photon shot noise dominates, the SNR depends only on the photon flux level ($SNR = \sqrt{n_{pe}}$). In Figure 4-19, dynamic range is defined as $N_{WELL}/<n_{FLOOR}>$. The instantaneous dynamic range, which includes all noise sources, is the SNR. The theoretical maximum SNR is $\sqrt{N_{WELL}}$. The SNR is equal to the dynamic range only when the system is noise floor limited. This occurs only for low signal values and small wells (less than about 10,000 electrons). For large well capacity arrays, the actual SNR can never reach the value suggested by the dynamic range. Dynamic range is used to select an appropriate analog-to-digital converter resolution (number of bits). This assures that low contrast targets can be seen.

The array signal-to-noise ratio was derived for a large target that covers many pixels. The noise rides on the signal. The camera SNR, which includes all the noise sources, is provided in Section 6.2.1., *Camera SNR*. In Chapter 12, *Minimum Resolvable Contrast*, the average noise associated with the target and the background is considered.

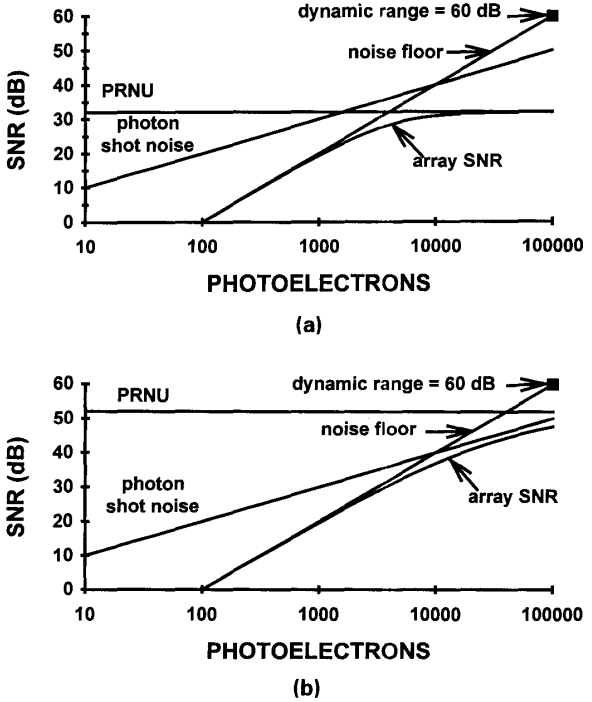


Figure 4-19. SNR as a function of noise for (a) U = 2.5% and (b) U = 0.25%. PRNU, photon shot noise, and the noise floor have slopes of 0, 10, and 20, respectively. Dark current shot noise is considered negligible.

4.4. CORRELATED DOUBLE SAMPLING

After a charge packet is read, the floating diffusion capacitor is reset before the next charge packet arrives. The uncertainty in the amount of charge remaining on the capacitor following reset appears as a voltage fluctuation (see Section 4.2.2., *Reset Noise*, page 127).

Correlated double sampling (CDS) assumes that the same voltage fluctuation is present on both the video and reset levels. That is, the reset and video signals are correlated in that both have the same fluctuation. The CDS circuitry may be integrated into the array package and this makes processing easier for the end user. For this reason it is included in this chapter.

CDS may be performed[20,21] in the analog domain by a clamping circuit with a delay-line processor or by the circuit illustrated in Figure 4-20. When done in the digital domain, a high speed ADC is required that operates much faster than the pixel clock. Limitations of the ADC may restrict the use of CDS in some applications. The digitized video and reset pulses are subtracted and then clocked out at the pixel clock rate.

The operation of the circuit is understood by examining the timing signals (Figure 4-21). When the valid reset pulse is high, the reset switch closes and the signal (reset voltage) is stored on capacitor C_R. When the valid video pulse is high, the valid video switch closes and the signal (video voltage) is stored on capacitor C_V. The reset voltage is subtracted from the video voltage in a summing amplifier to leave the CDS corrected signal. CDS also reduces amplifier 1/f noise.

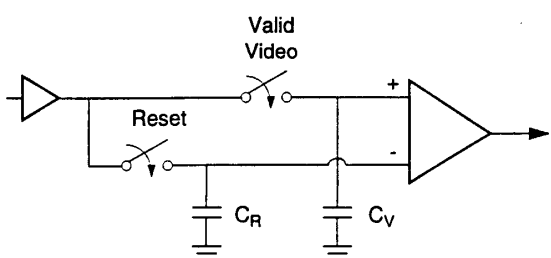


Figure 4-20. Correlated double sampling circuitry. The switches close according to the timing shown in Figure 4-21. (After Reference 20).

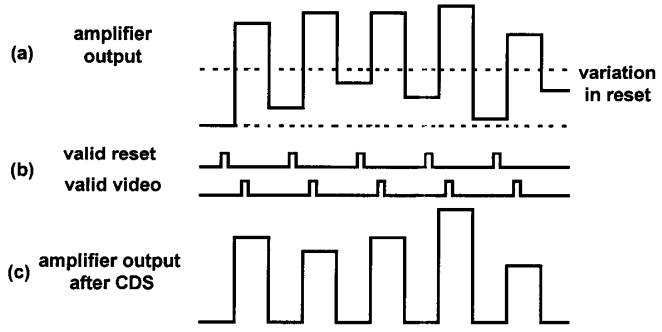


Figure 4-21. Correlated double sampling. (a) Amplifier output signal before CDS, (b) timing to measure video and reset levels, and (c) the difference between the video and reset levels after CDS. The dashed lines indicate the variation in the reset level. The signal level has been inverted compared to Figure 4-12 (page 124).

4.5. FRAME RATES

For a fixed output clock rate, the time to read an array increases as the array size increases. Very large arrays cannot be read in real time. Let the active array size be m × n pixels. Let there be an additional N_{READ} pixels before the sense node. Suppose the clock operates at f_{CLOCK} pixels/s. The minimum time to clock out the array is

$$t_{ARRAY} = \left[\frac{mn + N_{READ}}{f_{CLOCK}} \right] \tag{4-56}$$

and the theoretical frame rate (Figure 4-22) is

$$F_{ARRAY} = \frac{1}{t_{ARRAY}} . \tag{4-57}$$

The maximum frame rate is limited by N_{READ}. The actual frame rate may be less due to electron mobility, on-chip amplifier bandwidth, or clock rate limitations. For some CCDs and all CIDs and CMOS devices, subarrays (also called windows) can be read out. As the subarray size decreases, the frame rate increases, With increased frame rates, the integration time decreases and then more illumination is required to maintain a specific signal-to-noise ratio.

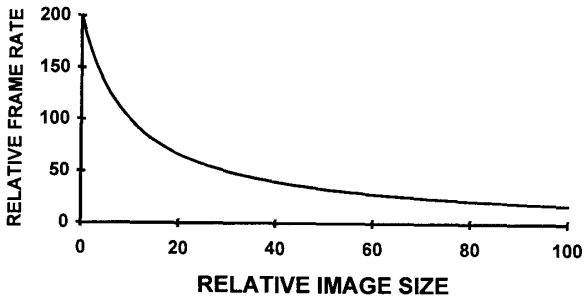


Figure 4-22. Frame rate a function of array size.

4.6. DEFECTS

Large arrays sometimes contain defects. Table 4-5 provides representative definitions. These definitions vary by manufacturer. Arrays with a few defects are more expensive than arrays with a large number. Depending upon the application, the location of the defects may be important. For example, the center ("sweet spot") must be fully operational with increasing defects allowed as the periphery is approached.

Defects are regions of reduced sensitivity. They can also be regions of increased dark current. For scientific application such as spectroscopy, the variation in dark current can affect results - particularly if long integration times are used. These pixels may be labeled as "hot" or "warm." Hot and warm detectors do not follow the same cooling curve as "normal" detectors.[22]

Table 4-5
ARRAY DEFECTS

DEFECTS	REPRESENTATIVE DEFINITION
Point defect	A pixel whose output deviates by more than 6% compared to adjacent pixels when illuminated to 70% saturation.
Hot point defects	Pixels with extremely high output voltages. Typically a pixel whose dark current is 10 times higher than the average dark current.
Dead pixels	Pixels with low output voltage and/or poor responsivity. Typically a pixel whose output is one-half of the others when the background nearly fills the wells.
Pixel traps	A trap interferes with the charge transfer process and results in either a partial or whole bad column (either all white or all dark).
Column defect	Many (typically 10 or more) point defects in a single column. May be caused by pixel traps.
Cluster defect	A cluster (grouping) of pixels with point defects.

4.7. REFERENCES

1. J. Janesick, "Introduction to Scientific Charge-Coupled Devices," Short Course SC-70 presented at the IS&T/SPIE's Symposium on Electronic Imaging: Science & Technology, San Jose CA (1995). J. Janesick presents this course at numerous symposia.
2. R. H. Dyck, "VLSI Imagers," in *VLSI Electronics: Microstructure Science, Volume 5*, N. G. Einspruch, ed., pp. 594, Academic Press, New York (1985).
3. R. H. Dyck, "Design, Fabrication, and Performance of CCD Imagers," in *VLSI Electronics: Microstructure Science, Volume 3*, N. G. Einspruch, ed., pp. 70-71, Academic Press, New York (1982).
4. A. J. P. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices*, pp. 205-210, Kluwer Academic Publishers, Dordrecht, The Netherlands (1995).
5. *The Leader in Spectroscopic Detection*, pp. 33-34, Princeton Instruments Catalog, 3660 Quakerbridge Road, Trenton, NJ 08619 (April 1997).
6. *The Leader in Spectroscopic Detection*, pp. 29-30, Princeton Instruments Catalog, 3660 Quakerbridge Road, Trenton, NJ 08619 (April 1997).
7. C. H. Sequin and M. F. Tompsett, *Charge Transfer Devices*, Academic Press, New York, NY (1975).

8. E. S. Yang, *Microelectronic Devices*, McGraw-Hill, NY (1988).
9. E. L. Dereniak and D. G. Crowe, *Optical Radiation Detectors*, John Wiley and Sons, New York, NY (1984).
10. D. G. Crowe, P. R. Norton, T. Limperis, and J. Mudar, "Detectors," in *Electro-Optical Components*, W. D. Rogatto, pp. 175-283. This is Volume 3 of *The Infrared & Electro-Optical Systems Handbook*, J. S. Accetta and D. L. Shumaker, eds., copublished by Environmental Research Institute of Michigan, Ann Arbor, MI, and SPIE Press, Bellingham, WA (1993).
11. J. R. Janesick, T. Elliott, S. Collins, M. M. Blouke, and J. Freeman, "Scientific Charge-coupled Devices," *Optical Engineering*, Vol. 26(8), pp. 692-714 (1987).
12. T. W. McCurnin, L. C. Schooley, and G. R. Sims, "Charge-coupled Device Signal Processing Models and Comparisons," *Journal of Electronic Imaging*, Vol. 2(2), pp. 100-107 (1994).
13. M. D. Nelson, J. F. Johnson, and T. S. Lomheim, "General Noise Process in Hybrid Infrared Focal Plane Arrays," *Optical Engineering*, Vol. 30(11), pp. 1682-1700 (1991).
14. T. S. Lomheim and L. S. Kalman, "Analytical Modeling and Digital Simulation of Scanning Charge-Coupled Device Imaging Systems," in *Electro-Optical Displays*, M. A. Karim, ed., pp. 551-560, Marcel Dekker, New York (1992).
15. J. M. Mooney, "Effect of Spatial Noise on the Minimum Resolvable Temperature of a Staring Array," *Applied Optics*, Vol. 30(23), pp. 3324-3332, (1991).
16. J. M. Mooney, F. D. Shepherd, W. S. Ewing, J. E. Murguia, and J. Silverman, "Responsivity Nonuniformity Limited Performance of Infrared Staring Cameras," *Optical Engineering*, Vol. 28(11), pp. 1151-1161 (1989).
17. G. C. Holst, *Electro-Optical Imaging System Analysis*, pp. 360-362 and pg. 400-401, JCD Publishing, Winter Park, FL (1995).
18. J. E. Murguia, J. M. Mooney, and W. S. Ewing, "Diagnostics on a PtSi Infrared Imaging Array" in *Infrared Technology XIV*, I. Spiro, ed., SPIE Proceedings Vol. 972, pp. 15-25 (1988).
19. B. Stark, B. Nolting, H. Jahn, and K. Andert, "Method for Determining the Electron Number in Charge-coupled Measurement Devices," *Optical Engineering*, Vol. 31(4), pp. 852-856 (1992).
20. A. J. P. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices*, pp. 228-231, Kluwer Academic Publishers, Dordrecht, The Netherlands (1995).
21. T. W. McCurnin, L. C. Schooley, and G. R. Sims, "Signal Processing for Low-light-level, High-precision CCD Imaging," in *Camera and Input Scanner Systems*, W. Chang and J. R. Milch, eds., SPIE Proceedings Vol. 1448, pp. 225-236 (1991).
22. *The Leader in Spectroscopic Detection*, pp. 31-32, Princeton Instruments Catalog, 3660 Quakerbridge Road, Trenton, NJ 08619 (April 1997).

5

CAMERAS

Solid state camera terminology is a mix of photographic, sensor, image processing, and television terminology. As solid state cameras enter emerging technologies, camera manufacturers adopt the marketplace terminology and add it to their specifications.

Many specifications include the known problems associated with vacuum image tubes such as image burn-in, distortion, or microphonics. These effects are not present in solid state devices. References to these vacuum tube disadvantages probably will disappear in the future as these tubes disappear.

Video standards were originally developed for televisions applications. The purpose of a standard is to ensure that devices offered by different manufacturers will operate together. The standards apply only to the transmission of video signals. It is the camera manufacturer's responsibility to create an output signal that is compatible with the standards. Similarly, it is the display manufacturer's responsibility to ensure that the display can recreate an adequate image.

Some agencies involved with standards include the Institute of Electrical and Electronic Engineers (IEEE), Electronics Industry Association (EIA), EIA-Japan (EIAJ), Joint Electron Devices Engineering Council (JEDEC), Society of Motion Picture and Television Engineers (SMPTE), International Electrotechnical Commission (IEC), Technical Center of the European Broadcast Union (EBU), Advanced Television Systems Committee (ATSC), Comite Consultatif International des Radiocommunications (CCIR) (now ITU-R), and International Telecommunications Union-Radiocommunications(ITU-R). Major manufacturers such as Panasonic, Sony, and JVC influence the standards.

There is no industry wide definition of low light level imaging. To some, it is simply a solid state camera that can provide a usable image when the lighting conditions are about 1 lux. Lower illumination values may be possible by cooling the device. To others it refers to an intensified camera. Intensified cameras are sometimes called low light level television (LLLTV) systems.

Because solid state devices are designed for specific applications, the array and camera use the same terminology: progressive scan chips are used in

progressive scan cameras, TDI arrays are used in TDI cameras, etc. Through appropriate clocking, progressive scan cameras can also provide composite video and be used with consumer displays. Scientific cameras can have both analog and digital outputs. The corresponding sections in Chapter 3, *Solid State Arrays*, should be read with the specific cameras listed here.

In this chapter, camera designs are separated into consumer/broadcast and industrial/scientific applications. The major features of each design are listed in the appropriate sections. However, as technology matures, the distinctions blur. That is, the same camera may be used for a variety of applications. New cameras offer variable frame rates and data rates that can be selected by the user. Therefore, the reader must review all sections before selecting a specific design. Chapter 6, *Camera Performance*, discusses minimum signal, maximum signal, dynamic range, and signal-to-noise ratio.

Both monochrome and color cameras are quite complex as evidenced by the 1478-page *Television Engineering Handbook*, K. B. Benson, ed., McGraw-Hill, New York (1986). This chapter is an overview of solid state camera operation. It highlights those features and specifications that are most often reported in camera data sheets. Together with the preceding chapters, it provides the necessary information to compare and analyze camera systems.

Selecting the right camera to meet specific requirements may be difficult. For consumer video, cost is very important. For broadcast systems, image quality and color are important. Scientific cameras offer low noise and high resolution. Issues not addressed include environmental conditions such as temperature, humidity, altitude, shock, vibration, the "-ilities" (e.g., reliability, maintainability, etc.) and physical attributes (e.g., weight and size).

The symbols used in this book are summarized in the *Symbol List* (page xviii) which appears after the *Table of Contents*.

5.1. CAMERA OPERATION

Figure 5-1 illustrates a generic solid state camera. The lens assembly images the light onto the detector array. General purpose consumer cameras typically have an infrared blocking filter to limit the camera's spectral response to the visible spectral band. Depending upon the application (military, industrial, or scientific), the camera may or may not have this filter. Color cameras have a spectral response limited to the visible region. Cameras may provide digital image processing and these cameras have an internal ADC. Analog cameras will not have an internal ADC. However, they still spatially sample the scene.

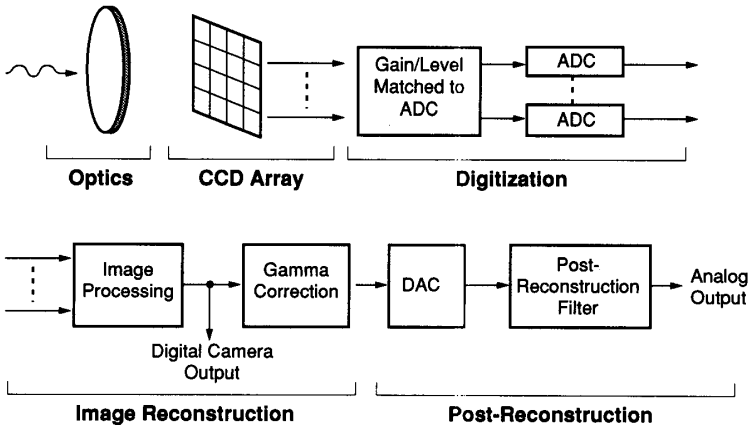


Figure 5-1. Generic functional block diagram. Not all components exist in every camera. The gamma correction circuitry is appropriate when the video is to be viewed on a CRT-based display. Analog cameras will not have an internal ADC.

A single chip color camera typically has an unequal number of red, green, and blue detectors resulting in different sampling frequencies for the different colors (discussed in Section 8.3.2., *Detector Array Nyquist Frequency*). Because the detectors are at discrete locations, target edges and the associated ambiguities are different for each color. A black-to-white edge will appear to have colors at the transition and monochrome periodic imagery may appear to have color. Aliasing is extremely annoying in color systems. A birefringent filter placed between the lens and the array reduces aliasing (discussed in Section 10.3.4., *Optical Anti-Alias Filter*). Any method to reduce aliasing generally reduces the overall MTF.

While aliasing is never considered desirable, it is accepted in scientific and military monochrome applications where the MTF must be maintained. Aliasing is reduced to an acceptable level in consumer color sensors. For specific applications such as medical imaging, it may be considered unacceptable. In medical imagery, an aliased signal may be misinterpreted as a medical abnormality requiring medication, hospitalization, or surgery.

Correlated double sampling reduces reset noise and minimizes 1/f noise. CDS circuitry may be integrated into the array package and this makes processing easier for the end user.

Image processing or image reconstruction reformats the data into signals and timing that are consistent with transmission and display requirements. The advantage of a digital camera is that the output can be directly fed into an image processor, machine vision system, or digital data capture system. If using an analog output camera for machine vision, the video signal must pass through an external analog-to-digital converter. These multiple conversions may modify the imagery (discussed in Section 8.4., *Aliasing in Frame Grabbers and Displays*). Thus, staying within the digital domain ensures maximum image fidelity.

To present a linear transformation from a scene to an observer, a gamma correction (discussed in Section 5.3.3., *Gamma Compensation*) algorithm provides compensation for the monitor's nonlinear response. The concern with gamma should be limited to applications where cathode ray tube (CRT) based displays are used. As the display industry moves from CRTs to flat panel displays, gamma correction is not required. However, gamma provides some image enhancement and therefore will probably be used with flat-panel displays. Gamma correction is generally not preferred for machine vision or scientific applications.

The digital-to-analog converter (DAC) and sample-and-hold circuitry converts the digital data into an analog stair-step signal. The post-reconstruction filter removes the stair-step appearance and creates a smooth analog signal. After the post-reconstruction filter, the spatial scene sampling by the detectors is barely noticeable.

5.2. VIDEO FORMATS

In 1941, the Federal Trade Commission (FCC) adopted the 525-line, 30-Hz frame standard for broadcast television that was proposed by the NTSC (National Television Systems Committee). The standard still exists today with only minor modifications. Although originally a monochrome standard, it has been modified for color transmission. Amazingly, color video fits within the video electronic bandwidth originally set aside for black-and-white television. NTSC initially selected a 60-Hz field rate. It was later changed to 59.94 Hz for color systems to avoid interactions (beat frequencies) between the audio and color subcarrier frequencies. The monochrome standard is often called EIA 170 (originally called RS 170) and the color format is simply known as NTSC

(originally called EIA 170A or RS 170A). The timing for these two standards is nearly the same.[1]

Worldwide, three color broadcast standards exist: NTSC, PAL, and SECAM. NTSC is used in the United States, Canada, Central America, some of South America, and Japan. The phase alteration line (PAL) system was standardized in most European countries in 1967 and is also found in China. Many of the PAL standards were written by the CCIR. As a result, PAL is sometimes called CCIR. In 1967, France introduced SECAM (Sequentiel colour avec mémoire). SECAM is used in France, Russia, and the former Soviet bloc nations. Reference 2 provides a complete worldwide listing. These standards are incompatible for a variety of reasons.

Video transmission bandwidth restrictions and the perception of good imagery on a display dictated the standard formats. That is, camera design is based upon the perception of displayed information. The interlace and scan times were selected to minimize electronic bandwidth and still provide a cosmetically pleasing image to the viewer with no flicker effects. Two fields are interwoven on the display. The eye's persistence time combined with the CRT phosphor persistence blends the two separate images into a perceived full image.

In the NTSC and EIA 170 standards, a frame is composed of 525 scan lines separated into two sequential fields (even and odd). Because time is required for vertical retrace, only 485 lines are displayed. There are 242.5 active lines per field. Because field 1 ends in a half-line, it is called the odd field. Figure 5-2 illustrates how the information is presented on a display. For commercial television, the aspect ratio (horizontal to vertical extent) is 4:3.

When a camera is designed to be compatible with a transmission standard, the manufacturer uses the broadcast terminology as part of the camera specifications. Similarly, consumer display manufacturers use the same terminology.

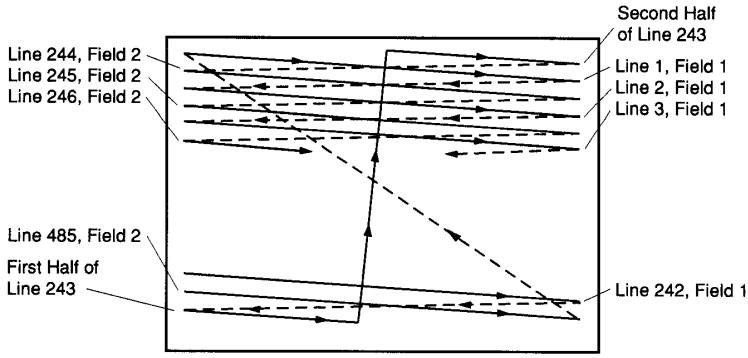


Figure 5-2. Presentation of EIA 170 video on a display. Standard timing is defined for video transmission only. Solid state arrays may not read out data in this precise order. The camera image processing algorithm formats the data into a serial stream that is consistent with the video standard.

5.2.1. VIDEO TIMING

Table 5-1 provides the relevant parameters for the monochrome and color video standards. The standards specify the line frequency and blanking period. The scanning frequency is the frame rate multiplied by the number of lines. The total line time is calculated from the line frequency (e.g., $1/15{,}750 = 63.492\,\mu s$) and the active line time is the total line time minus the horizontal blanking period. The blanking period has a tolerance of about $\pm 0.25\%$ so that the active time line also varies slightly. Table 5-1 provides the minimum active time line where the maximum blanking period has been selected.

Table 5-1
STANDARD VIDEO TIMING

FORMAT	FRAME RATE (Hz)	LINES PER FRAME	ACTIVE LINES	SCANNING FREQUENCY (kHz)	TOTAL LINE TIME (μs)	MINIMUM ACTIVE LINE TIME (μs)
EIA 170	30	525	485	15.750	63.492	52.092
NTSC	29.97	525	485	15.7343	63.555	52.456
PAL	25	625	575	15.625	64.0	51.7
SECAM	25	625	575	15.625	64.0	51.7

For the monochrome video transmission (See Section 5.2.2., *Broadcast/Non-broadcast Design*), the picture element was originally defined[3] as the smallest area being delineated by an electrical signal. Note that this is the definition of a resel (analog element) and it is referenced to an imaginary pixel. The impulse response to an ideal band-limited system is

$$A(t) = \frac{\sin(2\pi f_c t)}{2\pi f_c t} , \qquad (5\text{-}1)$$

where f_c is the video bandwidth cutoff frequency. The pixel length is defined as the minimum distance between two impulses such that the two are still discernible. This occurs when the maximum of one falls on the first minimum (zero) of the next. The separation of these two vertical lines in time is $\Delta t = 1/2f_c$. This is the smallest possible dimension in a linear system. For the video standard, the number of resels ("pixels") per line is

$$N = \frac{active\,line\,time}{\Delta t} = 2f_c(active\ line\ time) . \qquad (5\text{-}2)$$

The nominal[4] video bandwidths of NTSC, PAL, and SECAM are 4.2, 5.5, and 6 MHz, respectively. Table 5-2 lists the nominal number of horizontal elements per line. The quoted resolution (number of horizontal pixels) varies in the literature and depends upon the precise bandwidth and active line time selected by the author. Multiplying the pixel number by the number of active lines yields the elements per frame.

Table 5-2
NOMINAL NUMBER OF VIDEO RESELS ("PIXELS")
Based upon electrical transmission bandwidth only

FORMAT	NOMINAL BANDWIDTH (MHz)	ACTIVE LINE TIME (μs)	NUMBER of HORIZONTAL ELEMENTS	NUMBER of ACTIVE LINES	ELEMENTS PER FRAME
EIA 170	4.2	52.09	437	485	212,000
NTSC	4.2	52.45	440	485	213,000
PAL	5.5	51.7	569	575	327,000
SECAM	6.0	51.7	620	575	356,500

The full bandwidth is somewhat wider than the nominal value but high frequency components may be attenuated. For example, NTSC has a full bandwidth of approximately 4.5 MHz. The total bandwidth allocated for sound, chrominance, and luminance is 6, 8, and 8 MHz for NTSC, PAL, and SECAM, respectively. Using the allocated bandwidth, the maximum number of horizontal pixels is 625 for NTSC and 827 for PAL and SECAM.

Current camera design philosophy matches the number of detectors to the number of pixels supported by the video bandwidth. Based upon electrical bandwidth considerations, it appears that the number of horizontal detectors should range from 440 to 625 for NTSC, 560 to 827 for PAL, and 620 to 827 for SECAM compatible cameras.

The actual number of detector elements depends upon the manufacturer's philosophy. Image processing algorithm complexity is minimized when the spatial sampling frequency is equal in the horizontal and vertical directions. The 4:3 aspect ratio suggests that there should be 4/3 times more detectors in the horizontal direction compared to the vertical direction. If the number of elements in the vertical direction is simply equal to the number of active scan lines, then the number of horizontal detectors should be about 646 for NTSC, and 766 for PAL and SECAM.

The number of pixels per line is often chosen so that the horizontal clock rate is an integer multiple of the color subcarrier. For NTSC, 768 pixels equate to four times the NTSC color subcarrier (see Section 3.11.3., *Number of Detectors*, page 97).

This larger number provides more sampling for each color and reduces color aliasing. For single chip color cameras, resolution is more difficult to define. The primaries or their complements spatially sample the scene unequally (discussed in Section 8.3.2., *Detector Array Nyquist Frequency*). If the chip contains 768 detectors horizontally, then the number of green, red, and blue detectors is typically 384, 192, and 192, respectively. Because the eye is most sensitive to luminance values and the green component represents luminance, then the green resolution may be taken as the camera resolution. Most chips contain more green detectors and therefore the resolution is more than just one-third of an equivalent monochrome system. Through data interpolation, any number of output data points can be created. This number should not be confused with the camera spatial resolution. Single chip color cameras cannot provide the resolution listed in Table 5-2.

To reduce cost, a camera may contain fewer detector elements. Here, the full bandwidth of the video standard is not exploited. The imagery of these cameras will be poorer compared to those cameras that fully use the available bandwidth. Recall that the standard only specifies the timing requirements and a maximum bandwidth. The timing ensures that the camera output can be displayed on a standard display.

For the complete system, the horizontal resolution may be limited by the detector array, bandwidth of the video standard, or by the display. Display resolution is discussed in Section 7.5., *Resolution*. It is common for the display to limit the overall system resolution.

For NTSC and EIA 170 compatible arrays, it is common to have about 480 detectors in the vertical direction. The image formatting algorithm inserts blank lines (all black) to meet video standards (485 lines). Because many displays operate in an overscan mode, the blank (black) lines are not noticed.

Resolution is often expressed in TV lines. This visual psychophysical measure was developed to specify the resolution of monitors. Because of phasing effects, the number of individual horizontal lines that can be perceived, on the average, is less than the number of scan lines. The ratio of the average number to the total number is the Kell factor[5] and is assumed to be 0.7. Thus, the vertical resolution is 340 lines for NTSC and 402 lines for PAL and SECAM. This vertical resolution applies to the monitor capability and not the camera. However, the camera output must be displayed on a monitor so that the Kell factor is sometimes included in camera specifications.

For higher vertical resolution, more lines are required. EIA 343A (originally RS 343A) was created[6] for high resolution closed circuit television cameras (CCTV). Although the standard encompasses equipment that operates from 675 to 1023 lines, the recommended values are 675, 729, 875, 945, and 1023 lines per frame (Table 5-3). The military primarily uses 875 lines/frame but it requires "standards conversion" boxes that can record and present imagery on EIA 170 compatible equipment. EIA 343A is still a monochrome standard.

Table 5-3
RECOMMENDED EIA 343A VIDEO FORMATS

LINES/FRAME	ACTIVE LINES	SCANNING FREQUENCY (kHz)	LINE TIME μs	ACTIVE LINE TIME μs
675	624	20.25	49.38	42.38
729	674	21.87	45.72	38.72
875	809	26.25	38.09	31.09
945	874	28.35	35.27	28.27
1023	946	30.69	32.58	25.58

5.2.2. BROADCAST/NON-BROADCAST DESIGN

The EIA 170 standard was developed for the *broadcasting* of video signals over limited bandwidth channels. The FCC dictated that the bandwidth cannot exceed 6 MHz. To avoid flicker on the television receiver, a 2:1 interlace concept was adopted. The NTSC standard was constrained to the same bandwidth limitation.

For compatibility, all broadcast cameras and television receivers are designed to the timing and bandwidth of the broadcast standard. Using a wider bandwidth receiver does not improve image quality or resolution because it is limited by the 6 MHz bandwidth. An inexpensive receiver may have reduced bandwidth (as do many VHS recorders) and this reduces image quality. However, to function, this "lower quality" receiver still must operate under the *timing* requirements of the video standard.

Table 5-2 (page 153) listed the number of video resels supported by the video standard. If the camera has more pixels, then the received image resolution will be limited by the standard's bandwidth. If the camera has fewer

pixels, then the camera's resolution limits image quality. By using image compression, high definition television cameras can effectively transmit more resels than that suggested in Table 5-2. The number of transmitted resels remains the same (constrained by bandwidth). It is through coding (e.g., image compression) that the resels contain more information.

Because EIA 170 and NTSC are so pervasive, nearly all cameras are built to these video standards. This permits saving imagery on off-the-shelf video recorders and displaying it on just about any monitor. However, for non-broadcast operation, it is not necessary to adhere to the 6 Mhz bandwidth limitation. The only requirement is that the timing be identical to the standard timing. A high resolution monitor (wide bandwidth) can display the imagery from a camera that has more pixels and a wide bandwidth recorder can store that imagery.

Departure from the standard video timing requires special monitors and special recorders. Already digital recorder and digital monitors exist. Computer monitors now operate in the progressive scan mode. To avoid flicker, the frame rate exceeds 70 Hz. The frame rate is determined by the video drive electronics and is not generally user-selectable.

To see an image on a monitor, the camera's timing must be consistent with the monitor requirements. Because the video standards are popular, most cameras adhere to the appropriate timing. With progressive-scan cameras, the camera's internal electronics must reformat the data into an interlace mode. If the camera's output is fed directly into a computer, timing is no longer a concern. If digital data are transferred, only a sync signal is necessary to indicate the start of the data. If analog data are transferred, then the frame capture board should have a resolution that is greater than the camera's. That is, there should be more datels than pixels.

5.2.3. COMPONENT/COMPOSITE SIGNALS

The term composite is used to denote a video signal that encodes luminance, color, and timing information onto a single channel. It can be transmitted on a single wire such as cable television. NTSC, PAL, and SECAM are composite video signals. The signal must be decoded to create an image.

Component analog video (CAV) carries each part on a separate wire. RGB, or their complementary values are carried on three separate cables. It is common for the G to be a composite line and carry the sync signals. Component videos

do not require encoding and therefore there is less chance that there will be color or luminance cross coupling. This implies that component imagery will produce a higher quality image.

As equipment is miniaturized (e.g., hand-held video recorders) and resolution increased (e.g., HDTV), there seems to a proliferation of recommended standards.[7] Most standards deal with color encoding (chrominance encoding). The luminance channel is identical with the green channel and is labeled Y. The other signals are color and the exact composition depends upon the standard. SECAM uses the primaries whereas the other formats use the complements. The chrominance values are related through matrices but the signals are not compatible due to different line frequencies and sync signals. Because the eye is most sensitive to luminance, the luminance channel has the widest bandwidth and the chrominance channels have smaller bandwidths.

Matrixing can convert the RGB into any color coordinate system. With the ease of digital implementation, multiple outputs are available from many cameras. That is, a camera may offer RGB, NTSC, and Y/C outputs to provide flexibility to the user.

5.2.4. IRE UNITS

The Institute of Radio Engineers (IRE) assigned 140 IRE units to the peak-to-peak composite signal (Figure 5-3). This provides 100 IRE units for the active image signal and 40 IRE units for blanking. They assumed that the video is one volt peak-to-peak when terminated by 75 Ω. Thus, each IRE unit is equal to 1/140 = 7.14 mV.

To avoid seeing the retrace, NTSC introduces a small setup. With most systems, the video is usually clamped at the blanking level and the video signal with setup does not contain a reference black signal. That is, the video does not contain a signal to indicate the black level. This voltage only occurs when there is a black object within the field-of-view. The blanking level (retrace) is sometimes called "blacker-than-black" because it is 7.5 IRE units below the black level. With the signal standardized to 1 V peak-to-peak, the setup reduces the active video to 0.66 v.

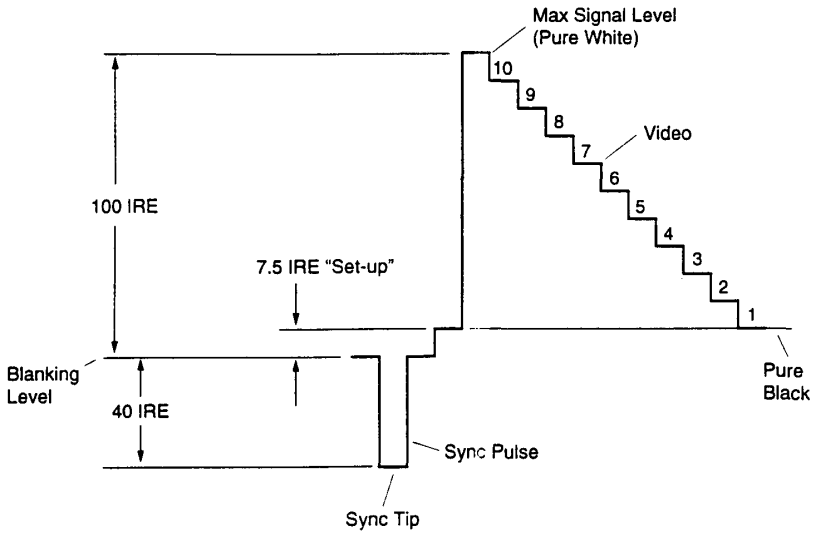


Figure 5-3. Definition of IRE units. The active video represents a gray scale divided into 10 equal steps. NTSC assigns 7.5 IRE units to setup.

By convention, the blanking level is assigned a value of O v so that the active video is positive while the sync pulse is negative. The peak-to-peak amplitude remains at 1 V. As new standards are introduced,[7] the voltage levels associated with IRE units change (Table 5-4). The composite signal with setup (NTSC) can introduce ambiguity in signal level because there is no fixed black reference. Without setup, radiometrically correct imagery is ensured and black-to-white measurements can be performed with higher accuracy. SMPTE and EBU have recommended the non-NTSC format for color component video.

When a video signal is expressed as a percentage, it is numerically equal to the IRE units. 80% video is the same at 80 IRE units. For example, 80% video is (0.8) (0.714) = 571 mV with the NTSC standard.

Table 5-4
VIDEO STANDARDS

FORMAT	VOLTS			IRE UNITS			
	Sync tip	Setup	Peak	Synch tip	Blanking	Setup	Peak
NTSC	-0.286	0.0536	0.714	-40	0	7.5	100
Composite without setup	-0.286	-	0.714	-40	0	-	100
Non-NTSC	-0.300	-	0.700	-40	0	-	100

5.2.5. DIGITAL TELEVISION

Digital image processing is already present in most transmitters and receivers. It seems natural to transmit digitally. The major advantage is that digital signals are relatively immune to noise. Either a bit is present or it is not. A digital voltage signal, which may have been reduced in value and had noise added, can be restored to its full value with no noise in the receiver. An all-digital system avoids the multiple analog-to-digital and digital-to-analog conversions that degrade image quality.

Digital signal "transmission" was first introduced into tape recorders. Because a bit is either present or not, multiple generation copies retain high image quality. In comparison, analog recorders, such as VHS, provide very poor quality after just a few generations. The first digital recorder used a format called D-1 and it became known as the CCIR 601 component digital standard. Digital recording formats[8] now include D-1, D-2, D-3, D-4, and D-5.

The nominal 4.2-MHz bandwidth required by NTSC is not a continuous spectrum but consists of many bands due to the interaction of the line frequency (59.94 Hz) and scan frequency (15.7343 KHz). It is this discrete nature that allows the interleaving of chrominance, luminance, and audio within the allocated bandwidth. A comb filter, present in the receiver, separates the signals to minimize the cross talk that appears as aliased color. The composite video signal can be equated to a signal whose sampling frequency is 3.58 MHz for NTSC and 4.43 MHz for PAL. SECAM has different effective sampling rates for each color difference: 4.25 MHz and 4.41 MHz. These are the subcarrier frequencies.[9]

The most popular sampling frequency is four times the subcarrier frequency and is simply called $4f_{sc}$ sampling. It is equal to 14.3 MHz for NTSC and 14.7 MHz for PAL. This allows the post-reconstruction filter to have a more gradual cutoff with less ripple. This filter is easier to design and is less expensive. Because the data exist at discrete frequencies, strict adherence to the Nyquist criterion (discussed in Chapter 8, *Sampling Theory*) is not required.

In component video there is no subcarrier. A sampling frequency of 13.5 MHz has been selected to accommodate both 525 and 625 line formats. The luminance channel is sampled at 13.5 MHz and the two color difference channels are sampled at 6.75 MHz each. Recall that there is less detail in the color channels so they are sampled at a lower rate. If 13.5 MHz is four times the frequency of some imaginary subcarrier, then 6.75 MHz is twice that same imaginary frequency. This is known as 4:2:2 digital component television. If the three signals are multiplexed into a serial stream, the new sample rate is 27 MHz ($13.5 + 6.75 + 6.75 = 27$). In the multiplexed form, component signals can be transmitted on a single cable.

An image, whose amplitude is digitized into 8 bits, can provide a cosmetically acceptable image for consumer camcorders. However, nonlinear circuitry is required to minimize the visibility of quantization contours at low signal levels. Otherwise, more bits are necessary. Photo quality imagery requires 12 or more bits. For composite signals, the sync pulse must also be digitized. The actual number of bits available to the imagery is conservatively less. CCIR 601 assigns 16 DN to the black level and 235 DN to the white level. The remaining values (1 to 15 and 236 to 254 DN) allow room for overshoots and control of the white and black levels. The values 0 and 255 DN are used for syncs.

5.2.6. HDTV/ATS

The current accepted meaning of high definition television (HDTV) is a television system providing approximately twice the horizontal and vertical resolution of present NTSC, PAL, and SECAM systems. While the term HDTV implies a receiver, the industry is equally concerned about transmission standards. The video format must be standardized before building cameras and receivers. The acronym HDTV is slowly being replaced with the advanced television system or ATS.

HDTV is envisioned as a system that provides high resolution on large displays and projection screens. Its resolution will not be fully appreciated in the

typical living room where the sofa is nine feet from the television (discussed in Section 7.1., *The Observer*). The Japan Broadcasting Corporation (NHK) provided the first HDTV demonstration in 1981. The aspect ratio was initially 5:3 but this was later changed to 16:9 to match the various aspect ratios used in motion pictures.

The goal is to have worldwide compatibility so that HDTV receivers can display NTSC, PAL, and SECAM transmitted imagery. With today's multimedia approach, any new standard must be compatible with a variety of imaging systems ranging from 35-mm film to the various motion picture formats. The desire is to create composite imagery[10,11] and special effects from a mixture of film clips, motion pictures, and 35-mm-film photos. Compatible frame rates allow the broadcasting of motion pictures and the filming of television.

As various companies entered the HDTV arena, they proposed different transmission formats[12]. To standardize the format, the Grand Alliance was created in May 1993. The seven members are: AT&T, David Sarnoff Research Center, General Instrument Corporation, MIT, North American Philips, Thomson Consumer Electronics, and Zenith Electronics. The Grand Alliance collaborates with a variety of agencies involved with standards (see list on page 146).

In April 1994, the Grand Alliance approved two different resolutions for HDTV (Table 5-5). Both formats have square pixels. The selected line rate should simplify conversion from NTSC and PAL into HDTV. The lines per frame of 1125 is 15/7 times 525 and 9/5 times 625. Similarly, 787.5 is 3/2 times 525 but is approximately 100/79 times 625 lines. Thus, with the lower line rate, it is more difficult to convert PAL into HDTV.

Doubling the resolution, increasing the aspect ratio, and separating color from luminance requires an increased bandwidth of eight-fold over current television requirements. With a nominal bandwidth of 4.2 MHz, the required HDTV bandwidth is approximately 33.6 MHz. A 10-bit system will have a data rate of 336 Mbit/s. At this data rate, transmission losses in conventional wire cable limit the maximum transmission distance to, perhaps, 1 km. This demands image compression.

Table 5-5
GRAND ALLIANCE RECOMMENDED FORMATS
(Recommended in April 1994)

RESOLUTION Pixels/line × active lines	LINES PER FRAME	FRAME RATE (Hz)	SCAN TECHNIQUE
1920 × 1080	1125	24	Progressive
		30	Progressive
		60	Interlaced
1280 × 720	787.5	24	Progressive
		30	Progressive
		60	Progressive

The FCC has allocated only a 6-MHz bandwidth for commercial television channels. Thus, image compression somewhere between 50:1 and 75:1 is required. The Grand Alliance recommends using MPEG-2 image compression. Although current cameras are analog devices,[13] the Grand Alliance recommended digital transmission.

In December 1996, the FCC adopted a standard proposed by the Advanced Television Systems Committee. This versatile standard defines four basic digital formats (Table 5-6). The higher resolution formats will probably be used to broadcast movies while news and commercials will be broadcasted at the lower resolution formats. While the standard allows both interlaced and progressive scan formats, only one will probably survive. It also mandates compressed digital transmission. The conversion from analog to digital television is a complex process.[14] Broadcasting should begin November 1, 1998. The status of HDTV/ATS cameras, standards, and receivers will appear in many journals over the next few years.

Table 5-6
PICTURE FORMATS SUPPORTED by ATSC STANDARD

RESOLUTION pixels/line × active lines	FRAME RATE (Hz)	ASPECT RATIO	SCAN TECHNIQUE
1920 × 1080	24	16:9	Progressive
	30	16:9	Progressive
	60	16:9	Interlaced
1280 × 720	24	16:9	Progressive
	30	16:9	Progressive
	60	16:9	Progressive
704 × 480	24	16:9 or 4:3	Progressive
	30	16:9 or 4:3	Progressive
	60	16:9 or 4:3	Progressive
	60	16:9 or 4:3	Interlaced
640 × 480	24	4:3	Progressive
	30	4:3	Progressive
	60	4:3	Progressive
	60	4:3	Interlaced

5.3. CONSUMER/BROADCAST CAMERAS

Figure 5-4 illustrates a typical color camera functional block diagram. The array output is processed to achieve the standard video signals. This section describes the knee, color correction, gamma, and aperture correction. Monochrome cameras tend to be used for scientific and industrial applications and are described in Section 5.4., *Industrial/Scientific Cameras*.

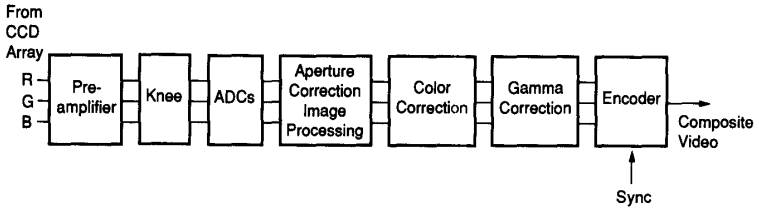


Figure 5-4. Typical color camera functional block diagram. Sync signals and other control logic are not shown. The subsystem location and design vary by manufacturer. Consumer cameras typically have a composite output that is created by the encoder. Scientific and industrial cameras may have a component output which is the signal before the encoder. To increase versatility, many cameras can provide either NTSC, RGB, or complementary color output. The RGB output provides easy interfacing with a computer monitor.

5.3.1. THE KNEE

The antibloom drain (see Section 3.8., *Antibloom Drain*, page 85) creates a knee in the linear response to brightness. The knee slope and location can be any value (Figure 5-5). While scientific cameras operate in a linear fashion, there is no requirement to do so with consumer cameras. Consumer camera manufacturers will provide a mapping that they think will provide a cosmetically pleasing image. The knee also simplifies requirements in that an 8-bit ADC can provide aesthetically pleasing imagery over a wide input dynamic range.

Scientific and industrial applications often require a linear input-output system. For example, linearity is required for radiometric analysis. Interestingly, as some manufacturers add more knee control, others are actively designing algorithms to "de-knee" the signal before image processing.

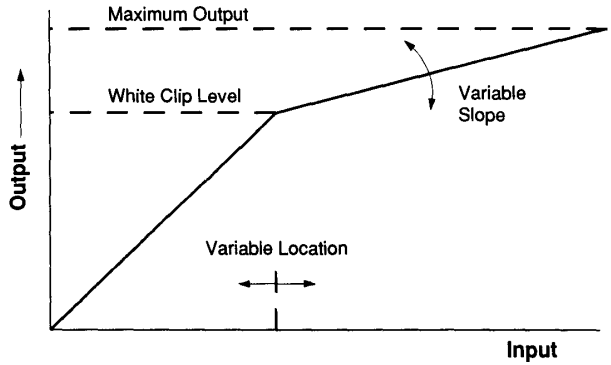


Figure 5-5. The knee extends the dynamic range by compressing high intensity signals. By decreasing the knee slope, the camera can handle a larger scene dynamic range.

5.3.2. COLOR CORRECTION

Figure 5-4 illustrated a typical color camera block diagram. The "color" signals may have originated from a single CFA (see Section 3.11.2., *Color Filter Arrays*, page 94) or three separate arrays as illustrated in Figure 5-6. Because of high cost, three-chip cameras were used almost exclusively for high quality broadcast applications. Now, three-chip color cameras are entering the scientific market. Consumer camera designs, driven by cost, tend to use only one chip. Although this is changing.

The color correction circuitry is an algorithm that can be represented as a matrix

$$\begin{pmatrix} R_2 \\ G_2 \\ B_2 \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \begin{pmatrix} R_1 \\ G_1 \\ B_1 \end{pmatrix}. \qquad (5\text{-}3)$$

When used to improve color reproducibility (i.e., matching the camera spectral response to the CIE standard observer), the process is called masking.[15,16] When manipulating signals into the NTSC standard, it is called matrixing.

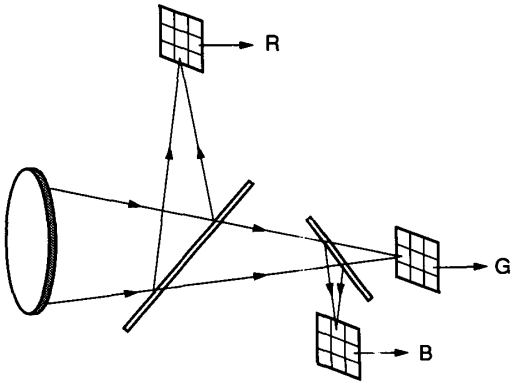


Figure 5-6. A high quality color camera has three separate arrays. One for each primary color or its complement. Light is spectrally separated by coatings on the beam splitters and additional filters. The three-chip camera is listed as CCD×3 or 3CCD in sales brochures. Each color samples the scene equally and this minimizes color moiré patterns when compared to a CFA. Although beam splitters are shown, most 3CCD cameras use prisms for compactness.

Matrixing can be used to alter the spectral response and intensities of the primaries to compensate for changes in scene color temperature. Intensity changes are achieved simply by multiplying a primary by a constant. Spectral changes are approximated by adding various ratios of the primaries together. For example, if an object is viewed at twilight, the matrix can be adjusted so that the object appears to be in full sun light (neglecting shadows). A scene illuminated with a 2600-K source may appear somewhat yellow compared to a scene illuminated with a 6500-K source. This occurs because a 2600-K source has less blue than a 6500-K source. With white balance (matrixing) a white object can appear to be illuminated by different blackbodies - usually ranging from 3200 to 5600 K.

Gamma correction is introduced into each of the three color channels. By convention, the analog signals are represented by E_G, E_R, and E_B. After gamma correction, they are represented by E_G', E_R', and E_B'.

With CFAs, detectors sensitive to the same color are separated resulting in a low fill factor. The optical anti-alias filter blurs the image and effectively increases the optical area of the detector. The birefringent filter is placed between the lens assembly and the detector array. Birefringent crystals break a beam into two components: the ordinary and the extraordinary. This makes the detector appear larger optically. While this changes the MTF (discussed in Section 10.3.4., *Optical Anti-Alias Filter*), the sensitivity does not change. The light that is lost to adjoining areas is replaced by light refracted to the detector.

With unequal sampling, monochrome edges may appear colored and small objects may change color. Sophisticated image processing is required to overcome the unequal sampling of the colors and to format the output into the NTSC standard. The algorithm selected depends upon the CFA architecture (see Figure 3-44, page 61). Algorithms may manipulate the ratios of blue-to-green and red-to-green,[17] separate the signal into high and low frequency components,[18] use a pixel shift processor,[19] or add two alternating rows together[20,21] to minimize color artifacts. The decoding algorithm is unique for each CFA configuration. Clearly, each CFA design requires a new decoding algorithm. Algorithm design is probably more of an art than a science because the acceptability of color aliasing is not known until after the camera is built.

Example 5-1
SINGLE COLOR CHIP ALGORITHM

Figure 5-7 illustrates a typical CFA[21] where rows are added to minimize aliasing. The NTSC odd field consists of CCD rows 1 + 2, 3 + 4, etc., and the NTSC even field consists of 2 + 3, 4 + 5, etc. Tables 5-7 and 5-8 provide the outputs.

Table 5-7
SUMMING TWO ROWS

CCD ROW	OUTPUT LINE
1 + 2	1 (odd field)
3 + 4	3 (odd field)
2 + 3	2 (even field)
4 + 5	4 (even field)

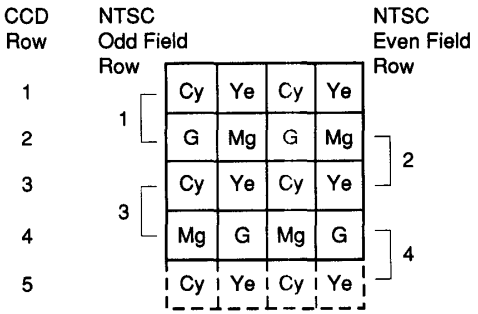


Figure 5-7. Typical CFA (From Reference 21).

Table 5-8
OUTPUT VALUES

TV LINE	PIXEL 1	PIXEL 2	PIXEL 3	PIXEL 4
1	Cy + G = 2G + B	Ye + Mg = 2R + G +B	Cy + G = 2G + B	Ye + Mg = 2R + G +B
3	Cy + Mg = R + G +2B	Ye + G = R + 2G	Cy + Mg = R + G +2B	Ye + G = R + 2G
2	Cy + G = 2G + B	Mg + Ye = 2R + G + B	Cy + G = 2G + B	Mg + Ye =2R + G + B
4	Mg + Cy = R + G +2B	Ye + G = R + 2G	Mg + Cy = R + G +2B	Ye + G = R + 2G

The NTSC signal requires equal R, G, and B information on each line and this requires line interpolation. For example, pixel 3 on the first line does not contain any red data. Red must be interpolated from adjoining pixels. The algorithm selected depends upon the mathematics and the human eye response. The selection depends upon the resultant imagery. That is, the interpolation tends to be empirical.

5.3.3. GAMMA COMPENSATION

A consumer video camera may be purposely made nonlinear to match a monitor response. Gamma correction optimizes the displayed image when using CRT-based monitors. The input predistortion (created at the camera) can only be effective over a strictly defined range. The overall system is linear only when the correct display unit is used. Gamma correction is nonlinear and should be inserted after any required image processing. As flat panel displays replace CRT-based displays, the concern with gamma correction should disappear.

The luminous output of a cathode ray tube is related to its grid voltage by a power law

$$L_{DISPLAY} = K(V_{GRID})^{\gamma} . \tag{5-4}$$

If the camera has the inverse gamma (Figure 5-8), then the radiometric fidelity is preserved at the display.

$$L_{DISPLAY} = K(V_{GRID})^{\gamma} = K\left(V_{SCENE}^{\frac{1}{\gamma}}\right)^{\gamma} = K V_{SCENE} . \tag{5-5}$$

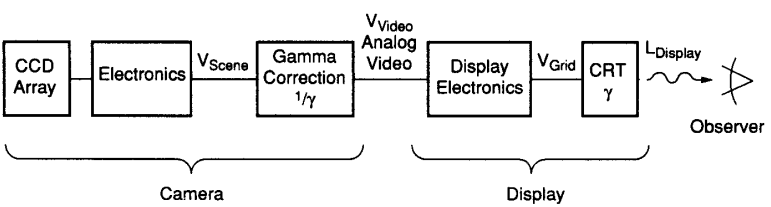


Figure 5-8. Gamma correction should only be used when the imagery is directly presented on a CRT-based display.

Although gamma correction circuitry is placed in the camera and it is numerically equal to the inverse of assumed CRT gamma, popular parlance simply calls it gamma. From now on, gamma will mean the inverse of the CRT gamma value. Gamma compresses small signals and some black detail is lost. In certain high contrast scenes, such as scenes with direct sunlight and shaded areas, the darker parts will lack contrast. It also changes the noise characteristics depending upon the scene level.

Figure 5-9 illustrates a machine vision system or a system where the imagery is evaluated by a computer. Gamma correction should be inserted before any image processing. Although standards recommend a gamma correction value, manufacturers modify these values to create what they perceive to be a cosmetically pleasing image. The output of an imaging processing work station is shown on a computer monitor. Although a machine vision system does not require a monitor, it adds convenience. A gamma correction, internal to the computer, assures that the correct image is shown on the computer monitor.

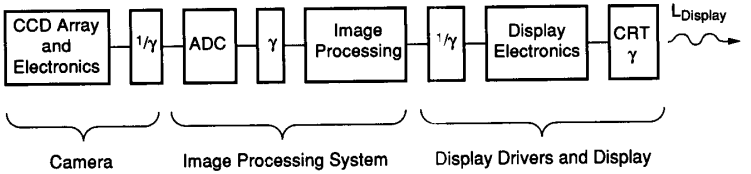


Figure 5-9. Image processing is simplified when gamma has been removed. The camera gamma should be one if the image is to be digitized and displayed on a computer monitor. Although not required for machine vision applications, the monitor offers convenience.

Although CRT displays have a range of gammas, NTSC, PAL, and SECAM "standardized" the display gamma to 2.2. 2.8, and 2.8, respectively. However, all three systems use 0.45 as the inverse. For NTSC, the monitor luminance is proportional to the detector output and target intensity $(2.2)(0.45) = 1$. With PAL and SECAM, the luminance still has[22] a residual gamma of 1.27. This provides some contrast enhancement.

The gamma slope becomes very large for low values of input brightness. It has therefore become accepted practice to limit the operation of the gamma compensator to an arbitrary level. SMPTE Standard 240M was developed[23] for high definition televisions but now is applied to many systems It has since been replaced by ITU 709.

Normalized to one, 240M states

$$V_{VIDEO} = 1.1115(V_{SCENE})^{0.45} - 0.1115 \quad when\ 0.0228 \le V_{SCENE} \le 1$$
$$V_{VIDEO} = 4V_{SCENE} \quad when\ V_{SCENE} < 0.0228 \ , \tag{5-6}$$

where V_{SCENE} is the voltage before the gamma corrector and V_{VIDEO} is the voltage after. The toe break intensity occurs at $V_{SCENE} = 0.0228$ and the toe slope is 4 (when $L_{SCENE} < 0.0228$). For gamma correction to be accurate, the black reference level must be known. If gamma is inserted after the 7.5% setup, some ambiguity exists in the black level resulting in some intensity distortion. While gamma correction may contain errors, the human eye is very accommodating. The consumer can also adjust the receiver contrast to optimize the image in accordance with his personal preferences. However, this luxury is not available for scientific applications.

Gamma compensation is often performed digitally using a look-up table. Figure 5-10 illustrates the desired output-input transformation recommended by SMPTE 240M. Alternate look-up tables are possible.[24] In fact, some camera manufacturers[21] developed an alternate gamma function *to enhance gray level in a high light intensity region and to obtain high S/N ratio in low light intensity regions simultaneously*. Changing the gamma to minimize the compression of dark areas is also called black stretching. Cameras may have a continuous variable gamma from 0.4 to 1 or may be preselected to 0.45 or 1. The user selects the appropriate value. However, the user rarely knows the CRT gamma. He will simply adjust the camera until he creates an aesthetically pleasing image. This setting will vary from individual to individual.

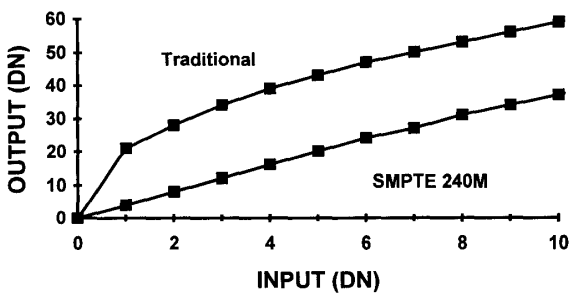


Figure 5-10. Gamma output-input transformation recommended by SMPTE 240M. Expansion of the first 10 levels of an 8-bit ADC system. Data only exist at integer values. Compression of blacks is obvious.

5.3.4. APERTURE CORRECTION

Image sharpness is related to the modulation transfer function (discussed in Chapter 9, *Linear System Theory*, and Chapter 10, *System MTF*). By increasing the MTF at high spatial frequencies, the image usually appears "sharper."

For analysis purposes, a camera is separated into a series of subsystems: optics, detector, electronics, etc. MTF degradation in any one subsystem will soften imagery. The degradation can be partially compensated with electronic boost filters (discussed in Section 10.6.5., *Boost*). Historically, each subsystem was called an "aperture." Therefore boost provides "aperture correction." This should not be confused with the optical system where the optical diameter is the entrance aperture.

Boost filters also amplify noise so that the signal-to-noise ratio remains constant. For systems that are contrast limited, boost appears to improve image quality. However, for noisy images, the advantages of boost are less obvious. As a result, these filters are used only in high contrast situations (typical of consumer applications) and are not used in scientific applications where low signal-to-noise situations are often encountered.

5.4. INDUSTRIAL/SCIENTIFIC CAMERAS

The operation of industrial and scientific cameras is straightforward. They are built around the solid state device capabilities. The array output is amplified and then either processed into a standard video format or left in digital format. Although these may be color cameras, the majority are monochrome.

Progressive scan simply means the noninterlaced or sequential line-by-line scanning of the image (see Section 3.3.5., *Progressive Scan*, page 68). Moving objects can be captured with a strobe light or a fast shutter. Only a progressive scan camera will give the full resolution when the illumination time is short. These cameras can capture imagery asynchronously. That is, the camera timing is initiated from an external sync pulse that is coincident with the strobe flash which may signify that the object is in the center of the field-of-view. Although the image is captured in a progressive scan manner, the camera output is processed into a standard video output so that it can be displayed on commercial monitors. For critical applications, a digital output is available.

TDI offers another method to capture objects moving at a constant velocity. The operation of TDI arrays (see Section 3.3.6, *Time Delay and Integration*, page 70) exactly describes how a TDI camera will operate. The TDI camera output may not be in a standard format and may need a dedicated computer or image processing work station. Reference 25 describes a TDI system used for remote sensing. Here, the aircraft or satellite provides the motion and the objects on the ground are stationary.

Consumer cameras are specified by the minimum detectable signal and dynamic range. In addition to these metrics, scientific and industrial cameras are specified by MTF and the signal-to-noise ratio. Scientific cameras often have a large dynamic range and may have 10-, 12-, 14-, or 16-bit analog-to-digital converters. Key to the operation is that the ADC is linear.

Example 5-2
TDI CAMERA

Bottles move on a conveyor belt at the rate of 1200 ft/min. An image processing system must detect defects without any human intervention. The bottle must be scrapped if a defect is larger than 0.1 inches. The camera sits 48 inches from the conveyor belt. If the detectors are 20 μm square, what focal length lens should be used? What is the camera line rate? If the bottle is 10 inches tall, what should the array size be?

Image processing software accuracy increases as the number of pixels on the target increases. While one pixel on the target may be sufficient for special cases, signal-to-noise ratio considerations and phasing effects typically require that the object cover at least three pixels. Thus, the image of the defect must be at least 60 μm.

The optical magnification is

$$M_{OPTICS} = \frac{R_2}{R_1} = \frac{y_2}{y_1} = \frac{60\ \mu m}{(0.1\ inch)(25{,}400\ \mu m/inch)} = 0.0236\ , \qquad (5\text{-}7)$$

where R_1 is the distance from the lens to the object and R_2 is the lens to array distance. The focal length is

$$fl = R_1 \frac{y_2}{y_1} = \frac{R_2 R_1}{R_2 + R_1} = 1.1\ inches. \qquad (5\text{-}8)$$

The image is moving at (1200 ft/min)(1 min/60 s)(304,800 μm/ft)(0.0236) = 143,865 μm/s. Because the detectors are 20 μm, the line rate is 7193 lines/s. This creates a detector dwell time of 1/7193 = 140 μs. That is, the edge of the defect has moved across the detector element in 140 μs. The pixel clock must operate at 7193 Hz and the integration time is 140 μs.

The image of the bottle is (10 inches)(25400 μm/inch)(0.0236) = 5994 μm. With 20 μm square detectors, the array must contain 300 detectors in the vertical direction. The number of TDI elements depends upon the detector noise level, light level, and detector integration time. The signal-to-noise ratio increases by $\sqrt{N_{TDI}}$.

5.4.1. ANALOG-to-DIGITAL CONVERTERS

An analog-to-digital converter is an essential component of all digital cameras. For consumer applications, an 8-bit converter is used. Sometimes seemingly slight deviations from perfection in an ADC can modify image fidelity[26] when examined by a trained photointerpreter or by an automatic target detection algorithm. While a nonlinear ADC may not be visually obvious with a single readout, it may become obvious if the array has multiple outputs with each output having its own ADC. Here, any object which spans the two areas with different ADCs will have a discontinuity in intensity.

ADC linearity is specified by differential nonlinearity (DNL) and integral nonlinearity (INL) (Figure 5-11). The differential nonlinearity is a test of missing codes or misplaced codes. Two adjacent codes should be exactly one LSB apart. Any positive or negative deviation from this is the DNL. If the DNL exceeds one LSB, there is a missing code. Many nonlinearity specifications

simply state "no missing codes." This means that the DNL is less than one LSB. However, 1.01 LSB DNL (fail specification) and 0.99 LSB DNL (pass specification) probably look the same visually. Poor DNL can affect the cosmetic quality of images.

Integral nonlinearity is the deviation of the transfer function from an ideal straight line. Most ADCs are specified with endpoint INL. That is, INL is specified in terms of the deviation from a straight line between the end points. Good INL is important for radiometric applications.

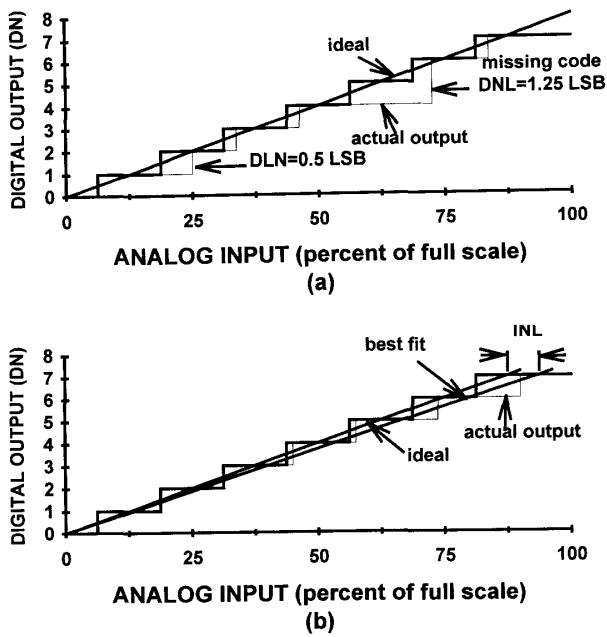


Figure 5-11. ADC nonlinearity for 3-bit ADC. (a) Differential nonlinearity and (b) integral nonlinearity. Missing codes can be caused by the accumulation of DNLs. Accumulation can also cause INL.

ADCs exist in most cameras. With commercially available cameras, the manufacturer has selected the ADC. To ensure that the ADC is linear, the user must experimentally characterize the performance in a manner consistent with his application.[27] This is rather difficult if the camera output is an analog signal. The analog output represents a smoothed output of the ADC. The smoothing can hide ADC nonlinearities.

The same problem may exist within a frame capture board. The user must test the board for any nonlinearities. Because the frame capture board usually resides within a computer, testing is relatively easy.

5.4.2. INTENSIFIED SOLID STATE CAMERAS

Although low light level televisions can be used for many applications, they tend to be used for military and scientific applications. The term *image intensifier* refers to a series of special imaging tubes[28] that have been designed to amplify the number of photons internally. These tubes increase the number of incident photons by several orders of magnitude. The more popular second generation (GenII) and third generation (GenIII) intensifier tubes use a microchannel plate (MCP) for intensification. Third generation is similar in design to second generation with the primary difference being higher quantum efficiency.

When photons hit the photocathode, electrons are liberated and thereby create an electron image. The electrons are focused onto the MCP and about 80% enter the MCP. The MCP consists of thousands of parallel channels (glass hollow tubes) that are about 10 μm in diameter. Electrons entering the channel collide with the coated walls and produce secondary electrons. These electrons are accelerated through the channel by a high potential gradient. Repeated collisions provide an electron gain up to several thousand. Emerging electrons strike the phosphor screen and a visible intensified image is created. The intensifier is an analog device and an observer can view the resultant image on the phosphor screen.

The intensifier output can be optically linked[29] to a solid state array with either a relay lens or a fiberoptic bundle (Figure 5-12). When coupled to a CCD, it is called an ICCD. The intensified camera spectral response is determined solely by the intensifier photocathode responsivity. Although Figure 5-13 illustrates GenII and GenIII spectral responsivities, a variety of photocathodes are available.[30] The window material controls the short wavelength cutoff whereas the photocathode determines the long wavelength cutoff.

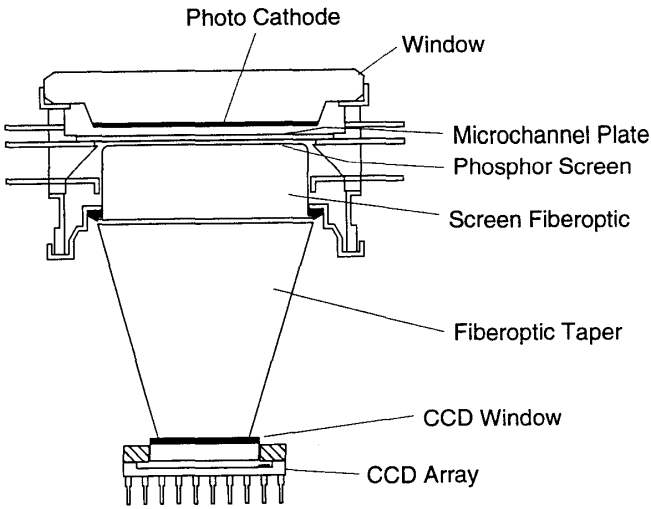


Figure 5-12. CCDs can be coupled to an image intensifier using either fiberoptic bundle or relay lens. The tapered fiberoptic bundle (shown here) tends to be the largest component of the ICCD.

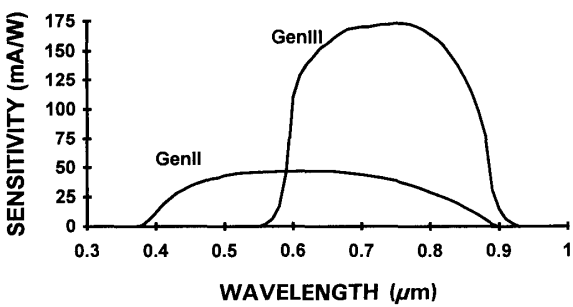


Figure 5-13. Nominal spectral response of militarized (fielded) GenII and GenIII image intensifier tubes. GenIII provides a better match to the night sky illumination (see Figure 2-13, page 39). The responsivity depends upon the window and photocathode material. GenII tubes use a multialkali photocathode whereas GenIII tubes use GaAs photocathodes.

Although many phosphors are available for the intensifier output, a phosphor should be selected whose peak emission is at the peak of the solid state array spectral responsivity. A color camera is not a good choice because it has lower quantum efficiency due to the color-separating spectral filters. Furthermore, single chip color cameras have unequal sampling densities and this can lead to a loss in resolution.

The intensified camera can provide high quality, real-time video when the illumination is greater than 10^{-3} lux (corresponding to starlight or brighter). For comparison, standard cameras (non-intensified) are sensitive down to about 1 lux. Thus, intensified cameras offer high quality imagery over an additional three orders of magnitude in illumination. A three-order magnitude increase in sensitivity implies that the MCP gain should be about a thousand. However, it must be an order of magnitude larger to overcome the transmission loss in the fiber optic bundle or relay lens. For very low illumination levels, it may be necessary to increase the integration time and then real-time video is no longer possible.

By simply changing the voltage on the image intensifier, the gain can be varied. Thus, operation over a wide range of luminance levels ranging from starlight to daylight is possible. Gating is probably the biggest advantage of the intensified camera. Electronic shutter action is produced by pulsing the MCP voltage. Gating allows the detection of low level signals in the presence of interfering light sources of much greater energy by temporal discrimination. Gate times as short as 5 ns FWHM (full width at half maximum) are possible. This allows the detection of laser diagnostic pulses against very intense continuous sources. Because full frame transfer or frame transfer CCDs are used for ICCD applications, the inherent gating capability of the intensifier is used as a shutter during transfer time. This eliminates image smear.

Because the detector size is typically smaller than the intensifier's resolution, the intensifier's image must be minified. Fiber-optic bundles reduce camera size and weight (lenses are relatively heavy). But fiberoptic coupling is complex because it requires critical alignment of the fibers to the individual detectors. The fiber area and pitch, ideally, should match the detector size and pitch for maximum sensitivity and resolution. However, to alleviate critical alignment issues, the fiber diameter is typically one-half the pixel pitch. This reduces sensitivity somewhat.

The lens-coupled system avoids the alignment problems associated with fibers and does not introduce moiré patterns. It offers more flexibility because an image intensifier with an output lens can be connected to an existing solid

state camera. The lens coupling represents a cost-effective way to add gating capability to an existing camera. However, a lens-coupled system provide a much lower throughput (2× to 10× lower) than a fiber coupled system.

There are four spatial samplers in the intensified camera: (1) sampling by the discrete channels in the MCP, (2) the fiberoptic output window within the intensifier, (3) the input to the fiberoptic minifier, and (4) the detector array. At these interfaces, the image is re-sampled. Depending upon the location (alignment) of the fiberoptic minifier, moiré patterns can develop and the MTF can degrade. In addition, manufacturing defects in the fiberoptic can appear[31] as a "chicken wire" effect.

The intensifier tube typically limits the intensified camera resolution. Increasing the array resolution may have little effect on the system resolution. The array is typically rectangular whereas the intensifier is round and the array cannot sense the entire amplified scene. Most commercially available intensified cameras are based on the standard military 18-mm image intensifier. Larger tubes (25 mm and 40 mm) are available but are more expensive and increase coupling complexity.

Intensifiers are electron tubes. Tubes can bloom so that scenes with bright lights are difficult to image. For example, it is difficult to image shadows in a parking lot when bright street lights are in the field-of-view. Intensifiers are subject to the same precautions that should be observed when using most photosensitive coated pickup tubes. Prolonged exposure to point sources of bright light may result in permanent image retention.

The intensifier specifications alone do not accurately portray how the system will work. Only a detailed analysis will predict performance. Resolution of the array will not describe system resolution if the image intensifier provides the limiting resolution. While the array photoresponse may be quite uniform, the overall nonuniformity is limited by the image intensifier. Standard military image intensifier photocathodes have nonuniformities approaching ±30%. Only carefully selected intensifiers with low nonuniformity are used as part of the intensified system.

5.5. REFERENCES

1. The timing requirements can be found in many books. See, for example, D. H. Pritchard, "Standards and Recommended Practices," in *Television Engineering Handbook*, K. B. Benson, ed., pp. 21.40-21.57, McGraw-Hill, New York, NY (1986).
2. D. H. Pritchard, "Standards and Recommended Practices," in *Television Engineering Handbook*, K. B. Benson, ed., pp. 21.73-21.74, McGraw-Hill, New York, NY (1986).

3. B. D. Loughlin, "Monochrome and Color Visual Information Transmission," in *Television Engineering Handbook*, K. B. Benson, ed., pp. 4.4 - 4.5, McGraw-Hill, New York, NY (1986).
4. D. H. Pritchard, "Standards and Recommended Practices," in *Television Engineering Handbook*, K. B. Benson, ed., pp. 21.10 - 21.11, McGraw-Hill, New York, NY (1986).
5. S. C. Hsu, "The Kell Factor: Past and Present," *SMPTE Journal*, Vol. 95, pp. 206-214 (1986).
6. "EIA Standard EIA-343A, Electrical Performance Standards for High Resolution Monochrome Closed Circuit Television Camera," Electronic Industries Association, 2001 Eye Street, NW Washington, D.C. 20006 (1969).
7. See, for example, "Solving the Component Puzzle," Tektronix Book # 25W-7009-1, dated 9/90, Tektronix Inc., Beaverton, OR.
8. J. Hamalainen, "Video Recording Goes Digital," *IEEE Spectrum*, Vol. 32(4), pp. 76-79 (1995).
9. B. D. Loughlin, "Monochrome and Color Visual Information Transmission," in *Television Engineering Handbook*, K. B. Benson, ed., pp. 4.24-4.42, McGraw-Hill, New York, NY (1986).
10. L. J. Thorpe, "HDTV and Film - Issues of Video Signal Dynamic Range," *SMPTE Journal*, Vol. 100(10), pp. 780-795 (1991).
11. L. J. Thorpe, "HDTV and Film - Digitization and Extended Dynamic Range," *SMPTE Journal*, Vol. 102(6), pp. 486-497 (1993).
12. B. L. Lechner, "HDTV Status and Prospects," in *Seminar Lecture Notes*, 1995 SID International Symposium, pp. F4/4-F4/74, Society for Information Display, Santa Ana, CA (1995).
13. J. Blankevoort, H. Blom, P. Brouwer, P. Centen, B. vd Herik, R. Koppe, A. Moelands, J. v Rooy, F. Stok, and A. Theuwissen, "A High-Performance, Full-Bandwidth HDTV Camera Applying the First 2.2 Million Pixel Frame Transfer CCD Sensor," *SMPTE Journal*, Vol. 103(5), pp. 319-329 (1994).
14. B. Bhatt, D. Birks, and D. Hermreck, "Digital Television: Making It Work," *IEEE Spectrum*, pp. 19-28 (October 1997).
15. K. B. Benson, ed., *Television Engineering Handbook*, Chapter 2 and pp. 22.29 - 22.39, McGraw-Hill, New York, NY (1986).
16. H. R. Kang, *Color Technology for Electronic Imaging Devices*, pp. 3-101, SPIE Press, Bellingham, WA (1997).
17. K. A. Parulski, L. J. D'Luna, and R. H. Hubbard, "A Digital Color CCD Imaging Systems using Custom VLSI Circuits," *IEEE Transactions on Consumer Electronics*, Vol. CE-35(8), pp. 382-388 (1989).
18. S. Nishikawa, H. Toyoda, V. Miyakawa, R. Asada, Y. Kitamura, M. Watanabe, T. Kiguchi, and M. Taniguchi, "Broadcast-Quality TV Camera with Digital Signal Processor," *SMPTE Journal*, Vol. 99(9), pp. 727-733 (1990).
19. Y. Takemura, K. Ooi, M. Kimura, K. Sanda, A. Kuboto, and M. Amano, "New Field Integration Frequency Interleaving Color Television Pickup System for Single Chip CCD Camera," *IEEE Transactions on Electron Devices*, Vol. ED-32(8), pp. 1402-1406 (1985).
20. H. Sugiura, K. Asakawa, and J. Fujino, "False Color Signal Reduction Method for Single Chip Color Video Cameras," *IEEE Transactions on Consumer Electronics*, Vol. 40(2), pp. 100-106 (1994).
21. J.-C. Wang, D.-S. Su, D.-J. Hwung, and J.-C. Lee, "A Single Chip CCD Signal Processor for Digital Still Cameras," *IEEE Transactions on Electron Devices*, Vol. 40(3), pp. 476-483 (1994).
22. D. H. Pritchard, "Standards and Recommended Practices," in *Television Engineering Handbook*, K. B. Benson, ed., p. 21.46, McGraw-Hill, New York, NY (1986).
23. SMPTE 240M-1988 appears in *SMPTE Journal*, Vol. 98, pp. 723-725 September 1989 and can be obtained from SMPTE, 595 West Hartsdale Ave. White Plains, NY 10607.

24. A. N. Thiele, "An Improved law of Contrast Gradient for High Definition Television," *SMPTE Journal*, Vol. 103(1), pp. 18-25 (1994).
25. T. S. Lomheim and L. S. Kalman, "Analytical Modeling and Digital Simulation of Scanning Charge-Coupled Device Imaging Systems," in *Electro-Optical Displays*, M. A. Karim, ed., pp. 551-560, Marcel Dekker, New York (1992).
26. C. Sabolis "Seeing is Believing," *Photonics Spectra*, pp. 119-124, October 1993.
27. D. H. Sheingold, *Analog-Digital Conversion Handbook*, pp. 317-337, Prentice-Hall, Englewood Cliffs, NJ (1986).
28. See, for example, I. Csorba, *Image Tubes*, Howard Sams, Indianapolis, Indiana (1985).
29. Y. Talmi, "Intensified Array Detectors," in *Charge-Transfer Devices in Spectropscopy*, J. V. Sweedler, K. L. Ratzlaff, and M. B. Denton, eds., Chapter 5, VCH Publishers, New York (1994).
30. I. Csorba, *Image Tubes*, p. 209, Howard Sams, Indianapolis, Indiana (1985).
31. G. M. Williams, "A High-Performance LLLTV CCD Camera for Nighttime Pilotage," in *Electron Tubes and Image Intensifiers*, C. B. Johnson and B. N. Laprade, eds., SPIE Proceedings Vol. 1655, pp. 14-32 (1992).

6

CAMERA PERFORMANCE

Camera performance metrics are conceptually the same as array metrics (see Chapter 4, *Array Performance*, page 102). The camera is limited by the array noise floor, dark current, and charge well capacity. The camera's fixed pattern noise (FPN) and photoresponse nonuniformity (PRNU) may be better than the array pattern noise when correction algorithms are present. Frame averaging and binning can reduce the random noise floor and thereby appear to increase the camera's dynamic range.

The signal is modified by the spatial response of the optics and detector whereas the noise originates in the detector and electronics. Noise characteristics may be different from signal characteristics. System design requires optimizing the signal-to-noise ratio.

This chapter covers maximum signal, minimum signal, dynamic range, and the signal-to-noise ratio. It assumes that the image size is large compared to the detector's area. Point source detection calculations are more difficult and can be found in References 1 and 2. Camera metrics, depend in part, on the application (see Section 5.2.2., *Broadcast/Non-broadcast Design*, page 155).

Historically, performance metrics were created for broadcast systems. These metrics may be applied to scientific, machine vision, and military systems. Some of these metrics are not appropriate for specific applications and therefore may be considered as superfluous information.

Image intensifier and array characteristics both affect the intensified camera signal-to-noise ratio. ICCDs have a dynamic range that varies with the image intensifier gain. This permits fabrication of general purpose ICCDs that operate in lighting conditions that range from starlight to full sunlight.

The symbols used in this book are summarized in the *Symbol List* (page xviii) which appears after the *Table of Contents*.

6.1. NOMENCLATURE

Figure 6-1 illustrates a functional block diagram of an analog camera. It is functionally identical to Figure 4-1 (page 104). Although the off-chip electronics is shown to have only one amplifier in Figure 4-1, it may consist of several amplifiers. One of the amplifiers may have a variable gain that allows the user to increase the signal under low light level conditions. This gain is turned off (i.e., $G_2 = 1$) when the maximum illumination is measured. The value of G_1 is superfluous because its only purpose is to provide a standard video voltage.

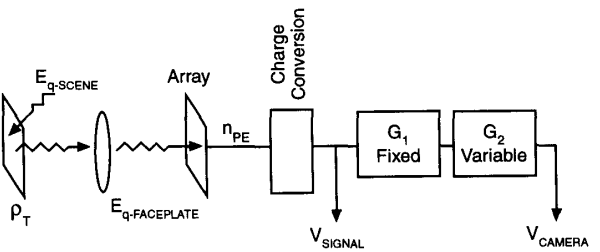


Figure 6-1. Camera function block diagram. The gain G_2 may be automatic (AGC) or manually controlled.

The input can be measured at two different places. With the first measurement, the lens focuses reflected scene illumination, $\rho_T E_{v\text{-SCENE}}$, onto the array where ρ_T is the target reflectivity. Tables 2-5 and 2-6 (page 28) provide typical values of $E_{v\text{-SCENE}}$. Here, the camera's output depends upon the lens f-number and transmittance (or equivalently the T-number). In the second measurement, the faceplate illumination is used. For radiometric inputs, the scene irradiance or faceplate irradiance is used. In Equation 4-2 (page 112), the source was represented by M_e and it is identical to $\rho_T E_{e\text{-SCENE}}$. Because the manufacturer has no control over the user's selection of lenses, the faceplate illumination is often specified.

Array parameters are based upon the photon flux irradiating the detector, $E_{q\text{-FACEPLATE}}$ and the measured output V_{SIGNAL}. The number of photons reaching the detector is

$$n_{DETECTOR} = E_{q\text{-}FACEPLATE} A_D t_{int} , \qquad (6\text{-}1)$$

which is identical to Equation 2-12 (page 34).

The array quantum efficiency is

$$R_q(\lambda) = \frac{n_{PE}(\lambda)}{n_{DETECTOR}(\lambda)} . \tag{6-2}$$

The array signal-to-noise is

$$SNR_{ARRAY} = \frac{n_{PE}}{\langle n_{SYS} \rangle} \tag{6-3}$$

and the array dynamic range is

$$DR_{ARRAY} = \frac{V_{SIGNAL}}{V_{NOISE}} = \frac{N_{WELL} - n_{DARK}}{\langle n_{SYS} \rangle} . \tag{6-4}$$

The array outputs, n_{PE} and $<n_{SYS}>$, are derived theoretically or are calculated from the measured values V_{SIGNAL} or V_{CAMERA}. Often only $<n_{FLOOR}>$ is used for $<n_{SYS}>$ to calculate DR_{ARRAY}. This provides the highest possible dynamic range. If the camera does not introduce any additional noise, modify the noise bandwidth, or minimize the maximum output, the camera SNR and DR_{CAMERA} will be identical to the array values.

For Lambertian targets,

$$E_{v\text{-}FACEPLATE} = \frac{\rho_T \, E_{v\text{-}SCENE}}{4 F^2 (1 + M_{OPTICS})^2} T_{OPTICS} \, T_{ATM} \quad lux. \tag{6-5}$$

On a clear day ($T_{ATM} \approx 1$), a white target ($\rho_T \approx 1$) which is far way ($M_{OPTICS} \approx 0$) provides

$$E_{v\text{-}FACEPLATE} \approx \frac{\rho_T E_{v\text{-}SCENE}}{4 F^2} T_{OPTICS} \quad lux. \tag{6-6}$$

For back-illuminated targets (such as those used for calibration), the flux emanating from the target is $\rho_T E_{v\text{-}SCENE}$. The measurement of faceplate illumination is usually performed with a calibrated lens. The value $\rho_T E_{v\text{-}SCENE}$ is measured with a calibrated radiometer or photometer, V_{CAMERA} is measured, and then by using Equation 6-6 or 6-7 $E_{v\text{-}FACEPLATE}$ is calculated.

With four possible inputs (scene irradiance, scene illuminance, faceplate irradiance, and faceplate illuminance) and three outputs (number of electrons, V_{SIGNAL}, and V_{CAMERA}) there potentially exists 12 input-output performance parameters. Each metric is appropriate for a specific application. This chapter will consider only the most widely used metrics. The derivation of the others is straightforward. For image intensified systems, the scene or faceplate values are used as the input to the image intensifier.

The maximum and minimum signals depend on the spectral output of the source and the spectral response of the detector. The source color temperature is not always listed but is a critical parameter for comparing systems. Although the CIE illuminant A is used most often, the user should not assume that this was the source used by the camera manufacturer. Manipulation of the source (CIE illuminant A, B, C, D_{65}, or another unspecified source) and optical filter characteristics can change the maximum and minimum signal values.

Based on signal detection theory, the minimum illumination would imply that the signal-to-noise ratio is one. However, the definition of minimum illumination is manufacturer dependent. Its value may be (a) when the video signal is, say, 30 IRE units, (b) when the SNR is one, or (c) when an observer just perceives a test pattern. Because of its incredible temporal and spatial integration capability, the human visual system can perceive SNRs as low as 0.05. This is further discussed in Section 12.1.1., *Perceived Signal-to-Noise Ratio*. Therefore, comparing cameras based on "minimum" illumination should be approached with care.

6.2. CAMERA METRICS

Both the signal and noise are usually measured. By using a measured value, all noise sources are included: shot noise, noise floor, reset noise, pattern noise, quantization noise, amplifier white noise, and amplifier 1/f noise. The minimum signal is then calculated using the variable amplifier gain, G_2, and signal-to-noise ratio.

6.2.1. CAMERA SNR

For many applications, the object is far away so that M_{OPTICS} approaches zero. When measured at the camera output, the SNR is

$$SNR_{SCENE} = \frac{\frac{G_2 G_1 G q}{C} \frac{A_D}{4F^2} \int_{\lambda_1}^{\lambda_2} R_q(\lambda)\, \rho_T(\lambda)\, M_{q\text{-}SCENE}(\lambda)\, t_{INT}\, T_{OPTICS}(\lambda)\, d\lambda}{\sqrt{\int_0^{\infty} S(f_{elec})\, |H_{SYS}(f_{elec})|^2\, df_{elec}}}, \quad (6\text{-}7)$$

where the subscript *scene* is used to indicate the input is the reflected scene flux. Noise is inserted by the detector and by the amplifiers (see Figure 4-11, page 123). The electronics MTF, $H_{SYS}(f_{elec})$, modifies the noise spectra. The noise power spectral density, $S(f_{elec})$, includes all the noise sources: shot noise, noise floor, reset noise, pattern noise, quantization noise, amplifier white noise, and amplifier 1/f noise. The subscript *elec* emphasizes that electrical frequency applies to the amplifiers and subsequent electronic filters. The MTFs are described in Chapter 10, *System MTF*.

Equation 6-8 is the generalized SNR equation. When equated to electrons,

$$SNR_{SCENE} = \frac{\frac{A_D}{4F^2} \int_{\lambda_1}^{\lambda_2} R_q(\lambda)\, \rho_T(\lambda)\, M_{q\text{-}SCENE}(\lambda)\, t_{INT}\, T_{OPTICS}(\lambda)\, d\lambda}{\langle n_{SYS} \rangle}. \quad (6\text{-}8)$$

For a small spectral band

$$SNR_{SCENE} \approx \frac{\frac{A_D}{4F^2} R_q(\lambda_o)\, \rho_T(\lambda_o)\, M_{q\text{-}SCENE}(\lambda_o)\, t_{INT}\, T_{OPTICS}(\lambda_o)\, \Delta\lambda}{\langle n_{SYS} \rangle}. \quad (6\text{-}9)$$

The SNR is proportional to the detector area and integration time. Figure 6-2 provides the relative SNR as a function of these parameters. Binning is identical to increasing the detector area. The output of four binned detectors is the same as one detector four times larger. For fixed area, increasing the integration time increases the SNR. For fixed integration time, binning increases the SNR. Shorter integration times permit higher frame rates (See Section 4.5., *Frame Rates*, page 142). Shorter integration times also decrease the SNR.

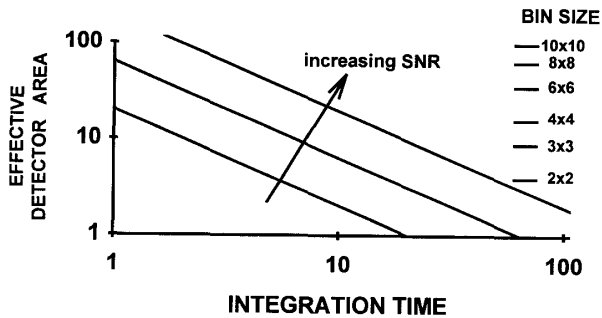


Figure 6-2. Relative SNR as a function of bin size (detector area) and integration time. Only shot noise is considered.

6.2.2. CAMERA DYNAMIC RANGE

Dynamic range is the maximum signal divided by the noise. It depends on integration time, binning, and frame integration. In addition, the spectral content of the source and the array spectral responsivity affect the camera output voltage. Thus, the camera dynamic range can be quite variable depending on test conditions.

Although the theoretical maximum signal just fills the charge wells, the manufacturer may limit the maximum output voltage to a lower stored-charge value. Because the dynamic range is often expressed in decibels,

$$DR_{CAMERA\text{-}dB} = 20 \log\left(\frac{V_{MAX}}{V_{Noise}}\right) \tag{6-10}$$

and the noise voltage is

$$V_{NOISE} = V_{MAX} \cdot 10^{\left(-\frac{DR_{CAMERA-dB}}{20}\right)} . \qquad (6\text{-}11)$$

Many camera manufacturers list the dynamic range as the signal-to-noise ratio. This value should not be confused with the actual SNR (see Section 4.3., *Array Signal-to-noise Ratio*, page 139). With most cameras, the noise level is approximately equivalent to the array read noise. The electronics may increase the camera noise, and image processing techniques may reduce it somewhat. Usually the dynamic range is calculated from the measured signal and noise. Because the read noise is amplifier white noise, it is appropriate to include the bandwidth as part of the measurements. For standard video compatible cameras, the bandwidth is equal to the video format bandwidth (see Section 5.2.1., *Video Timing*, page 151).

Example 6-1
DIGITAL DYNAMIC RANGE

A device with a well capacity of 50,000 electrons has a noise floor of 60 electrons rms. What size analog-to-digital converter should be used?

Adding noise powers, the total camera noise is related to the array noise and ADC noise by

$$\langle n_{SYS} \rangle = \sqrt{\langle n_{FLOOR}^2 \rangle + \langle n_{ADC}^2 \rangle} . \qquad (6\text{-}12)$$

If an ADC is selected so that an LSB is equal to the array noise level, the ADC increases the camera noise by 4%. The array dynamic range is 50,000/60 = 833 but the smallest available ADC has 10 bits or 1024 levels. This ADC devotes 1.23 bits to array noise. But if the LSB was one-quarter of the array noise level, the camera noise would increase by only 1%. The ADC dynamic range must be greater than 3332:1. The smallest ADC that covers this range contains 12 bits.

6.2.3. MAXIMUM SIGNAL

If specifying scene illumination, a f-number must be selected. Historically, film-based cameras are compared when the lens f-number is 5.6. Figure 6-3 illustrates the relative change in the maximum signal as the f-number is varied. For consumer video and industrial cameras, the integration is selected to match the video standard (t_{INT} = 1/30 for EIA 170). Automatic irises and variable integration times can increase the maximum level.

The photon level that saturates the charge wells is the camera's maximum input signal and $V_{WELL} = G_1 G q N_{WELL}/C$ where $G_2 = 1$. However, there is no requirement that the camera manufacturer set the maximum output voltage equal to V_{WELL}. If the array has an antibloom drain, the maximum output may be limited to the white clip level (see Figure 3-35, page 87).

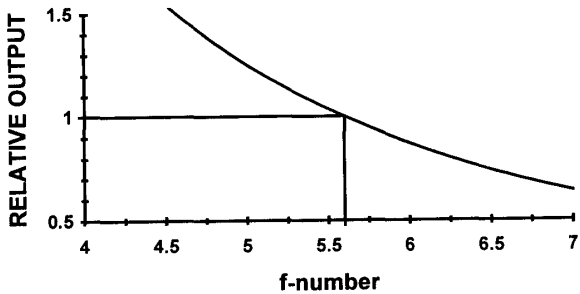


Figure 6-3. Relative maximum output normalized to f/5.6.

6.2.4. MINIMUM SIGNAL

The minimum input signal suggests something about the lowest detectable signal. When using scene illumination, the minimum signal is often specified with different optics than the optics specified for the maximum signal (Figure 6-4). Faster optics is a reasonable choice because these lenses are used in low illumination conditions.

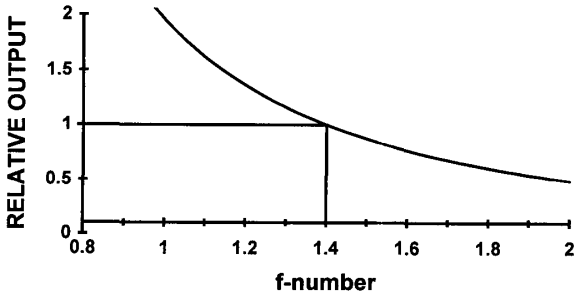


Figure 6-4. Relative minimum output normalized to f/1.4. A camera can appear to have very high sensitivity if a very low f-number lens is used.

The minimum signal provided with gain "on" (G_2 greater than one) is usually calculated due to the difficulty of performing accurate low level radiometric and photometric measurements. These values may be provided at 30%, 50%, or 80% video levels. That is, the illumination given produces an output video that gives 30, 50, or 80 IRE units, respectively (see Section 5.2.4., *IRE Units*, page 157). Although a higher gain amplifier could provide 100% video, the user can optimize the image by adjusting the gain and level of the display. That is, the display's internal amplifiers can be used for additional gain. In this context, the camera *system* consists of the camera and display.

If G_2 is expressed in decibels,

$$G_2 = 10^{\left(\frac{G_{2\text{-}dB}}{20}\right)} . \qquad (6\text{-}13)$$

The scene illumination that creates a 30% video signal is

$$M_{v\text{-}SCENE}(30\%video) = \frac{0.3 \cdot M_{v\text{-}SCENE}(MAX)}{G_2} , \qquad (6\text{-}14)$$

where $M_{v\text{-}SCENE}(MAX)$ is the scene illumination that produced the maximum output. For 50% and 80% video, the factor becomes 0.5 and 0.8, respectively.

The input signal that produces a signal-to-noise ratio of one is

$$M_{v\text{-}SCENE}(SNR = 1) = M_{v\text{-}SCENE}(Max) \cdot 10^{\left(-\frac{DR_{CAMERA\text{-}dB}}{20}\right)} . \qquad (6\text{-}15)$$

For faceplate illumination, $M_{v\text{-}SCENE}$ is replaced with $E_{v\text{-}FACEPLATE}$. Actual performance may deviate from the calculated values. The gain used in the equations is typically a nominal value. Usually more gain is available than that specified in the data specifications.

The examples in Chapter 4 (pages 116-118, 122, and 131) should be reviewed prior to reading the following examples.

Example 6-2
MINIMUM SIGNAL ESTIMATION

A camera provides maximum output when the faceplate illumination is 0.65 lux. The dynamic range is 55 dB. What is the equivalent noise signal?

Using Equation 6-15, the noise signal is 1.16×10^{-3} lux. A signal-to-noise of one provides an extremely noisy image. While a human observer can detect some objects when the SNR is one, a machine vision system cannot.

Example 6-3
VIDEO SIGNAL ESTIMATION

The camera described in Example 6-2 provides a gain of 20 dB. What is the SNR at 30% and 80% video?

Using Equations 6-14 and 6-15, at 30% video the faceplate illumination is 0.0195 lux and at 80% video, the faceplate illumination is 0.052 lux. The SNR is 17 and 45, respectively. With a SNR of 17, the image will appear slightly noisy but probably would be considered acceptable to most observers. However, if no additional gain is used, the image will appear very dark on the display. Here, the display gain must be increased to make the image viewable. Depending upon the task, a machine vision system may need a higher SNR. While computer programs are usually insensitive to absolute signal values, the frame grabber's ADC may not provide enough levels for adequate image

processing. For example, if the ADC is an 8-bit system, 30% video only provides digital numbers ranging from 0 to 77.

Example 6-4
NOISE EQUIVALENT INPUT

In Example 6-2, the minimum signal was specified in photometric units. In the more general radiometric case, the input that provides SNR = 1 is the noise equivalent input (NEI). That is, $\rho_T M_{q\text{-SCENE}}$ is found from Equation 6-9 when $SNR_{q\text{-SCENE}}$ is one. Over a small spectral band,

$$NEI \approx \frac{4F^2}{A_D\, t_{INT}\, \eta\, \Delta\lambda\, T_{OPTICS}} \langle n_{SYS} \rangle \quad \frac{photons}{s\text{-}m^2\text{-}\mu m}. \tag{6-16}$$

When the SNR is unity, the number of photons is small and photon noise is negligible (see Figure 4-16b, page 136). If the detector is cooled, dark current is minimized. With PRNU and FPN correction, the minimum NEI is

$$NEI \approx \frac{4F^2}{A_D\, t_{INT}\, \eta\, \Delta\lambda\, T_{OPTICS}} \langle n_{FLOOR} \rangle \quad \frac{photons}{s\text{-}m^2\text{-}\mu m}. \tag{6-17}$$

If TDI is used, then N_{TDI} is multiplied in the denominator. With binning or super pixeling, the number of pixels combined is multiplied in the denominator.

Example 6-5
NOISE EQUIVALENT REFLECTANCE

The NEI assumes that the target is against a black background. For real scenes, both the target and background create signals. The target is only visible when a sufficient reflectance difference exists. For detecting objects at long ranges, the atmospheric transmittance loss must be included. The signal difference between the target and its background is

$$\Delta n_{pe} = \int_{\lambda_1}^{\lambda_2} R_q(\lambda) \frac{(\rho_T - \rho_B) M_q(\lambda)\, t_{INT}\, A_D}{4F^2} T_{OPTICS}(\lambda)\, T_{ATM}(\lambda) d\lambda \ . \quad (6\text{-}18)$$

Both the target and background are illuminated by the same source. The signals from both depend on the spectral reflectances that are taken as constants over the spectral region of interest. Using the same simplifying assumptions as before, solving for $\rho_T - \rho_B$ and calling this value the noise equivalent reflectance difference (NE$\Delta\rho$) provides

$$NE\ \Delta\rho \approx \frac{4F^2}{A_d\, t_{INT}\, \eta\, M_q(\lambda_o)\, \Delta\lambda\, T_{OPTICS}\, T_{ATM}} \langle n_{SYS} \rangle \ , \quad (6\text{-}19)$$

where T_{ATM} is the total atmospheric transmittance and it depends on range. The value T_{ATM} is usually expressed as $T_{ATM} = \exp(-\sigma_{ATM} R)$. Then

$$NE\ \Delta\rho \approx \frac{4F^2 e^{\sigma_{ATM} R}}{A_D\, t_{INT}\, \eta\, M_q(\lambda_o)\, \Delta\lambda\, \tau_{OPTICS}} \langle n_{SYS} \rangle = \frac{NEI}{M_q(\lambda_o)} e^{\sigma_{ATM} R} . \quad (6\text{-}20)$$

Thus, NE$\Delta\rho$ depends on the source illumination, system noise, and the atmospheric conditions. As the range increases, the target reflectance must increase to keep the SNR at one. The maximum range occurs when the NE$\Delta\rho$ is one. At this maximum range,

$$NEI = M_q(\lambda_o)\, e^{-\sigma_{ATM} R_{MAX}} \ . \quad (6\text{-}21)$$

The variable $M_q(\lambda)[\rho_T(\lambda) - \rho_B(\lambda)]$ is a complicated function that depends on the cloud cover, atmospheric conditions, sun elevation angle, surface properties of the target, and the angle of the target with respect to the line-of-sight and sun angle.

These back-of-the-envelope approximations are still not complete. They assume that the total received energy is from the target only. In fact, over long path lengths, path radiance contributes to the energy detected. Path radiance, specified by visibility, reduces contrast. The additional light reflected by the atmospheric particulates introduce photon noise that is independent of the signal from the target. It is quantified by the sky-to-background ratio (discussed in Section 12.6.1, *Contrast Transmittance*).

Example 6-6
MINIMUM PHOTOMETRIC SIGNAL ESTIMATION

A CCD array has a noise floor of 150 electrons rms. If placed into a camera, what is the estimated minimum faceplate detectable signal?

The minimum signal occurs when the signal-to-noise ratio is one. That is, $n_{pe} = \langle n_{FLOOR} \rangle$. Several simplifying assumptions are required to obtain a back-of-the-envelope expression for the minimum signal. If the CIE illuminant D_{6500} is used, the spectral photon sterance is approximately constant over the spectral region, $\Delta\lambda$. If the quantum efficiency is constant,

$$E_{q\text{-}FACEPLATE}(\lambda) \approx \frac{\langle n_{FLOOR} \rangle}{\eta A_D \Delta\lambda t_{INT}} \quad \frac{photons}{s\text{-}m^2\text{-}\mu m} . \tag{6-22}$$

Because the energy of one photon is hc/λ, the faceplate spectral power is

$$E_{p\text{-}FACEPLATE}(\lambda) \approx \frac{\langle n_{FLOOR} \rangle}{\eta \Delta\lambda t_{INT}} \frac{h c}{\lambda} \quad \frac{W}{\mu m} . \tag{6-23}$$

The faceplate illumination is

$$E_{v\text{-}FACEPLATE} = \frac{683}{A_D} \int_{0.4\,\mu m}^{0.7\,\mu m} V(\lambda) E_{p\text{-}FACEPLATE}(\lambda) d\lambda \quad lux \tag{6-24}$$

or

$$E_{V\text{-}FACEPLATE} \approx \frac{\langle n_{FLOOR} \rangle h c}{\eta \Delta\lambda t_{INT}} \frac{683}{A_D} \int_{0.4}^{0.7} \frac{V(\lambda)}{\lambda} d\lambda \quad lux. \tag{6-25}$$

From Figure 4-5 (page 108) $\eta \approx 0.3$. For EIA 170 compatibility, $t_{INT} = 1/30$ s, and the spectral bandwidth is $\Delta\lambda = (0.7 - 0.4)\,\mu m = 0.3\,\mu m$. The photometric integral is approximately 0.2. The constant, hc, is approximately 2×10^{-19} J-μm. Then

$$E_{v\text{-}FACEPLATE} = \frac{1.36 \times 10^{-12}}{A_D} \quad lux. \tag{6-26}$$

Assume a 50% fill factor interline transfer array is available. From Table 3-3 (page 99), a ½-inch format detector will have an active size of approximately 4.17 μm × 10 μm ($A_D = 41.7 \times 10^{-12}$ m^2). Then the estimated minimum signal is 3.3×10^{-2} lux.

Consider a target that is pure white, ($\rho_T = 1$) on a black background ($\rho_B = 0$). At low light levels, a f/1.4 lens is often used. Let $T_{OPTICS} = 0.9$ and $T_{ATM} = 1.0$. Using Equation 6-6 (page 184) provides $E_{v\text{-}SCENE} = 0.29$ lux. This target will have a signal-to-noise ratio of one and therefore be very difficult to see. Real targets and backgrounds have reflectivities between 0 and 1. When including the signal-to-noise ratio, $E_{v\text{-}SCENE} = 0.29\,(SNR)/(\rho_T - \rho_B)$. Let the required SNR be 20, $\rho_T = 0.6$, and $\rho_B = 0.4$, then $E_{v\text{-}SCENE} = 29$ lux. From Table 2-5 (page 28), this target can be seen during twilight.

While the assumptions are not unreasonable, the statement that the target can be seen during twilight is an extrapolation. The spectral output of the D_{6500} source is substantially different than twilight with color temperature of approximately 2000 K. Changing the spectral input increases the minimum illumination. Note that these calculations were based solely upon SNR considerations. The camera must have sufficient gain so that these signals can be displayed.

6.3. INTENSIFIED CCD CAMERA

The construction of the ICCD was described in Section 5.4.2., *Intensified Solid State Cameras*, page 176. As illustrated in Figure 6-5, the photocathode creates photoelectrons that are amplified by the MCP. The amplified electrons are converted back into photons by the phosphor screen. These photons are relayed to the CCD by either a fiberoptic bundle or a relay lens. The CCD creates the photoelectrons that are measured. Each conversion has a quantum efficiency less than unity and each optical element yields a finite transmittance loss.

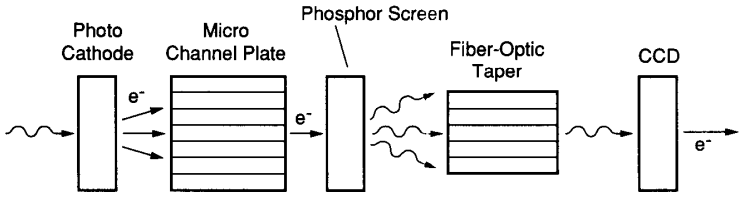


Figure 6-5. Multiple photon-electron conversions occur in an ICCD.

While the image intensifier amplifies the image, the SNR may be reduced due to quantum losses and additive noise. For moderate light levels, the ICCD offers no advantage in SNR performance over the standard CCD camera. Its key attribute is the ability to gate the scene temporally.

6.3.1. ICCD SIGNAL

The ICCD uses a lens to image the target onto the intensifier faceplate. The photon flux reaching the photocathode is

$$\Phi_{CATHODE} = \frac{M_q(\lambda)\, A_{CATHODE}}{4\, F^2 (1 + M_{OPTICS})^2}\, T_{OPTICS} T_{WINDOW} \cdot \tag{6-27}$$

Using the same approximations as given in the preceding section, the number of signal electrons generated by the photocathode is

$$n_{CATHODE} = \frac{\eta_{CATHODE} M_q(\lambda_o)\, \Delta\lambda}{4F^2}\, t_{INT}\, T_{OPTICS} A_{PIXEL}\,, \tag{6-28}$$

where A_{PIXEL} is the projected area of a CCD pixel onto the photocathode. Geometrically, it is the CCD detector area multiplied by the magnification of the fiberoptic bundle. However, A_{PIXEL} depends on fiberoptic bundle characteristics and microchannel pore size. It may be somewhere between 50% of the CCD detector geometric projected area to several times larger.[3] As such, it represents the largest uncertainty in signal-to-noise ratio calculations. The value λ_o is the center wavelength of the photocathode spectral responsivity and $\Delta\lambda$ is the photocathode's spectral bandwidth.

After the microchannel plate, the number of electrons is

$$n_{MCP} = G_{MCP}\, n_{CATHODE} \cdot \tag{6-29}$$

These electrons are then converted to photons by the phosphor screen

$$n_{SCREEN} = \eta_{SCREEN} n_{MCP} \cdot \tag{6-30}$$

The photons travel to the CCD via a fiberoptic bundle whose effective transmittance is T_{fo}. This is a composite value that includes the fiberoptic transmittance and a coupling loss.

$$n_{PHOTON\text{-}CCD} = T_{fo} n_{SCREEN} \cdot \tag{6-31}$$

The number of electrons sensed in the CCD is proportional to its quantum efficiency

$$n_{pe} = \eta_{CCD} n_{PHOTON\text{-}CCD} \tag{6-32}$$

or

$$n_{CCD} = \eta_{CCD}\eta_{SCREEN}\eta_{CATHODE} T_{fo} G_{MCP} \left[\frac{M_q(\lambda_o)\,\Delta\lambda\, t_{INT}\, A_{PIXEL}}{4F^2} \right] . \tag{6-33}$$

For a lens-coupled system, T_{FO} is replaced by an effective relay lens transmittance.[4] Each has its own merits. Transmission losses in a fiberoptic bundle increase proportionally to the square of the minification ratio. This leads to a practical minification of about 2:1. Light is lost because the screen is a Lambertian source whereas the fiberoptic can only accept light over a narrow cone (acceptance angle). But lenses are large and bulky and may provide a throughput as low as 2%. Relay lenses also suffer from vignetting. Here, the "relayed" intensity decreases as the image moves from the center to the periphery of the image intensifier tube.

Because the lens coupling is less efficient than the fiberoptic coupling, the number of electrons produced is lower. The fiberoptic coupled system, with its higher effective transmittance and better sensitivity, can detect lower signals.

6.3.2. ICCD NOISE

The photocathode produces both photon noise and dark noise:

$$\left\langle n^2_{PC\text{-}SHOT} \right\rangle = \left\langle n^2_{CATHODE} \right\rangle + \left\langle n^2_{PC\text{-}DARK} \right\rangle \cdot \tag{6-34}$$

The MCP amplifies these signals and adds excess noise so that

$$\left\langle n_{MCP}^2 \right\rangle = k_{MCP} G_{MCP}^2 \left\langle n_{PC\text{-}SHOT}^2 \right\rangle , \tag{6-35}$$

where k_{MCP} is typically taken as 1.8 but varies with bias voltage.[5,6] The value k_{MCP} is also reported as an SNR degradation. The noise power is transformed into photons by the phosphor screen

$$\left\langle n_{SCREEN}^2 \right\rangle = \eta_{SCREEN}^2 \left\langle n_{MCP}^2 \right\rangle . \tag{6-36}$$

It is attenuated by the fiberoptic bundle

$$\left\langle n_{CCD\text{-}PHOTON}^2 \right\rangle = T_{fo}^2 \left\langle n_{SCREEN}^2 \right\rangle . \tag{6-37}$$

This value is converted to electrons by the CCD and the CCD noise is added

$$\left\langle n_{CCD}^2 \right\rangle = \eta_{CCD}^2 \left\langle n_{CCD\text{-}PHOTON}^2 \right\rangle + \left\langle n_{CCD\text{-}DARK}^2 \right\rangle + \left\langle n_{FLOOR}^2 \right\rangle + \left\langle n_{PRNU}^2 \right\rangle . \tag{6-38}$$

The total noise is

$$\left\langle n_{CCD}^2 \right\rangle = \eta_{CCD}^2 T_{fo}^2 \eta_{SCREEN}^2 k G_{MCP}^2 \left\langle n_{PC\text{-}SHOT}^2 \right\rangle + \left\langle n_{CCD\text{-}SHOT}^2 \right\rangle + \left\langle n_{FLOOR}^2 \right\rangle + \left\langle n_{PRNU}^2 \right\rangle . \tag{6-39}$$

6.3.3. ICCD SNR

The resultant signal-to-noise ratio is

$$SNR = \frac{\eta_{CCD} \eta_{SCREEN} T_{fo} G_{MCP} n_{CATHODE}}{\sqrt{\eta_{CCD}^2 T_{fo}^2 \eta_{SCREEN}^2 k G_{MCP}^2 \left\langle n_{PC\text{-}SHOT}^2 \right\rangle + \left\langle n_{CCD\text{-}SHOT}^2 \right\rangle + \left\langle n_{FLOOR}^2 \right\rangle + \left\langle n_{FPN}^2 \right\rangle}} . \tag{6-40}$$

Several limiting cases are of interest. When the gain is very high, the intensifier noise overwhelms the CCD noise:

$$SNR = \frac{n_{CATHODE}}{\sqrt{k\left[\left\langle n_{CATHODE}^2 \right\rangle + \left\langle n_{PC\text{-}DARK}^2 \right\rangle\right]}} . \tag{6-41}$$

The SNR is independent of gain and CCD characteristics. Under high gain conditions, an expensive, high quantum efficiency or thinned back illuminated CCD is not necessary.

Cooling the photocathode reduces the dark current:

$$SNR = \sqrt{\frac{n_{CATHODE}}{k}} = \sqrt{\frac{\eta_{CATHODE} M_q(\lambda_o)\,\Delta\lambda}{4F^2 k}\, t_{INT}\, T_{OPTICS} A_{PIXEL}}\,. \tag{6-42}$$

High gain is used for photon counting applications. SNR, in this case, is limited by the photocathode quantum efficiency, the excess noise factor, and the optical system.

Under moderate gain situations where the CCD noise is greater than the intensifier noise,

$$SNR = \frac{\eta_{CCD}\eta_{SCREEN} T_{fo} G_{MCP} n_{CATHODE}}{\sqrt{\langle n^2_{CCD\text{-}SHOT}\rangle + \langle n^2_{FLOOR}\rangle + \langle n^2_{PRNU}\rangle}}\,. \tag{6-43}$$

Comparing this with Equation 6-8 (page 186), the ICCD will have lower SNR than a comparable standard camera. The advantage of the ICCD under these conditions is the ability to gate the incoming photon flux temporally.

☞

Example 6-7
IMAGE INTENSIFIER NOISE

Is there a relationship between the image intensifier noise equivalent input and the ICCD noise?

The noise in an image intensifier is due to random emissions (dark current noise) at the photocathode. When referred to the input (SNR = 1), it is called the equivalent background illumination (EBI). Using the same methodology as in Example 6-4, the EBI for an area A_{PIXEL} is

$$EBI = M_q(\lambda_o)\,\Delta\lambda\, t_{INT} A_{PIXEL} = \frac{4\sqrt{k}\,F^2 \langle n_{PC\text{-}DARK}\rangle}{\eta_{CATHODE} T_{OPTICS}}\ photons\,. \tag{6-44}$$

The ICCD will have the same equivalent background input under high gain situations. Cooling the image intensifier reduces the EBI and it may be negligible in gated applications. Under low gain situations, the CCD noise dominates.

Example 6-8
CCD versus ICCD

Under moderate gain situations, what is the relative SNR of the ICCD to the CCD?

The ratio of SNRs is

$$\frac{SNR_{ICCD}}{SNR_{CCD}} = \eta_{SCREEN} T_{fo} G_{MCP} \eta_{PHOTOCATHODE} . \qquad (6\text{-}45)$$

The P20 phosphor screen has a peak emission at 0.56 μm with a conversion efficiency of 0.063. An S20 bialkali photocathode's efficiency is 0.2 at 0.53 μm. The value of T_{fo} is about 0.25. Combined, these factors give 0.00315. The MCP gain must be greater than 1/0.00315 = 317 for the SNR_{ICCD} to be greater than the SNR_{CCD}.

6.4. REFERENCES

1. G. C. Holst, *Electro-Optical Imaging System Performance*, pp. 48-54, JCD Publishing (1995).
2. M. A. Sartor, "Characterization and Modeling of Microchannel Plate Intensified CCD SNR Variations with Image Size," in *Electron Tubes and Image Intensifiers*, C. B. Johnson and B. N. Laprade, eds., SPIE Proceedings Vol. 1655, pp. 74-84 (1992).
3. Mark Sartor, Xybion Corporation, private communication.
4. Y. Talmi, "Intensified Array Detectors," in *Charge-Transfer Devices in Spectroscopy*, J. V. Sweedler, K. L. Ratzlaff, and M. B. Denton, eds., Chapter 5, VCH Publishers, New York (1994).
5. I. Csorba, *Image Tubes*, pp. 120-124, Howard Sams, Indianapolis, Indiana (1985).
6. R. J. Hertel, "Signal and Noise Properties of Proximity Focused Image Tubes," in *Ultrahigh Speed and High Speed Photography, Photonics, and Videography*, G. L. Stradling, ed., SPIE Proceedings Vol. 1155, pp. 332-343 (1989).

7

CRT-BASED DISPLAYS

Displays are used for consumer television receivers, computer monitors, scientific applications, and military applications. A display may be called a monitor, visual display unit (VDU), video display terminal (VDT), or simply a television receiver.

The major performance requirements are listed in Table 7-1. This table is not meant to be all inclusive. It just indicates that different users have different requirements. As displays are built for each user, they are specified in units familiar to that user. For example, consumer displays will not have the MTF or spot size listed. These parameters are provided for militarized displays.

Table 7-1
DISPLAY REQUIREMENTS

APPLICATION	REQUIREMENTS
Consumer TV	"Good" image quality
Computer	Alphanumeric character legibility "Good" graphical capability
Scientific and military	High MTF Unity gamma

Television receivers have been available for over 50 years and the technology has matured. They are considered adequate and little improvement in cathode ray tube (CRT) technology is expected. Advances in television displays currently focus on high definition television, flat panel displays, the legibility of alphanumeric characters, and computer graphics capability. As a result, display specifications are a mix of CRT terminology, video transmission standards, alphanumeric character legibility, and graphics terminology.

System resolution depends on the camera response and limitations imposed by the video standard, display characteristics, and observer acuity. High resolution displays do not offer any "extra" resolution from a system point of view if display is not the limiting factor. Rather, high resolution displays just ensure that all the information available is displayed. In many applications,

system resolution is limited by the human eye. If the observer is too far from the screen, not all of the image detail can be discerned.

In almost all cases, the output device sets the system dynamic range. Often, the display is the limiting factor in terms of image quality and resolution. No matter how good the camera is, if the display resolution is poor, then the overall system resolution is poor.

Flat panel displays are an emerging technology with liquid crystal displays (LCDs) prevalent in notebook (laptop) computers. LCDs require only a fraction of the power required by cathode ray tubes and therefore lend themselves to battery operated systems. However, CRTs will probably dominate the display market due to low cost, high resolution, wide color gamut, high luminance, and long life.

The commercially important flat panels are the light emitting diode (LED) displays, AC plasma display panels (AC PDP). DC plasma display panels (DC PDP), AC thin film electroluminescent (ACTFEL) display, DC thin film electroluminescent (DCTFEL) displays, vacuum fluorescent displays (VFD), and the liquid crystal displays (LCD).

LCD arrays can either be active matrix (AMLCD) or passive arrays. The active matrix contains switching devices on each disel to maintain high brightness and high contrast control. The liquid crystals may be either twisted-nematic (TN) or supertwisted-nematic (STN). The details of these displays can be found in References 1 through 6. Although this chapter discusses CRTs, many concepts apply to all displays.

Emerging technology, new designs, and characterization methods can be found in the *Society of Motion Picture and Television Engineers Journal* and the *Proceedings of the Society of Information Display* symposia. The Society of Information Display offers numerous tutorials, complete with published notes, at their symposia.

In accordance with industry convention, the display sizes and viewing distances are expressed in inches. The majority of this book uses the metric system. The symbols used in this book are summarized in the *Symbol List* (page xviii) which appears after the *Table of Contents*.

7.1. THE OBSERVER

The observer is the starting point for display design. Design is based on both perceptual and physical attributes (Table 7-2). For most applications, the design is driven by perceptual parameters. These parameters are partially related to the physical parameters. For example, sharpness is related to MTF. But the precise relationship has not been quantified. While the physical parameters are important to all applications, they tend to be quantified only for scientific and military displays. Tables 7-1 and 7-2 illustrate different ways of specifying the same requirement: the need for a high quality display.

Maximum image detail is only perceived when the observer is at an optimum viewing distance. If the observer is too far away, the detail is beyond the eye's limiting resolution. It the observer moves closer to the display than the optimum, the image does not become clearer because there is no further detail to see. Display size and resolution requirements are based on an assumed viewing distance.

Table 7-2
PERCEPTIBLE AND PHYSICAL PARAMETERS

PERCEPTUAL PARAMETERS	PHYSICAL PARAMETERS
	Luminance
	Resolution
	Uniformity
Brightness	Addressability
Contrast	Gamma
Sharpness	Color saturation
Color rendition	Color accuracy
Flicker	Color convergence (CRT)
	Distortion (CRT)
	Refresh rate
	MTF

For monochrome systems, the optimum viewing distance occurs when the raster pattern is barely perceptible. At closer distances, the raster pattern becomes visible and interferes with image interpretation. That is, it is considered annoying by most. For color CRTs, the observer must be far enough away so that the individual dots of the three primaries are imperceptible. Otherwise, the individual color spots become visible and the impression of full color is lost.

For design purposes, perceptibility occurs when targets are larger than 1 arc minute. This corresponds to the normal visual acuity of 20/20. This design criterion is quite good. An observer can adjust his distance accordingly. An observer with poorer acuity (e.g., 20/40) will simply move closer to the screen whereas someone with better acuity will probably move further away. Most systems are observer-limited in that the observer cannot see all the detail because he is typically too far from the display.

Figure 7-1 illustrates the optimum viewing distance as a function of display diagonal size for monochrome displays. Television receivers and many computer monitors are specified by the diagonal dimension. Each active line (raster line) subtends 1 arc minute (0.291 mrad). At closer distances, the individual raster lines may become visible. The viewing ratio is the viewing distance divided by the display height. The typical viewing ratio for computer monitors is three.

Television viewers sit about seven times the picture height or 4.2 times the diagonal picture dimension. Knowing the dimensions of a typical living room (sofa on one side, television on the other) suggests that consumer televisions should have a diagonal picture size of 21 to 25 inches. These were the most popular television receiver picture sizes. Sizes have changed reflecting changing television viewing habits.

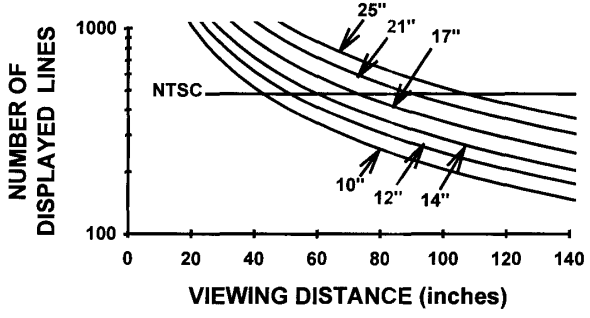


Figure 7-1. Viewing distance as a function of CRT diagonal picture size and number of displayed lines per picture height for monochrome displays. Most CRTs have a 4:3 aspect ratio so that the vertical extent is 60% of the diagonal.

A major design consideration is the elimination of flicker. EIA 170 and NTSC use a 2:1 interface technique which is updated at 30 Hz. For computer monitors which display progressive scan imagery, the frame rate (refresh rate) must be greater than 70 Hz to avoid flicker. Flicker and other design features are not covered here but can be found in numerous books.[2-6]

7.2. CRT OVERVIEW

Figure 7-2 illustrates a typical color cathode ray tube display. Although a color display is illustrated, the block diagram applies to all CRT displays. The input video signal is decoded into its components and then digitized. While digital electronics may be used to process the signal, the CRT is an analog device. Similarly, displays accepting digital inputs must convert the signal into an analog voltage to drive the CRT. Numerous transformations occur to create a visible image. The display matrix and display gain controls allow linear mapping from input image to visible image.[7]

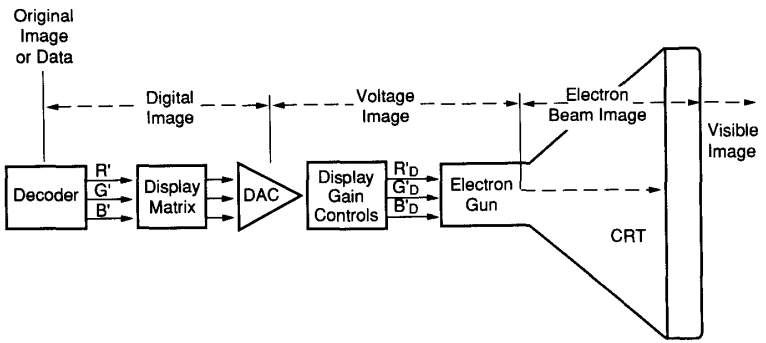


Figure 7-2. Although a color display is illustrated, the block diagram applies to all CRT displays. The digital-to-analog converter (DAC) provides the analog voltages that modulate the electron beam.

The display matrix provides the mapping between the standard video color components and the actual colors emitted by the phosphors on the screen. This color correction circuitry is an algorithm that can be represented as a matrix

$$\begin{pmatrix} R_2 \\ G_2 \\ B_2 \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \begin{pmatrix} R_1 \\ G_1 \\ B_1 \end{pmatrix}. \qquad (7\text{-}1)$$

With the NTSC format, the electron beams paint an image on the CRT phosphor in the scan format shown in Figure 5-2, page 151. The color of the visible image depends on the specific phosphors used. Electron beam density variations, which represent scene intensity differences, create the visible image.

Monochrome displays have a single electron gun and a single phosphor. A color CRT must produce three images (three primary colors) to give a full range of color. To do so requires three electron guns and three phosphors. The three electron beams scan the CRT screen in the same way as a monochrome beam, but arrive at the screen at slightly different angles. The beams strike three different phosphors dots. When the dots (*sub*disels) are beyond the eye's limiting resolution, the eye blends them to create the illusion of a full color spectrum.

The output brightness of a CRT depends on the grid voltage and follows a power law relationship. The slope of this curve, on a log-log scale, is the display gamma (γ):

$$L_{DISPLAY} = k\left(V_{GRID}\right)^{\gamma} . \qquad (7\text{-}2)$$

Gamma can range[8] from one to five. For design purposes, NTSC standardized gamma at 2.2. PAL and SECAM standardized gamma at 2.8.

To provide a linear relationship in the complete camera-display system, the inverse gamma is inserted in consumer cameras as an image processing algorithm (see Section 5.3.3., *Gamma Compensation*, page 169). For scientific applications and computer monitors, the displays will have an internal gamma correction circuit. Here, the user simply supplies the signal with no gamma compensation (e.g., $1/\gamma = 1$). For color displays, each color will have its own gamma look-up table. This operation is performed in the display matrix.

Digital displays provide excellent alphanumeric and graphic imagery. There is no distortion and the disels are accurately located because there is no jitter in timing signals compared to what might be expected with EIA 170. "Digital" seems to imply high resolution but the displayed resolution is limited by the electron beam spot size. Although called "digital," the CRT is an analog device.

7.2.1. MONOCHROME DISPLAYS

In many industrial and scientific applications, monochrome cameras are used. The military uses monochrome monitors extensively to display imagery from non-visible scenes (e.g., thermal and radar imagery). Monochrome displays are found in cockpits and medical imaging instruments. For television receivers, these are simply called black-and-white TVs. Depending on the phosphor used, monochrome displays can provide nearly any color desired. Certain colors (e.g., orange) tend to reduce eye strain. Others (e.g., blue) tend not to disrupt scoptic vision adaption. Although computers have color monitors, word processing is often performed in a monochrome format.

Black-and-white scenes can be imaged on a color monitor by assigning a color palette (false color or pseudo color) to various intensity levels. This does not increase resolution but draws the observer's attention to selected features. The eye can easily differentiate color differences but may have trouble distinguishing small intensity differences.

7.2.2. COLOR DISPLAYS

A color monitor has three electron beams that scan the CRT faceplate. Near the faceplate is a metal mask that has a regular pattern of holes (Figure 7-3). With the dot arrangement, it is called as shadow mask. With slots, it is called an aperture grille. The three guns are focussed onto the same hole and the emerging beams strike the CRT faceplate at different locations. At each location there is a different phosphor. One fluoresces in the blue, one in the red, and one in the green region of the spectrum. The arrangement is such that the beam corresponding to the desired color will strike only the phosphor dot producing that color. The electron guns emit only electrons and have no color associated with them. Color is only created when the beam hits the phosphor even though the guns are called red, green, and blue guns.

Hexagonal structure masks (the dots) are inherently more rigid than slots. However, slots are easier to manufacture and therefore are usually found in television receivers. The slots have a larger area and therefore aperture-grille CRTs tend to be brighter than shadow-mask CRTs. With this brightness it is said that the colors are richer. The distance between dots is less and more difficult to perceive. Shadow-mask CRTs tend to provide sharper images. Because computer monitors are viewed at a close distance, the dots are used most often. The selection of slots over dots depends upon the application.

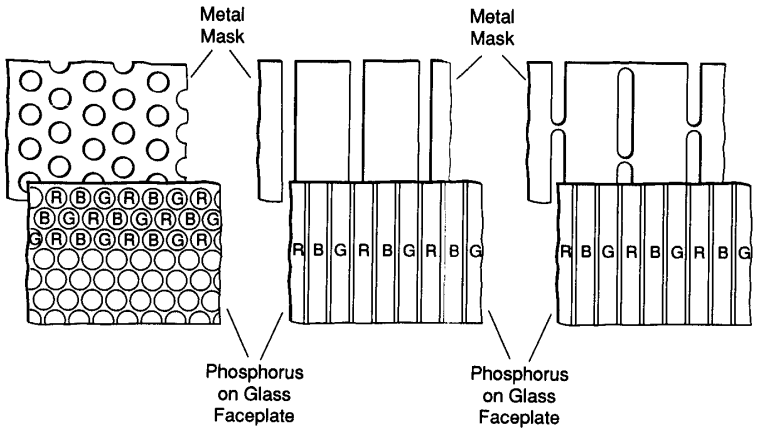


Figure 7-3. Color CRT mask patterns. (a) Dotted, (b) stripes, and (c) slotted. Most consumer television receivers use slots whereas computer displays use dots. The mask design varies with manufacturer.

Dots and stripes cannot be easily compared because the distance between adjacent spots is different and depends upon orientation and the way measured. Because there is $\sqrt{3}/2$ difference, a 0.24-mm aperture-grille pitch is roughly equal to a 0.28-mm shadow mask.

Figure 7-4 illustrates one possible arrangement of color dots. The dot pitch is the center-to-center spacing between adjacent like-colored dots. The three dots create a triad. Because each electron beam passes through the same hole, the mask pitch is the same as the triad pitch. Table 7-3 relates the triad dot pitch to the more common definitions of CRT resolution.

Table 7-3
RESOLUTION OF COLOR DISPLAYS

RESOLUTION	TRIAD DOT PITCH
Ultra-high	< 0.27 mm
High	0.27 - 0.32 mm
Medium	0.32 - 0.48 mm
Low	> 0.48 mm

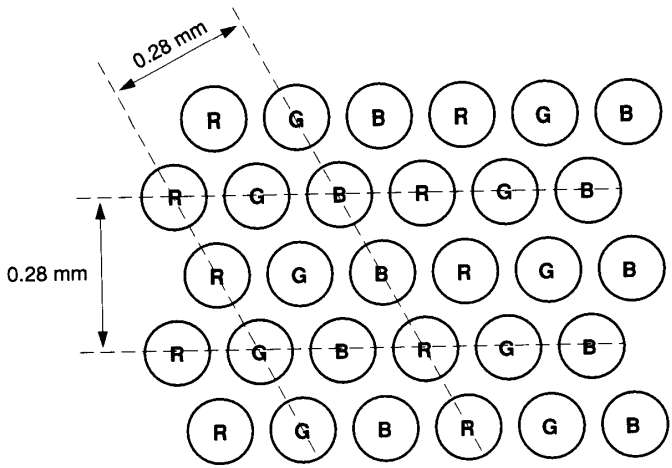


Figure 7-4. Arrangement of phosphor dots on a CRT. There is one shadow-mask hole for each triad. In 1997, most computer monitors had a triad pitch of 0.28 mm. The triad is outlined.

By varying the ratio of red, green, and blue, a variety of effective color temperatures can be achieved. The video signal is simply a voltage with no color temperature associated with it. The display electronics can be adjusted so that a white target appears as if illuminated by a source with color temperature range somewhere between 3200 K and 10,000 K. Most displays are preset to either 6500 K or 9300 K. As the color temperature increases, whites appear to change from a yellow tinge to a blue tinge. The perceived color depends on the adapting illumination (e.g., room lights). Setting the color temperature to 9300 K provides aesthetically pleasing imagery and this setting is unrelated to the actual scene colors (see Section 5.3.2., *Color Correction*, page 165).

Usually there is no relationship between the color temperature and contrast of a scene and what is seen on a display. This is not the fault of the camera and display manufacturers; they strive for compatibility. The observer, who usually has no knowledge of the original scene, will adjust the display for maximum visibility and aesthetics.

For consumer television and consumer computer graphics applications, adequate color imagery is obtained when each primary has 8 bits of intensity (256 shades of gray). The combination creates 24 bits or 16 million colors. Here, each intensity level of each primary is considered a color. It does not indicate that the eye can distinguish all 16 million - only that 16 million combinations are possible. Finer intensity increments may be used in commercial graphics displays. These systems may operate at 10 bits or 12 bits per primary color and provide 2^{30} or 2^{36} "colors," respectively.

7.2.3. HDTV

As the name implies, high definition television (HDTV) provides more resolution than conventional television receivers. The HDTV display has a 16:9 aspect ratio compared to the NTSC 4:3 aspect ratio. It is designed to be viewed at three times the picture height. While the horizontal visual angle subtended by NTSC systems is about 11 degrees, the HDTV subtends about 33 degrees. This larger format gives the feeling of teleprescence - that you are there! The operation of a CRT-based HDTV is the same as any other CRT.

7.3. SPOT SIZE

Display spot size is a complex function of design parameters, phosphor choice, and operating conditions. It is reasonable to assume that the spot intensity profile is Gaussian distributed and is radially symmetric with radius r:

$$L(r) = Ae^{-\frac{1}{2}\left(\frac{r}{\sigma_{SPOT}}\right)^2} . \tag{7-3}$$

As illustrated in Figure 7-5, the spot diameter is defined as the full width at half-maximum (FWHM) and is:

$$S = FWHM = \sqrt{8\ln(2)}\,\sigma_{SPOT} = 2.35\,\sigma_{SPOT} . \tag{7-4}$$

Then

$$L(r) = Ae^{-4\ln(2)\left(\frac{r}{S}\right)^2} . \tag{7-5}$$

The advantage of assuming a Gaussian spot is that many resolution measures are easily compared and the display MTF is easily specified (discussed in Section 10.7., *CRT-based Displays*).

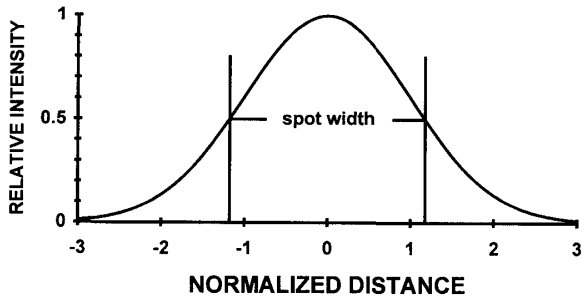


Figure 7-5. The spot diameter is the full width at one-half maximum intensity. The profile has been normalized to σ_{SPOT}.

While the shrinking raster method is listed as a resolution test method, its primary value is in determining the FWHM spot size. With the shrinking raster method, every line is activated (Figure 7-6). If the lines are far apart, the valleys between the lines are visible and the raster pattern is obvious. The lines are slowly brought together until a flat field is encountered. That is, the entire screen appears to have a uniform intensity. The resolution is the number of displayed lines divided by the picture height of the shrunk raster:

$$R_{FF} = \frac{X}{Y} \quad \frac{lines}{cm} \; . \tag{7-6}$$

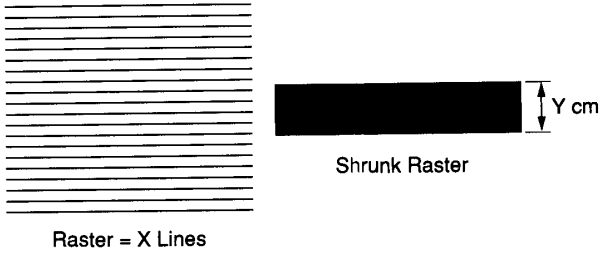


Figure 7-6. Shrinking raster resolution test method. The raster is shrunk until the discrete lines can no longer be discerned.

As the lines coalesce, the MTF drops. Under nominal viewing conditions, experienced observers can no longer perceive the raster when the luminance variation (ripple) is less than 5%. The line spacing has been standardized[9] to $2\sigma_{SPOT}$ (Figure 7-7). Given a shrunk raster resolution of R_{FF} lines/cm, the line spacing in the shrunk raster is $1/R_{FF}$ cm. Then $\sigma_{SPOT} = 1/2R_{FF}$ cm and the spot size is

$$S = FWHM = \frac{2.35}{2R_{FF}} . \qquad (7\text{-}7)$$

While a flat field condition is desirable, the line spacing may be larger. The shrinking raster methodology is applicable for CRTs only and cannot be used for flat panel displays because the "raster" is fixed in those displays.

Sometimes the spot profile tends to be flat-topped. A Gaussian approximation is adequate if the standard deviation, σ_{SPOT}, is calculated from the full width at the 5% intensity level[10]:

$$\sigma_{SPOT} = \frac{FW @ 5\% \ level}{\sqrt{8 \ln(20)}} \approx \frac{FW @ 5\% \ level}{4.9} . \qquad (7\text{-}8)$$

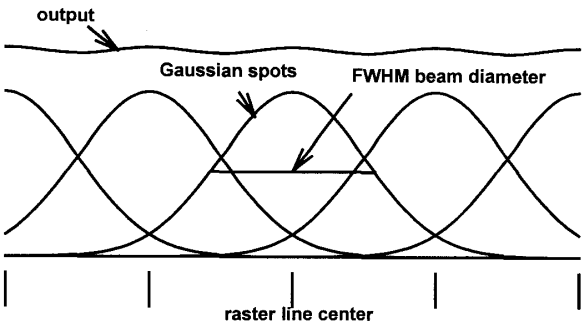


Figure 7-7. When the line spacing is $2\sigma_{SPOT}$, the spots appear to merge and this produces a flat field. The peak-to-peak ripple is 3% and the MTF associated with the residual raster pattern is 0.014. The output is the sum of all the discrete line intensities.

The beam cannot be truly Gaussian because the Gaussian distribution is continuous from $-\infty$ to $+\infty$. Amplifiers may reasonably reproduce the Gaussian beam profile within the $-3\sigma_{SPOT}$ to $+3\sigma_{SPOT}$ limits or a smaller portion. No simple relationship exists between electronic bandwidth and spot size. That is, monitor performance cannot be determined simply by the electronic bandwidth. The bandwidth number quoted depends on the manufacturer's philosophy on rise time issues. Monitors with the same electronic bandwidth may have different resolutions.

7.4. DISELS

Monitor disels are not related to the camera disels. Each is generated by the respective designs. Cameras with a standard video output may not have the resolution suggested by the video standard. For example, a camera containing an array of 320 × 240 pixels may have its output formatted into EIA 170 timing. This standard suggests that it can support approximately 437 × 485 pixels (see Table 5-2, page 153). The signal may be digitized by a frame grabber that creates 1000 × 485 datels. The displayed image still represents 320 × 240 pixels. See Section 5.2.2., *Broadcast/Non-broadcast Design* (page 155), for a discussion of pixels, datels, and disels.

A disel is the smallest imaginary area that can exist on a display (see Section 1.5., *Pixels, Datels, Disels, and Resels*, page 14). In the display literature, "pixel" is used. If necessary, the reader should mentally insert pixel wherever disel is written. The disel size indicates how many individual pieces of information the system can carry. It does not indicate the extent to which the detail will be resolved. The distance between adjoining disels is related to resolution. As the distance decreases, disels overlap and the resolution becomes poorer. This is further discussed in Section 7.6., *Addressability*.

The vertical disel size is defined by the raster pitch. It is simply the display height divided by the number of active scan lines. This is dictated by the video standard. With an optimum system, the electron beam diameter and video bandwidth are matched in the horizontal direction. For television receivers, the number of monitor disels is matched to the video standard. These values were given in Table 5-2, page 153.

The definition of a disel for color systems is more complicated. The representation shown in Figure 7-3 (page 208) implies that the electron beam passes through one hole only. Due to alignment difficulties between the mask and phosphor dots and because the phosphor may have imperfections, CRTs have electron beams that encompass several mask holes. The beam width at the 5% intensity level typically illuminates about 2.5 holes. Encompassing a non-integral number of holes minimizes aliasing and color moiré patterns.[11] The shadow mask samples the continuous electron beam to create a sampled output (Figure 7-8). The phasing effects created by the shadow mask are obvious.

Figure 7-9 illustrates the relationship between the shadow-mask pitch and the beam profile. Assume the shadow-mask pitch is 0.28 mm. If the 5% diameter covers 2.5 holes, then, using Equation 7-9, $\sigma_{SPOT} = (0.28)(2.5)/4.9 = 0.143$ mm. The spot size (Equation 7-4) is S = (2.35)(0.143) = 0.336 mm. In Figure 7-9, the spots are contiguous and each is considered a disel.

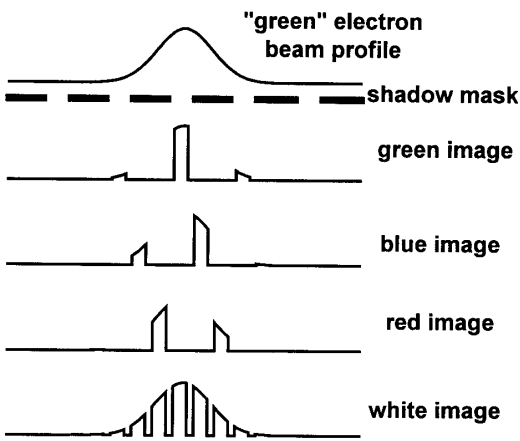


Figure 7-8. Representative image of a single spot on a color CRT. The "green" electron beam is shown, The red and blue beams (not shown) arrive at slightly different angles and provide different intensity profiles. The "sum" appears white when the dot size is less than the eye's resolution limit.

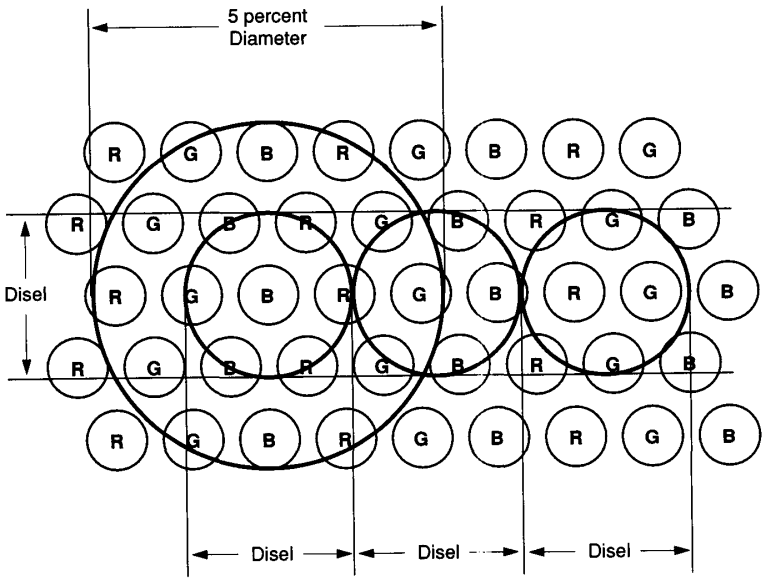


Figure 7-9. Five percent intensity level illuminating 2.5 shadow-mask holes. The spot size (FWHM) is equal to the disel size. If the shadow-mask pitch is X mm, then the beam 5% diameter is 2.5X mm, σ_{SPOT} = 0.51X mm, and the disel is 1.2X mm.

The disel format is the arrangement of disels into horizontal and vertical rows. Table 7-4 lists the common formats found in computer color monitors. The format represents the CRT controller's ability to place the electron beam in a particular location. The CRT is an analog device even though the format is listed as discrete values. These monitors are said to be digitally addressed. The disel format is (erroneously) called the resolution. The relationship between resolution and format is discussed in Section 7.6., *Addressability*.

Disel density is usually specified in units of dots per inch. Table 7-5 provides the approximate relationship between "resolution" and disel density. Confusion exists because dots also refer to the individual phosphor sites. In this context, a "dot" is a disel. The horizontal disel density is the number of horizontal disels (Table 7-4) divided by the horizontal display size.

Table 7-4
FORMAT OF COLOR DISPLAYS

TYPE	FORMAT
CGA (color graphics card adapter)	340 × 200 disels
EGA (extended graphics adapter)	640 × 350 and 640 × 400
VGA (video graphics array)	640 × 480
SVGA (super video graphics array)	800 × 600
XGA (extended graphics array) Typical of multi-sync displays	1024 × 768
High resolution	>1024 × 1024

Table 7-5
TYPICAL DISEL DENSITY

"RESOLUTION"	DISEL DENSITY (dpi)
Ultra-high	> 120
High	71 - 120
Medium	50 - 70
Low	< 50

Figure 7-10 illustrates the "resolution" of three displays. For a fixed format, as the display size increases, the disel size also increases to yield a decreasing disel density. Plotted also is the human eye's resolution capability of 1 arc minute per line expressed at a viewing distance three times the display height. The 1280 × 960 display provides about the highest perceivable resolution. A higher disel density will be beyond the eye's resolution limit and therefore ineffective. While consumer receivers are viewed at 4.2 times the display height, computer monitors are viewed at a closer distance. Figure 7-10 illustrates that the observer is an integral component of perceived resolution. When monitors are designed, a nominal viewing distance is assumed.

The user is typically interested in performance and not in specific design specifications. For example, the dot pitch is typically superfluous information to the user. Disel size and the display modulation transfer function (discussed in Section 10.7., *CRT-based Displays*) are usually more important.

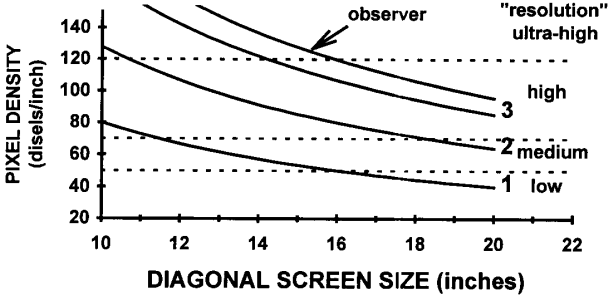


Figure 7-10. "Resolution" of three displays: (1) 640 × 480, (2) 1024 × 768, and (3) 1280 × 960. The observer's resolution is plotted for a viewing ratio of three. The 1280 × 960 display is near the observer's ability to see the disels.

7.5. RESOLUTION

Resolution suggests something about the highest spatial frequency that can be seen. As an electron beam moves across the screen, its intensity can be modulated. The maximum rate at which the beam can be modulated from full off to full on defines the maximum number of (monitor) disels that can be written in a row. The beam moves down vertically to the next line and again travels left to right. The distance the beam moves down is the vertical extent of a disel. While this can be any distance, it is usually selected so that the individual raster lines cannot be seen. Resolution and addressability specify monitor performance. They are combined into the resolution/addressability ratio (RAR). The RAR provides guidance to monitor design.

Resolution metrics are intertwined with television standards such that they are often presented together. Nearly all televisions are built to the same basic design so that by specifying EIA 170 or NTSC, the television receiver bandwidth and spot size is implied. In this sense, a video standard has a "resolution."

Resolution is independent of display size. Displays can be built to any size and the viewing distance affects the *perceived* resolution. For example, if a screen is designed for viewing at 45 inches, then it may have a 10-inch diagonal (see Figure 7-1, page 204). If viewing will be at 90 inches, the entire display can simply be made larger (20-inch diagonal). The larger display has larger raster lines. But the video electronic bandwidth is the same and the resolution remains constant. When viewed at 45 inches, the larger display will appear to have poorer image quality. During design, the selected display size is based upon an assumed viewing distance.

7.5.1. VERTICAL RESOLUTION

The vertical resolution is limited by the number of raster lines. For example, NTSC compatible monitors display 485 lines. However, to see precisely this number, the test pattern must be perfectly aligned with the raster pattern. Because the test pattern may be randomly placed, the pattern at Nyquist frequency can have a zero output if 180° out-of-phase (discussed in Section 8.3.3., *Image Distortion*). To account for random phases, the Kell factor[12] is applied to the vertical resolution to provide an average value. A value of 0.7 is widely used:

$$R_{VERTICAL} = (active\,scan\,lines)\,(Kell\,factor)\ . \qquad (7\text{-}9)$$

Timing considerations force the vertical resolution to be proportional to the number of vertical lines in the video signal. Therefore, NTSC displays have an average vertical resolution of (0.7)(485) = 340 lines. PAL and SECAM displays offer an average of (0.7)(575) = 402 lines of resolution.

7.5.2. THEORETICAL HORIZONTAL RESOLUTION

While the number of horizontal disels was given in Table 5-2, page 153 (see also Section 5.2.2., *Broadcast/Non-broadcast Design*, page 155), the theoretical horizontal resolution is usually normalized to the picture height. If the video electronics has a bandwidth of f_{VIDEO}, then the theoretical number is

$$R_{TVL} = \frac{(2f_{VIDEO})(active\,line\,time)}{aspect\,ratio}\ \frac{TVL}{PH}\ . \qquad (7\text{-}10)$$

The units are television lines/picture height or TVL/PH. Unfortunately, careless usage results in using only "lines." Note that there are two TV lines per cycle. While the video bandwidth can be any value, for consumer applications, it is matched to the video standard. For NTSC television receivers, the nominal horizontal resolution is (2)(4.2 MHz)(52.45 μs)/(4/3) = 330 TVL/PH. PAL and SECAM values are approximately 426 and 465 TVL/PH, respectively. These values change slightly depending on the active line time and bandwidth selected.

Displays designed for scientific applications often operate with standard video format but have a much wider video bandwidth. These monitors can display more horizontal TV lines per picture height (TVL/PH) than that suggested by, say, the NTSC video transmission bandwidth. The horizontal bandwidth can be any value and therefore the theoretical horizontal resolution can be any value.

7.5.3. TV LIMITING RESOLUTION

Whereas the flat field condition exists when two adjacent lines cannot be resolved, the TV limiting resolution is a measure of when alternate vertical bars are just visible (on-off-on-off). The standard resolution test target is a wedge pattern with spatial frequency increasing toward the apex of the wedge. It is equivalent to a variable square wave pattern.

The measurement is a perceptual one and the results vary across the observer population. The industry selected[9] the limiting resolution as a bar spacing of $1.18\sigma_{SPOT}$. This result is consistent with the flat field condition. The flat field condition was determined for two Gaussian beams whereas square waves are used for the TV limiting resolution test. Suppose the Gaussian beams in Figure 7-7 (page 213) were separated by an imaginary "black" beam. Then the beams would be separated by just σ_{SPOT} but the intensity distribution would be identical to Figure 7-7. The black beam does not contribute to the visible image. For TV limiting resolution, the "on" lines are separated by (2)(1.18 σ_{SPOT}). Because bar targets are used, the resultant image is not a precise Gaussian beam and a larger line separation is required ($1.18\sigma_{SPOT}$ versus σ_{SPOT}) for the TV limiting resolution.

The flat field condition and high TV limiting resolution are conflicting requirements. For high TV limiting (horizontal) resolution, σ_{SPOT} must be small. But raster pattern visibility (more precisely, invisibility) suggests that σ_{SPOT} should be large. This is further discussed in Section 7.6., *Addressability*.

Example 7-1
SPOT SIZE versus TV LIMITING RESOLUTION

Consider a 14-inch diagonal monitor used to display an image encoded in the EIA 170 format. Assume the raster line spacing is $2\sigma_{SPOT}$. What is the horizontal resolution?

The vertical extent of the 4:3 aspect ratio monitor is 8.4". With 485 lines, each line is separated by 0.0173". The value σ_{SPOT} is 0.0173/2 = 0.00865" and the spot size (FWHM) is 0.0204".

EIA 170 contains approximately 440 elements in the horizontal direction. The display width is 11.2 inches and each element width is 11.2/440 = 0.0254". But TV limiting resolution suggests that a bar width of (1.18)(0.0865) = 0.102" can be perceived. Because the element width due to the video bandwidth is narrower than the TV limiting resolution, resolution is limited by the spot size. That is, resolution is not limited by the EIA 170 bandwidth.

7.5.4. MTF

TV limiting resolution occurs when the bar spacing is $1.18\sigma_{SPOT}$. Because a cycle consists of two TV lines, the square wave fundamental frequency is $1/(2.35\sigma_{SPOT})$ and the MTF is 0.028. Conversely, the display resolution is selected as that spatial frequency that provides an MTF of 0.028. The MTF equation is found in Section 10.6., *CRT-based Displays*. Any spatial frequency could be used to specify resolution and the slightly different values are cited in the literature. For commercial applications, the resolution is often specified by the spatial frequency where the MTF is 10%.

But square waves are the most popular test patterns. A specification could be: The CTF must be greater than 10% when viewing 320 line pairs. This is equivalent to 640 disels in an on-off-on-off configuration). Note that by using square wave targets, the system response is the contrast transfer function and not the MTF (discussed in Section 9.5., *Contrast Transfer Function*). Careless usage results in calling the system response to square waves, the MTF. This makes it extremely difficult to compare specifications or to calculate performance from published specifications.

When designing a simulation system, the distance to the display and display size are selected first. These are based on physical constraints of the overall system design. The simulation industry generally accepts resolution as that spatial frequency (line pair) which provides a CTF of 10% at this viewing ratio. Because the number of lines depends on what is seen visually, the simulation industry uses units of optical line pairs (OLP). This translates into arc-minutes per line pair.

7.6. ADDRESSABILITY

Addressability is a characteristic of the display controller and represents the ability to select and activate a unique area on the screen. It defines how precisely an electron beam can be positioned on the CRT screen. Addressability is given as the number of discrete lines per picture height or disels per unit length. With this definition, the inverse of addressability is the center-to-center spacing between adjacent disels. Addressability is the image format discussed in Section 7.4., *Disels* (page 213), and provided in Table 7-4 (page 217).

Addressability and resolution are independent of each other. If the resolution is low, successive lines will overwrite preceding lines. If addressability is low, adjacent raster lines will not merge and they will appear as stripes. This was illustrated in Figure 7-7 (page 213) where the line spacing is inverse of the addressability and resolution is the inverse of σ_{SPOT}. Figure 7-7 illustrates good addressability (because lines can be placed close together) but poor resolution (because lines cannot detected when the modulation is low).

There are two opposing design requirements. The first design requirement is that the raster pattern be imperceptible to the observer. This is called the adjacent disel requirement. If the display meets this requirement, the picture will appear uniform and solid. It provides the flat field condition. Alphanumeric characters will appear well constructed and highly legible. The second design requirement is the alternating disel requirement. Here, individual lines (one disel on and one disel off) should be highly visible.

Increased addressability favors the adjacent disel requirement. Placing large spots close together will eliminate the visibility of the raster. But this also reduces the visibility of alternating disels. Similarly an increase in resolution favors the alternating requirement but may make the raster pattern visible.

The resolution/addressability ratio[13] (RAR) is the ratio of FWHM spot size to disel spacing, P:

$$RAR = \frac{S}{P} = \frac{2.36\,\sigma_{SPOT}}{P}, \qquad (7\text{-}11)$$

where P is the disel pitch. For example, if a display is 27.5 cm high and it displays (addresses) 1024 lines, then P is 27.5 cm/1024 lines = 0.27 mm/line. Assuming a 0.28-mm-wide spot then RAR is 0.28/0.27 = 1.04.

As shown in Figure 7-7 (page 213), the summation of periodically spaced spots creates a modulated output. Modulation is defined as $(V_{MAX} - V_{MIN})/(V_{MAX} + V_{MIN})$ where V_{MAX} and V_{MIN} are the maximum and minimum values in the output. Figure 7-11 illustrates the adjacent disel (on-on-on-on) and alternating disel (on-off-on-off) modulation as a function of the RAR. A RAR of 1.0 is considered desirable. In Figure 7-7, the RAR is 1.18.

Before the RAR was defined, display design was based on the experimentally determined flat field condition. The flat condition occurs when $P = 2\sigma_{SPOT}$ or RAR = 1.18. Here, the adjacent disel modulation is near zero - the lines are not visible. The TV limiting resolution describes the alternating disel situation. Now, $P = 1.18\sigma_{SPOT}$, RAR = 2, and the alternating disel modulation is near zero (e.g., the lines are not discernible).

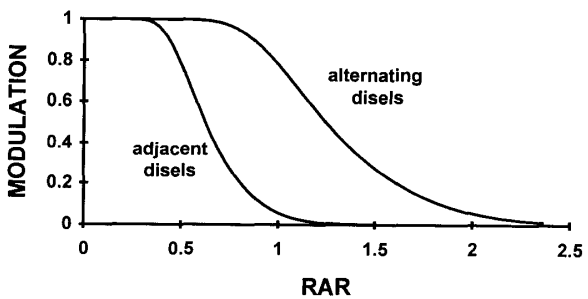


Figure 7-11. Modulation based on Gaussian beam profiles for the adjacent disel and alternating disels. A RAR of one is considered desirable. The modulation of the alternating disel is associated with the display's Nyquist frequency. The modulation of the adjacent disel is associated with the display's sampling frequency. These are imaginary frequencies because the CRT is an analog device. They represent the ability to place a spot at the desired location on the CRT faceplate.

Note that the modulation transfer function is traditionally defined for sinusoidal inputs. The residual ripple is not truly sinusoidal and therefore the modulation is not exactly what would be expected if the input was sinusoidal. That is, the modulation shown in Figure 7-11 is not a true MTF. The measured display response may provide different values than those shown in Figure 7-11.

The alternating disel modulation represents one Gaussian beam on and one off. This is different from one bar on and one bar off. Resolution tests are performed with square waves. For an analog system, the contrast transfer function is $4/\pi$ times greater than the MTF (discussed in Section 9.5., *Contrast Transfer Function*).

If the horizontal format supports 640 disels (see Table 7-4, page 217), then 320 cycles should be perceived. But the modulation is not specified at this resolution. For high quality monitors, the 320 cycles are easily visible (RAR $<$ 1). With average monitors, the 320 cycles are visible but with reduced modulation (RAR $\approx$ 1). On a poorly designed monitor, the cycles will be barely perceived (RAR $>$ 1).

While any disel format can be specified, the RAR determines if the disels are visible. For the configuration shown in Figure 7-9, (page 216), if the shadow mask pitch is 0.28 mm, then the disels will be separated by 0.336 mm when the RAR is one. If using a 13-inch diagonal display, the width is 264 mm and 786 horizontal disels are viewable when the RAR = 1. More disels can be displayed, but with reduced modulation (Table 7-6 and Figure 7-12). Thus the disel format, by itself, does not fully characterize the display capability. For optimized displays, it is reasonable to assume that the RAR is about one.

Table 7-6
ALTERNATING DISEL MODULATION
13-inch diagonal display and 0.28-mm shadow-mask pitch
5% intensity level illuminating 2.5 shadow-mask holes

DISELS	RAR	MODULATION
800	1.02	0.75
900	1.15	0.61
1000	1.27	0.47
1100	1.40	0.35
1200	1.53	0.25

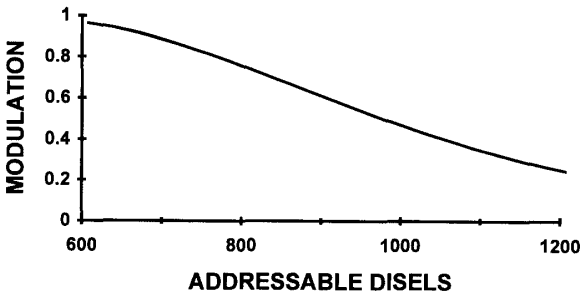


Figure 7-12. 13-inch diagonal display alternating disel (Gaussian beam) modulation. The shadow-mask pitch is 0.28 mm. The 5% intensity level illuminates 2.5 shadow-mask holes.

ANSI/HFS 100-1988 requires that the adjacent disel modulation be less than 20%. This deviates significantly from the flat field condition (adjacent disel modulation is 0.014). ISO 9241 recommends that the alternating disel modulation be greater than 0.4 for monochrome and greater than 0.7 for color systems. For monochrome systems, the RAR will vary from 0.8 to 1.31. For color systems the RAR varies from 0.8 to 1.04. These standards were developed to assure legibility of alphanumeric characters.

If the RAR is too high, then the output of two adjacent "on" lines is much greater then the output of two separated lines due to the summing effect (Figure 7-13). For the flat field condition (Figure 7-7, page 213), the RAR is 1.18 and the summed lines are 27% brighter than an individual line. If the RAR is very large, the intensity of several adjacent lines "on" can be significantly brighter than a single line "on". Thus, there is also a tradeoff in intensity with RAR.

Multi-sync monitors automatically adjust the internal line rate to the video line rate. These monitors can display any format from 525 lines up to 1225 lines. Because the displayed image size remains constant, the vertical line spacing changes. At the 1225 line format, the raster lines are close together and form a uniform (flat) field. At the lower line rates, the individual raster lines are separated and may be perceived. Because the same video amplifiers are used for all line rates, the horizontal resolution is independent of the line rate.

That is, the RAR in the horizontal direction stays constant whereas the RAR in the vertical direction changes with the number of lines. If the vertical RAR is one at 1225 lines, then it must be 525/125 = 0.43 at 525 lines.

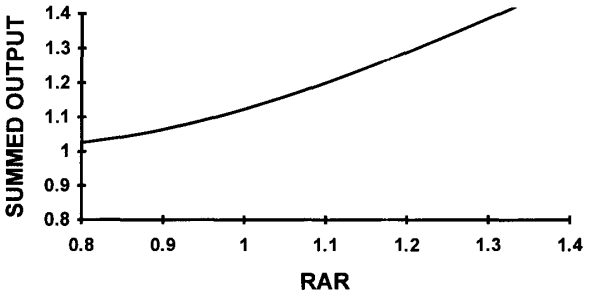


Figure 7-13. Summed output of contiguous lines "on". Ideally, the intensity should remain constant whether one or multiple lines are "on".

Example 7-2
DISPLAY RESOLUTION

A multi-sync monitor has a horizontal resolution of 800 TV lines per picture height. If the display diagonal is 14", what is the spot size? Are the raster lines visible?

For a 4:3 aspect ratio display, the CRT vertical extent is 284 mm. The estimated spot size is 284/800 = 0.355 mm and σ_{SPOT} is S/2.35 = 0.335/2.35 =0.15 mm. If the number of displayed lines is N_{LINE}, the RAR is

$$RAR = \frac{S N_{LINE}}{H_{MONITOR}} = \frac{0.355}{284} N_{LINE} = \frac{N_{LINE}}{800} . \qquad (7\text{-}12)$$

With 485 displayed lines (EIA 170), the RAR is 0.61. From Figure 7-11 the raster pattern (adjacent disels) is quite visible. For EIA 343A with 809 active lines, the RAR is at the desired value of one. At 946 active lines, the RAR increases to 1.18. Here the alternating disel modulation starts to drop and the vertical resolution starts to decrease.

7.7. SHADES OF GRAY

An often used, but widely misunderstood, concept is "shades of gray." While displays may be specified in shades of gray where each step is square root of 2, human vision cannot. The misunderstanding originated with the thought that the eye requires a $\sqrt{2}$ increment to just detect an increment in luminance. Studies of visual perception clearly indicate that just noticeable increments in luminance vary with background luminance level, size and shape of the target, the ambient luminance, and a variety of other variables. To use a single number to characterize the human eye is too simplistic. The $\sqrt{2}$ approach is merely another way to indicate display dynamic range. No other inference should be made about the definition.

7.8. CHARACTER RECOGNITION

Computer monitors are designed to for maximum legibility of alphanumeric characters and enhanced graphical capability. Readable characters can be formed in any block of disels from 5×7 to 9×16 and still be aesthetically pleasing. More complex symbols such as Japanese characters may require 16×16 disels.

The ability to see detail is related to the RAR. Figure 7-14 illustrates three letters and the resultant intensity traces. The RAR must be near one so that a reasonable contrast ratio exists between "on" and "off" disels. With reasonable contrast, the inner detail of the character is seen and the character is legible. Similarly, with a reasonable contrast ratio, adjacent letters will appear as separate letters. With characters, the alternating disel pattern is called one stroke separated by a space.

Figure 7-15 illustrates how reduced modulation affects character legibility. For easy interpretation, the system contains a shadow mask but only white and black can be seen. The image is magnified so that the individual phosphor dots can be seen. The sampled nature of the Gaussian electron beam is evident. As the beam width increases, the modulation decreases and the RAR increases. The characters become blurry as the RAR increases.

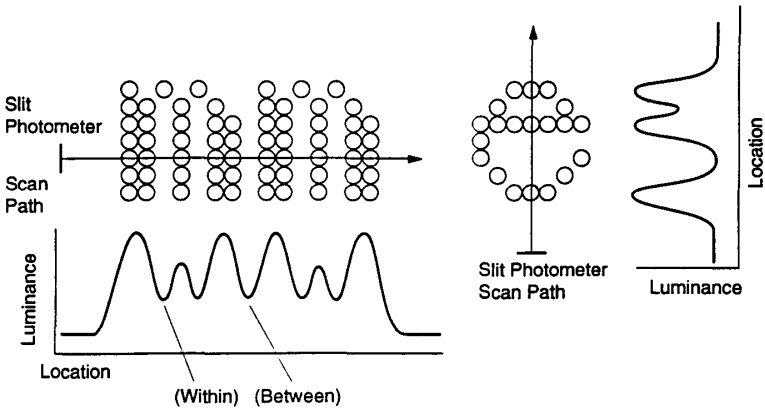


Figure 7-14. The RAR must be near one so that the inner details of characters are visible and the double width appears solid. Similarly with RAR near one, two adjacent characters appear separate. Each dot represents a disel. Lower RARs will make the characters appear very sharp. However, the raster pattern will become visible and the characters will not appear solid.

7.9. CONTRAST

Selection of one minute of arc as a measure of human visual acuity provided back-of-the-envelope calculations for selecting disel size, display size, and viewing distance. However, the eye's response varies according to the displayed contrast.

The air-glass interface and the phosphor reflect ambient light. This reduces contrast between disels and reduces the display's color gamut. As a result, most displays are measured in a darkened room where the reflected light is less than 1% of the projected light. Typically the contrast is measured with a checkerboard pattern consisting of 16 alternating rectangles. The average value of the eight white rectangles and the average value of the eight black rectangles are used to calculate the contrast ratio.

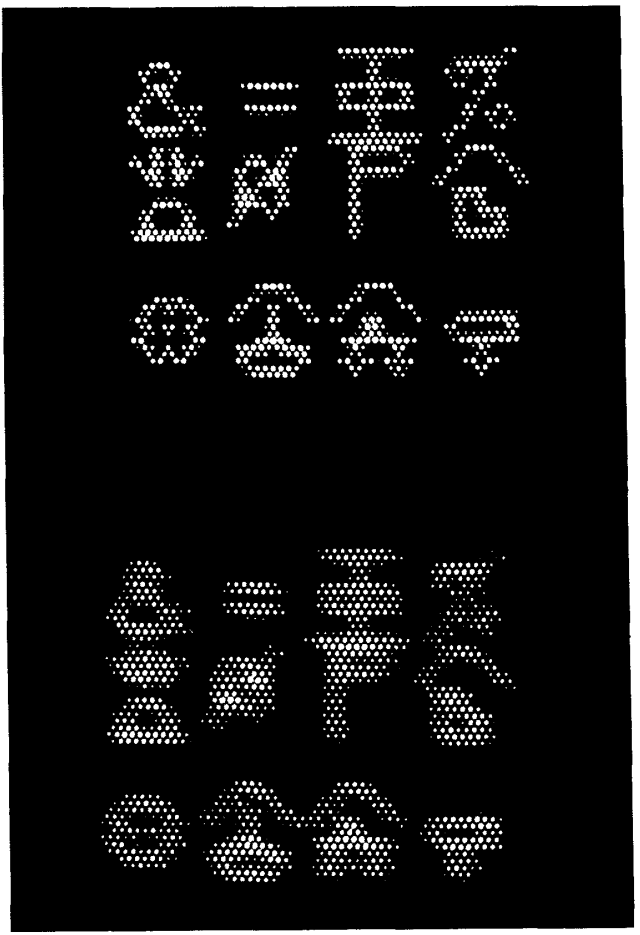


Figure 7-15. Complex symbols are more legible when the display modulation is high. Each dot represents a hole in the shadow mask. Top: modulation = 0.5 at the Nyquist frequency. Bottom: modulation = 0.1 at the Nyquist frequency. When viewed at 10 feet, the individual dots are no longer perceptible. At this distance, the top characters appear sharp (RAR $\approx$ 1.25) and the bottom ones appear fuzzy (RAR $\approx$ 1.8). By courtesy of the Mitre Corporation.

The perceived contrast ratio is

$$Contrast\ ratio = \frac{L_{WHITE} + L_{AMBIENT} \cdot reflectivity}{L_{BLACK} + L_{AMBIENT} \cdot reflectivity}, \qquad (7\text{-}13)$$

where L_{WHITE} and L_{BLACK} represent the luminances from the white and black rectangles, respectively. The reflections are the same regardless if a disel is on or off. About 4 percent of the ambient light ($L_{AMBIENT}$) is reflected at the air-glass interface. The CRT phosphor is an excellent Lambertian reflector with a typical reflectivity of 70%. "Black" displays minimize these reflections and thereby increase the perceived contrast.

Unfortunately, specifications that are based on no ambient lighting can only be achieved in a totally darkened room. Safety considerations require some lighting in all rooms so that the specification can never be achieved in practice.

The contrast ratio for typical displays is about 20:1 under normal ambient lighting conditions. This decreases as the ambient light increases. The eye operates in the reverse way. As the ambient lighting increases, more contrast is necessary to see target-background differences. Thus, as the displayed contrast decreases in bright light, the eye requires more contrast. This explains why television imagery is so difficult to see in sunlight and why it appears so bright in a darkened room.

7.10. REFERENCES

1. L. E. Tannas Jr., ed., *Flat panel Displays and CRTs*, Van Nostrand, New York, NY (1985).
2. J. A. Castellano, *Handbook of Display Technology*, Academic Press, New York, NY (1992).
3. S. Sherr, *Electronic Displays*, Second edition, Wiley and Sons, New York, NY (1993).
4. J. C. Whitaker, *Electronic Displays*, McGraw-Hill, New York, NY (1994).
5. J. Peddie, *High-Resolution Graphics Display Systems*, Windcrest/McGraw-Hill, NY (1994).
6. L. W. MacDonald, ed., *Display Systems, Design and Applications*, Wiley, NY (1997).
7. W. Cowan, "Displays for Vision Research," in *Handbook of Optics*, M. Bass, ed., Second edition, Vol. 1, Chapter 27, McGraw-Hill, New York. NY (1995).
8. T. Olson, "Behind Gamma's Disguise," *SMPTE Journal*, Vol. 104(7), pp. 452-458 (1995).
9. L. M. Biberman, "Image Quality," in *Perception of Displayed Information*, L. M. Biberman, ed., pp. 13-18, Plenum Press, New York (1973).
10. P. G. J. Barten, "Spot Size and Current Density Distributions of CRTs," in Proceedings of the SID, Vol. 25(3), pp. 155-159 (1984).
11. A. Miller, "Suppression of Aliasing Artifacts in Digitally Addressed Shadow-mask CRTs," *Journal of the Society of Information Display*, Vol. 3(3), pp 105-108 (1995).
12. S. C. Hsu, "The Kell Factor: Past and Present," *SMPTE Journal*, Vol. 95, pp. 206-214 (1986).
13. G. M. Murch and R. J. Beaton, "Matching Display Resolution and Addressability to Human Visual Capacity," *Displays*, Vol. 9, pp. 23-26 (Jan 1988).

8

SAMPLING THEORY

Sampling is an inherent feature of all electronic imaging systems. The scene is spatially sampled in both directions due to the discrete locations of the detector elements. This sampling creates ambiguity in target edges and produces moiré patterns when viewing periodic targets. Aliasing becomes obvious when (a) the image size approaches the detector size and (b) the detectors are in a periodic lattice (the normal situation). Spatial aliasing is rarely seen in photographs or motion pictures because the grains are randomly dispersed in the photographic emulsion.

If the detector size and spacings are different in the horizontal and vertical directions, sampling effects will be different in the two directions. This leads to moiré patterns that are a function of the target orientation with respect to the array axis. Although sampling is two-dimensional, for simplicity sampling effects will be presented in one-dimension.

The highest frequency that can be faithfully reconstructed is one-half the sampling rate. Any input signal above the Nyquist frequency, f_N (which is defined as one-half the sampling frequency, f_S), will be aliased down to a lower frequency. That is, an undersampled signal will appear as a lower frequency after reconstruction (Figure 8-1).

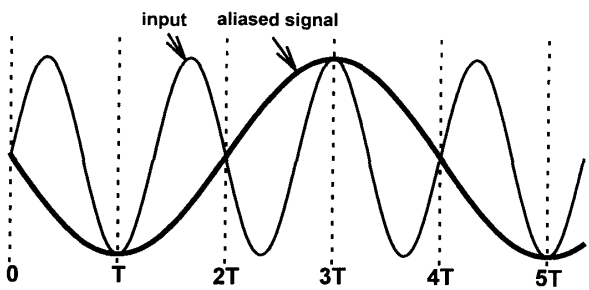


Figure 8-1. An undersampled sinusoid will appear as a lower frequency after reconstruction. The sampling frequency is $f_S = 1/T$. When T is the measured in mrad, mm, or time, the sampling frequency is expressed in cycles/mrad, cycles/mm, or Hz, respectively.

After aliasing, the original signal can never be recovered. It would appear that aliasing is an extremely serious problem. However, the extent of the problem depends upon the final interpreter of the data.

Signals can be undersampled or oversampled. Undersampling is a term used to denote that the input frequency is greater than the Nyquist frequency. It does not imply that the sampling rate is inadequate for any specific application. Similarly, oversampling does not imply that there is excessive sampling. It simply means that there are more samples available than that required by the Nyquist criterion.

After aliasing, the original signal can never be recovered. Undersampling creates moiré patterns (Figure 8-2). Diagonal lines appear to have jagged edges or "jaggies." Periodic structures are rare in nature and aliasing is seldom reported when viewing natural scenery although aliasing is always present. It may become apparent when viewing periodic targets such as test patterns, picket fences, plowed fields, railroad tracks, and Venetian blinds. Distortion effects can only be analyzed on a case-by-case basis. Because the aliasing occurs at the detector, the signal must be band-limited by the optical system to prevent it.

We have become accustomed to the aliasing in commercial televisions. Periodic horizontal lines are distorted due to the raster. Cloth patterns such as herringbones and stripes produce moiré patterns. Cross-color effects occur in color imagery. Many video tape recordings are undersampled to keep the price modest and yet the imagery is considered acceptable when observed at normal viewing distances.

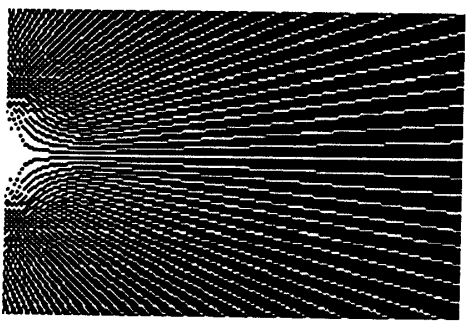


Figure 8-2. A raster scan system creates moiré patterns when viewing wedges or star bursts.

Aliasing is reduced to an acceptable level in consumer color sensors. For specific applications such as medical imaging, it may be considered unacceptable. In medical imagery, an aliased signal may be misinterpreted as a medical abnormality that requires medication, hospitalization, or surgery.

While aliasing is never considered desirable, it is accepted in scientific and military monochrome applications where the MTF must be maintained. High MTF is related to image sharpness. Aliasing becomes bothersome when the scene geometric properties must be maintained as with mapping. It affects the performance of most image processing algorithms.

Sampling theory states that the *frequency* can be unambiguously recovered for all input frequencies below Nyquist frequency. It was developed for band-limited electrical circuits. The extension to imaging systems is straightforward. However, the eye is primarily sensitive to intensity variations.

Sampling theory is traditionally presented from a modulation transfer function viewpoint. MTF theory is presented in Chapter 9 and the subsystem MTFs are discussed in Chapter 10. Linear system theory is used for system analysis because of the wealth of mathematical tools available. Approximations are used to account for sampling effects. While these approximations are adequate to describe the MTFs, they do not indicate how imagery may be distorted.

Sampled data systems are nonlinear and do not have unique MTFs.[1-6] The "MTF" depends on the phase relationships of the scene with the sampling lattice. Superposition does not hold and any MTF derived for the sampler cannot, in principle, be used to predict results for general scenery. To account for the nonlinear sampling process, a sample-scene MTF is added (discussed in Section 10.10., *Sampling Effects*). As the sampling rate increases, the MTF becomes better defined. As the sampling rate approaches infinity, the system becomes an analog system and the MTF is well defined. Equivalently, as the detector size and detector pitch decrease, signal fidelity increases. Reference 7 provides an in-depth discussion on sampling effects. This chapter highlights those results.

The symbols used in this book are summarized in the *Symbol List* (page xviii) and it appears after the *Table of Contents*.

8.1. SAMPLING THEOREM

The sampling theorem, as introduced by Shannon,[8] was applied to information theory. He stated that if a time-varying function, v(t), contains no frequencies higher than f_{MAX} (Hz), it is completely determined by giving its ordinates at a series of points spaced $1/2f_{MAX}$ sec apart. The original function can be reconstructed by using an ideal low-pass filter. To just satisfy Shannon's theorem, the sampling frequency, f_S, must be $2f_{MAX}$. Shannon's work is an extension of others[9] and the sampling theorem is often called the Whittaker-Shannon theorem.

Three conditions must be met to satisfy the sampling theorem. The signal must be band-limited, the signal must be sampled at an adequate rate, and a low-pass reconstruction filter must be present. When any of these conditions are not present, the reconstructed analog data will not appear exactly as the originals.

In a sampled-data system, the sampling frequency interacts with the signal to create sum and difference frequencies. Any input frequency, f_o, will appear as $nf_S \pm f_o$ after sampling ($n = -\infty$ to $+\infty$). Figure 8-3 illustrates a band-limited system with frequency components replicated by the sampling process. The base band ($-f_H$ to f_H) is replicated at nf_S. To avoid distortion, the lowest possible sampling frequency is that value where the base band adjoins the first side band (Figure 8-3c). This leads to the sampling theorem that a band-limited system must be sampled at twice the highest frequency ($f_S \geq 2f_H$) to avoid distortion in the reconstructed image. The sampling theorem applies to any periodic input frequency. The units can be either spatial frequency or electrical frequency.

After digitization, the data reside in data arrays (e.g., computer memory location) with nonspecific units. The user assigns units to the arrays during image reconstruction. That is, the data are read out of the memory in a manner consistent with display requirements. Because monitors are analog devices, the data are transformed into an analog signal using a sample-and-hold circuit. This creates a "blocky" image and can be removed by a reconstruction filter (discussed in Section 10.6.4., *Post-Reconstruction Filter*). If the original signal was oversampled and if the post-reconstruction filter limits frequencies to f_N, then the reconstructed image can be identical to the original image.

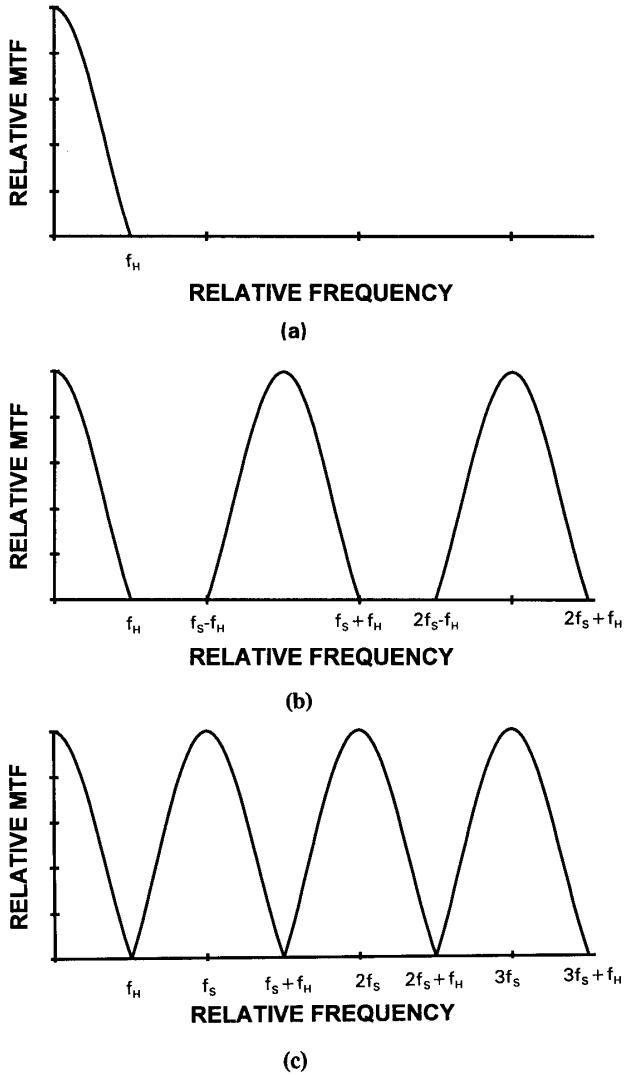


Figure 8-3. Sampling replicates frequencies at $nf_S \pm f$. (a) Original band-limited signal, (b) frequency spectrum after sampling, and (c) when $f_S = 2f_H$, the bands just adjoin.

8.2. ALIASING

As the sampling frequency decreases, the first side band starts to overlap the base band. It is the summation of these power spectra that create the distorted image (Figure 8-4). The overlaid region creates distortion in the reconstructed image. This is aliasing. Once aliasing has occurred, it cannot be removed. Within an overlapping band, there is an ambiguity in frequency. It is impossible to tell whether the reconstructed frequency resulted from an input frequency of f_o or $nf_S \pm f_o$.

All frequency components above f_N are folded back into the base band so that the base band contains

$$O_{BASE\text{-}BAND} = \sum_{n=-\infty}^{\infty} O(nf_S \pm f) \quad where\ nf_S \pm f \le f_N\ . \tag{8-1}$$

To avoid aliasing in the electronic domain, the signal may be passed though a low-pass filter (anti-aliasing filter). These filters cannot remove the aliasing that has taken place at the detector. They can only prevent further aliasing that might occur in the downstream analog-to-digital converter. Figure 8-5 illustrates the cutoff features of an ideal filter. The ideal filter passes all frequencies below f_N and attenuates all frequencies above f_N. The ideal filter is, of course, unrealizable.[10] It also can produce undesirable effects such as ringing.

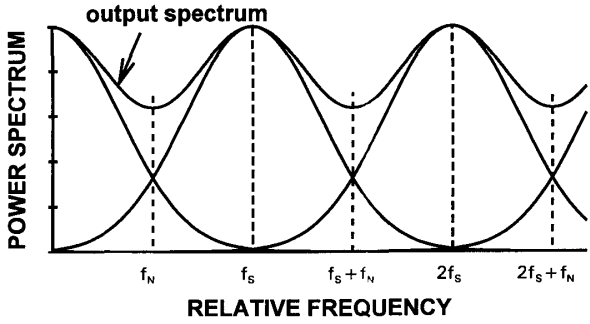


Figure 8-4. Aliasing alters both the signal and noise base band spectra. In a real system, the signal and noise are not necessarily band-limited and some aliasing occurs.

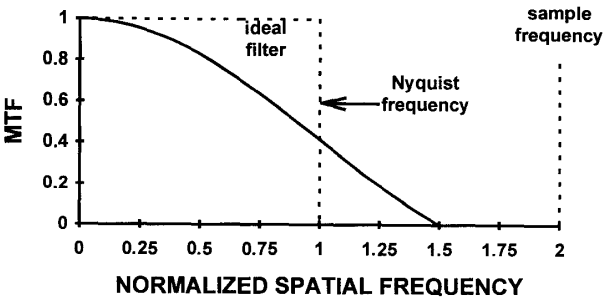


Figure 8-5. An undersampled system must be band-limited to avoid aliasing. The ideal anti-aliasing filter passes all the signals below f_N and no signal above f_N. Whether aliasing is considered undesirable depends on the application.

In solid state cameras, the sampling frequency is determined by the detector pitch. That is, the scene is optically sampled by the detectors. If the scene contains frequencies greater then the array Nyquist frequency, aliasing will occur. Optical band-limiting can be achieved by using small diameter optics, defocusing, blurring the image, or by inserting a birefringent crystal between the lens and array. The birefringent crystal changes the effective sampling rate and is found in almost all single chip color cameras (discussed in Section 10.3.4., *Optical Anti-Alias Filter*). Unfortunately, these approaches also degrade the MTF (reduce image sharpness) in the base band and typically are considered undesirable for scientific applications. For medical applications, where aliased imagery may be misinterpreted as an abnormality, reduced aliasing usually outweighs any MTF degradation.

According to the sampling theorem, any sampled frequency, f, where $f \leq f_N$, can be uniquely recovered. However, the square wave is the most popular test target and it is characterized by its fundamental frequency f_o. When expanded into a Fourier series, the square wave consists of an infinite number of frequencies (discussed in Section 9.5., *Contrast Transfer Function*). Although the square wave fundamental may be oversampled, the higher harmonics will not. The appearance of the square wave after reconstruction depends on the relative values of the optical, detector, and electronic MTFs.

Although no universal method exists for quantifying aliasing, it is reasonable to assume that it is proportional to the MTF that exists above f_N. In Figure 8-6 a significant MTF exists above the Nyquist frequency. Whether this aliasing is objectionable, depends on the camera application.

If a system is Nyquist frequency limited, then the Nyquist frequency is used as a measure of resolution. Because no frequency can exist above the Nyquist frequency, many researchers represent the MTF as zero above the Nyquist frequency (Figure 8-7). This representation may be too restrictive for modeling purposes.

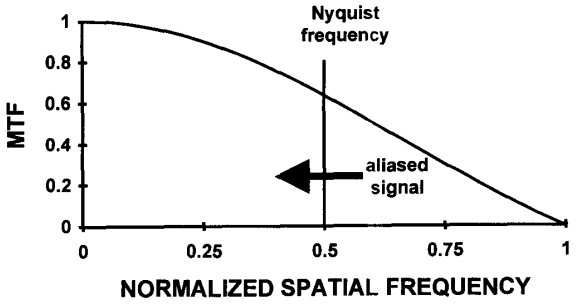


Figure 8-6. Signals above Nyquist frequency are aliased down to the base band. The area bounded by the MTF above f_N may be considered as an "aliasing" metric.[11]

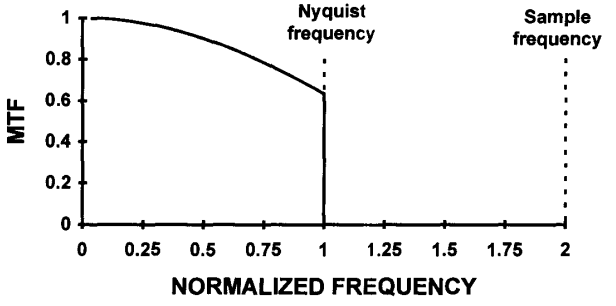


Figure 8-7. MTF representation of an undersampled system as a function of f/f_N. This restrictive representation erroneously suggests that there is no response above f_N.

8.3. The DETECTOR as a SAMPLER

Figure 8-1 (page 231) illustrated a signal sampled at discrete points. It typifies extremely small detectors and flash analog-to-digital converters. Detectors have a finite size and integrate the signal spatially. This spatial integration reduces the modulation and is the detector MTF.

Sections 8-1 and 8-2 described sampling in generic terms. The variable f was used to denote frequency. There are four different frequency domains that can exist with an electro-optical imaging system (discussed in Section 10.1., *Frequency Domains*). Imagery at the detector plane is described by image-space spatial frequency. The horizontal and vertical variables are represented by u_i and v_i, respectively. They have units of cycles/mm.

8.3.1. DETECTOR MTF

Figure 8-8 illustrates the spatial integration afforded by a staring array. In Figure 8-8a, the detector width is one-half of the center-to-center spacing (detector pitch). That is, only 50% of the input sinusoid is detected. The detector averages the signal over the range indicated and the heavy lines are the detector outputs. The output MTF is $(V_{MAX} - V_{MIN})/(V_{MAX} + V_{MIN})$ where V_{MAX} and V_{MIN} are the maximum and minimum voltages, respectively. As the detector size increases, it integrates over a larger region and the MTF decreases (compare Figure 8-8a with 8-8b). As the phase changes between the detector location and the peak of the sinusoid, the output voltages change. This creates a variable MTF caused by sampling.

The detector MTF cannot exist by itself. Rather, the detector MTF must also have the optical MTF to make a complete imaging system. In the horizontal direction, the MTF of a single rectangular detector is

$$MTF_{DETECTOR}(u_i) = \left| \frac{\sin(\pi d_H u_i)}{\pi d_H u_i} \right| , \qquad (8\text{-}2)$$

where d_H is the physical extent of the photosensitive surface.

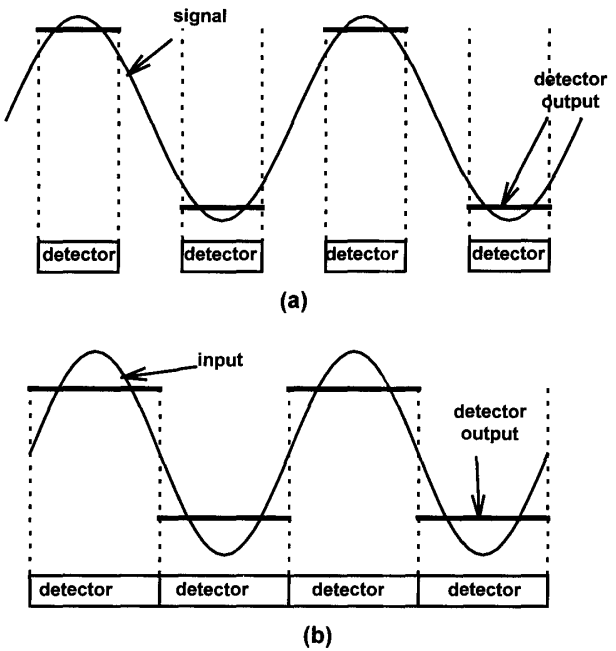


Figure 8-8. A detector spatially integrates the signal. Figure 8-1 (page 231) represents the output of extremely small detectors. (a) Detector width is one-half of the center-to-center spacing and (b) detector width is equal to the center-to-center spacing. (a) Typifies an interline transfer CCD array in the horizontal direction (see Figure 3-16, page 64) and (b) is representative of frame transfer arrays. The heavy lines are the detector output voltages. As the phase changes, the output also changes.

Figure 8-9 illustrates the detector MTF in one dimension. The MTF is equal to zero when $u_i = k/d_H$. The first zero ($k = 1$) is considered the detector cutoff, u_{iD}. Any input spatial frequency above detector cutoff will be aliased down to a lower frequency. It is customary to plot the MTF only up to the first zero (Figure 8-10). The optical system cutoff limits the absolute highest spatial frequency that can be faithfully imaged. The detector cutoff may be either higher (optics-limited system) or lower (detector-limited system). Most CCD-based imaging systems are detector-limited (discussed in Section 10.4., *Optics-Detector Subsystem*).

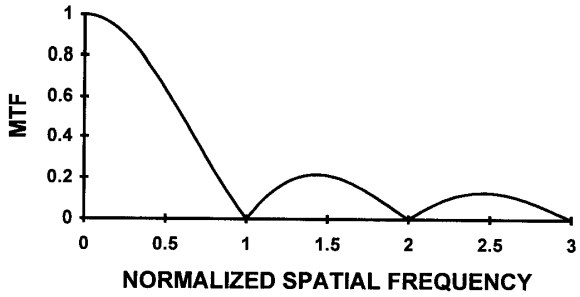


Figure 8-9. Detector MTF as a function of normalized spatial frequency u_i/u_{iD}. Detector cutoff is $u_{iD} = 1/d$.

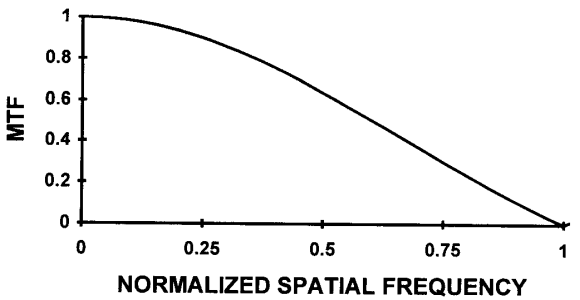


Figure 8-10. Typical detector MTF representation as a function of normalized spatial frequency u_i/u_{iD}. This representation erroneously suggests that the detector does not respond to frequencies greater than u_{iD}.

8.3.2. DETECTOR ARRAY NYQUIST FREQUENCY

A staring array acts a sampler whose sampling rate is $1/d_{CC}$ (cycles/mm) where d_{CC} is the detector center-to-center spacing (detector pitch). The pitch in the horizontal and vertical directions, d_{CCH} and d_{CCV}, respectively, may be different. Detector arrays are specified by the fill factor which is the detector area divided by the cell area. It is $d_H d_V / d_{CCH} d_{CCV}$. Staring arrays can faithfully reproduce horizontal signals up to $u_{iN} = 1/(2d_{CCH})$. Although the individual detectors can reproduce higher spatial frequencies, the horizontal scene spectrum is sampled at $u_{iS} = 1/d_{CCH}$ (Figure 8-11).

Microlenses increase the effective detector size but do not affect the Nyquist frequency. By increasing the detector area, the detector cutoff decreases. With microlenses, the array in Figure 8-11b can optically appear as the array in Figure 8-11a.

Staring arrays are inherently undersampled when compared to the detector spatial frequency cutoff. Any scene spatial frequency above u_{iN} or v_{iN} will appear as a lower spatial frequency in the reconstructed image. To avoid aliasing, the scene frequencies must be optically band-limited similar to that shown in Figure 8-5 (page 237).

Most single chip color filter arrays have an unequal number of red, green, and blue detectors. A typical array designed for NTSC operation will have 768 detectors in the horizontal direction with 384 detectors sensitive to green, 192 sensitive to red, and 192 sensitive to the blue region of the spectrum. Suppose the arrangement is G-B-G-R-G-B-G-R. The "green" detectors will have a center-to-center spacing as shown in Figure 8-11b.

The spacing of the "blue" and "red" detectors is twice the "green" detector spacing. This produces a "blue" and "red" array Nyquist frequency that is one-half of the "green" array Nyquist frequency (Figure 8-12). Other detector layouts will create different array Nyquist frequencies. The "color" Nyquist frequencies can be different in the horizontal and vertical directions. These unequal array Nyquist frequencies create color aliasing in single chip cameras that is wavelength specific. Black-and-white scenes can appear as green, red, or blue imagery.[12] A birefringent crystal, inserted between the lens and the array, changes the effective sampling rate and minimizes color aliasing (discussed in Section 10.3.4., *Optical Anti-Alias Filter*).

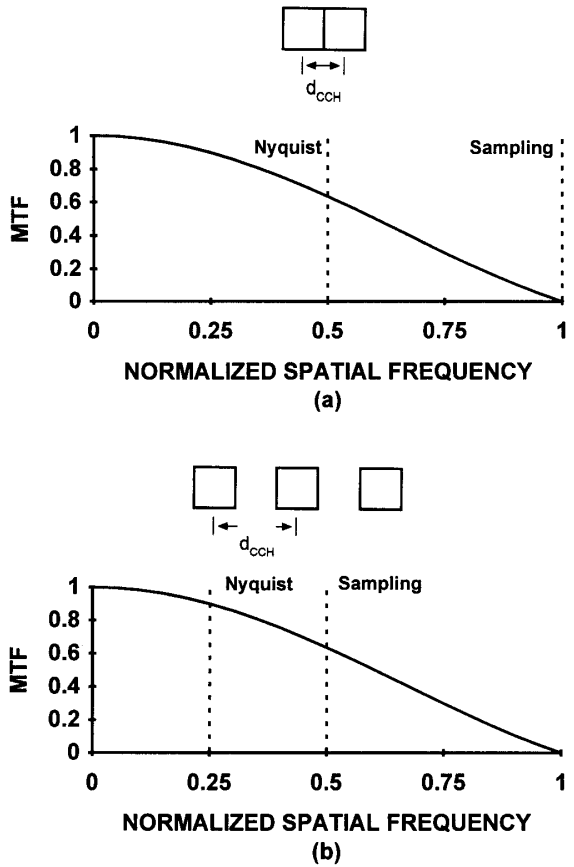


Figure 8-11. Two arrays with different horizontal center-to-center spacings. The detector size, d_H, is the same for both. (a) $d_H/d_{CCH} = 1$. This typifies frame transfer devices. (b) $d_H/d_{CCH} = 0.5$. This is representative of interline transfer CCD arrays. The detector MTF is plotted as a function of u_i/u_{iD}.

Aliasing in the vertical direction can be reduced by summing alternate detector outputs (field integration) in a two-field camera system (see Figure 3-19, page 67). Here, the effective detector height is $2d_V$ and the cutoff is $1/2d_V$ (Figure 8-13). This eliminates aliasing but also reduces the vertical MTF and, therefore, reduces image sharpness in the vertical direction. Horizontal sampling still produces aliasing.

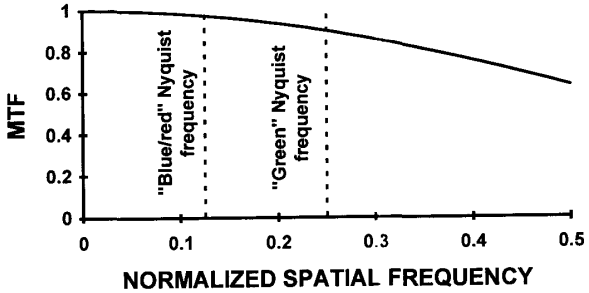


Figure 8-12. Horizontal detector MTF normalized to u_i/u_{iD}. Unequally spaced red, green, and blue sensitive detectors creates different array Nyquist frequencies. This creates different amounts of aliasing and black-and-white scenes may break into color. The spacing between the "red" and "blue" detectors is twice the "green" detector spacing.

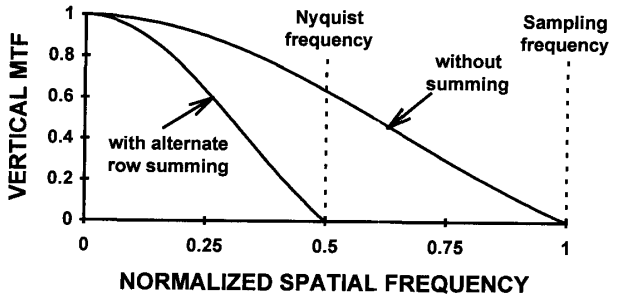


Figure 8-13. Alternate detector summing can reduce aliasing in the vertical direction. Vertical MTF with and without summing normalized to u_i/u_{iD}.

8.3.3. IMAGE DISTORTION

The square wave (bar pattern) is the most popular test target. It is characterized by its fundamental frequency only. The expansion of a square wave into a Fourier series clearly shows that it consists of an infinite number of sinusoidal frequencies. Although the square wave fundamental frequency may be oversampled, the higher harmonics may not. During digitization, the higher order frequencies will be aliased down to lower frequencies and the square wave will change its appearance. There will be intensity variations from bar-to-bar and the bar width will not remain constant.

Sampling replicates frequencies. When two frequencies are close together, they create a beat frequency that is equal to the difference. The fundamental and its replication create a beat frequency of $f_{BEAT} = (f_S - f_o) - f_o$. The beat frequency period lasts for N input frequency cycles:

$$N = \frac{f_o}{2(f_N - f_o)} . \tag{8-3}$$

Figure 8-14 illustrates N as a function of f_o/f_N. As f_o/f_N approaches one, the beat frequency becomes obvious. Lomheim et. al.[13] created these beat frequencies when viewing multiple bar patterns with a CCD camera.

An ideal reconstruction filter will eliminate the beat frequency. However, ideal filters do not exist. Real filters will reduce the MTF and soften edges. Here a square wave may appear as a sinusoid. Thus, there is a tradeoff between edge sharpness and aliasing.

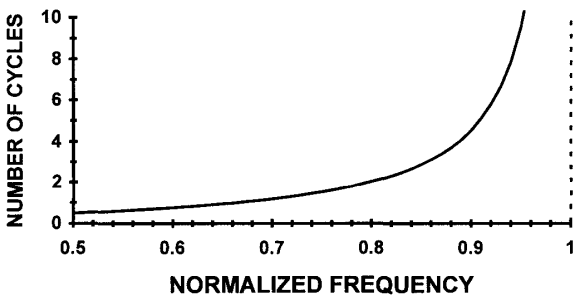


Figure 8-14. Number of input frequency cycles required to see one complete beat frequency cycle as a function of f_o/f_N.

Figures 8-15 through 8-18 illustrate an ideal staring array output when the optical MTF is unity for all spatial frequencies and therefore represents a worst case scenario. That is, all higher harmonics that would have been attenuated by the optical system are available for aliasing by the detector. For these figures, d_H/d_{CCH} is one. The individual detectors sample the square wave in the same manner as illustrated in Figure 8-8b (page 240). Section 10.4., *Optics-Detector Subsystem* describes how the optical MTF modifies the imagery.

If $u_o/u_{iN} = 0.952$, the beat frequency is equal to 9.9 cycles of the input frequency (Figure 8-15). Although a square wave consists of an infinite number of frequencies, it is characterized by its fundamental, u_o. The target must contain at least 10 cycles to see the entire beat pattern. A 4-bar pattern may either be replicated or nearly disappear depending on the phase.

Standard characterization targets, however, consist of several bars. Therefore the beat pattern may not be seen when viewing a 3-bar or 4-bar target. The image of the bar pattern would have to be moved $\pm$½ d (detector extent) to change the output from a maximum value (in-phase) to a minimum value (out-of-phase). This can be proven by selecting just four adjoining bars in Figure 8-15.

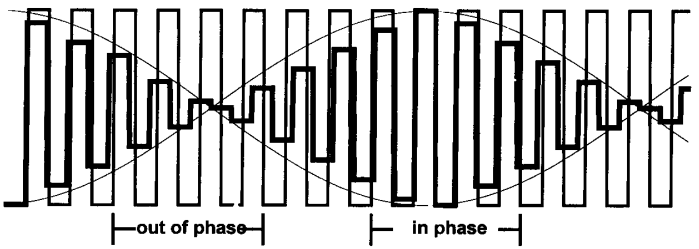


Figure 8-15. Beat frequency produced by an ideal staring system when $u_o/u_{iN} = 0.952$. The value N = 9.9 and $d_H/d_{CCH} = 1$. The light line is the input and the heavy line is the detector output. The beat frequency envelope is shown.

When u_o/u_{iN} is less than about 0.6 (Figure 8-16), the beat frequency is not obvious. Now the output nearly replicates the input but there is some slight variation in pulse width and amplitude. In the region where u_o/u_{iN} is approximately between 0.6 and 0.9, adjacent bar amplitudes are always less than the input amplitude (Figure 8-17).

When u_o/u_{iN} is less than about 0.6, a 4-bar pattern will always be seen (select any four adjoining bars in Figure 8-16). When $u_o/u_{iN} < 0.6$, phasing effects are minimal and a phase adjustment of $\pm$½d in image space will not affect an observer's ability to resolve a 4-bar target. In the region where f_o/f_N is approximately between 0.6 and 0.9, 4-bar targets will never *look* correct (Figure 8-17). One or two bars may be either much wider than the others or one or two bars may be of lower intensity than the others. This phasing effect is a measurement concern and is not normally included in modeling.

Input frequencies of $u_o = u_{iN}/k$, where k is an integer, are faithfully reproduced (i.e., no beat frequencies). When k = 1, as the target moves from in-phase to out-of-phase, the output will vary from a maximum to zero. Selection of u_{iN}/k targets avoids the beat frequency problem but significantly limits the number of spatial frequencies selected.

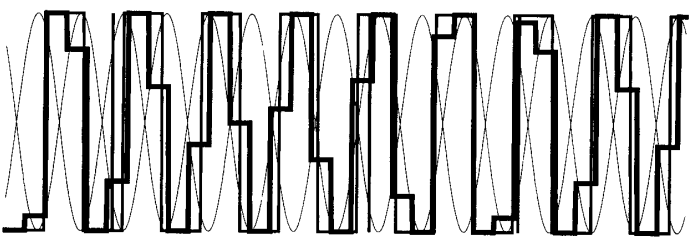


Figure 8-16. Ideal staring system output when $u_o/u_{iN} = 0.522$. The value $N = 0.54$ and $d_H/d_{CCH} = 1$. The light line is the input and the heavy line is the detector output. The output nearly replicates the input when $u_o/u_{iN} < 0.6$.

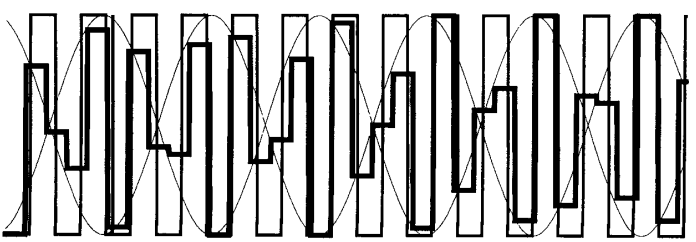


Figure 8-17. Ideal staring system output when $u_o/u_{iN} = 0.811$. The value $N = 2.14$. The light line is the input and the heavy line is the detector output. The output never *looks* quite right when u_o/u_{iN} is between 0.6 and 0.9.

Signals with frequencies above Nyquist frequency will be aliased down to lower frequencies. This would be evident if an infinitely long periodic target was viewed. However, when u_o is less than about $1.15\,u_{iN}$, it is possible to select a phase such that four adjoining bars *appear* to be faithfully reproduced (Figure 8-18). These targets can be *resolved* although the underlying fundamental frequency has been aliased to a lower frequency.

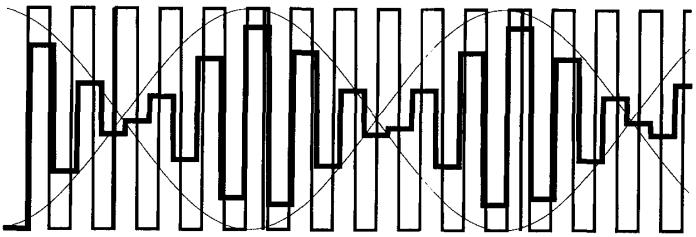


Figure 8-18. Ideal staring system output when $u_o/u_{iN} = 1.094$. The value $N = 5.8$. The light line is the input and the heavy line is the detector output. By selecting the appropriate phase, the output *appears* to replicate an input 4-bar pattern.

Systems can *detect* signals with spatial frequencies above cutoff but cannot faithfully reproduce them. For example, patterns above the Nyquist frequency are aliased to a frequency below Nyquist and a 4-bar pattern may appear as a distorted 3-bar pattern.

The phasing effects shown in Figures 8-15 through 8-18 "disappear" when the target is moving. Here, each frame provides a different image. For each target whose spatial frequency is above Nyquist frequency, one, two, or three bars are visible depending on the relative phase. As the pattern moves, a different set of bars is visible. The eye blends these individual patterns so that you perceive four bars. This prompted some researchers[14] to develop a dynamic test in which the 4-bar target moves.

Example 8-1
SYSTEM CUTOFF

A staring array consists of detectors that are 10 μm × 10 μm in size. The detector pitch is 15 μm. The focal length is 15 cm. The aperture diameter is 3 cm and the average wavelength is 0.5 μm. What is the system cutoff?

System cutoff is the smaller of the optical cutoff, detector cutoff, or Nyquist frequency. In image space, the optical cutoff is $u_{iC} = D_o/(\lambda fl) = 400$ cycles/mm. The detector cutoff is $u_{iC} = 1/d_H = 100$ cycles/mm. The detector pitch provides sampling every 15 μm for an effective sampling rate at 66.7 cycles/mm. Because the Nyquist frequency is one-half the sampling frequency, the system cutoff is 33.3 cycles/mm.

Because object space is related to image space by the lens focal length, the optical cutoff is 30 cycles/mrad, detector cutoff is 15 cycles/mrad, sampling is 10 cycles/mrad, and the Nyquist frequency is 5 cycles/mrad.

Because the MTF remains zero for all frequencies higher than the optical cutoff, any scene frequency above u_{iC} will appear as a uniform blob with no modulation. Spatial frequencies higher than the detector cutoff or the Nyquist frequency (but less than u_{iC}) will be aliased down to lower frequencies.

8.4. ALIASING in FRAME GRABBERS and DISPLAYS

For many applications, the camera output is an analog signal that is then re-sampled by a frame grabber (frame capture device). If the frame grabber's sampling rate is sufficiently high and the analog signal is band-limited, the reconstructed signal will replicate the analog signal in frequency, amplitude, and pulse width.

In principle, the clock rate of an analog-to-digital converter can be set at any rate. However, to conserve on memory requirements and minimize clock rates, some frame grabbers tend to just satisfy the Nyquist frequency of a standard video signal. That is, if the video bandwidth is f_{BW}, then the frame grabber sampling clock operates at $2f_{BW}$ (see Table 5-2, page 153, for f_{BW} values). A frame grabber, with its flash converter, can distort the signal as shown in Figure 8-19. The variation in edge locations affects all image processing algorithms.

The variations in pulse width can only be seen on a monitor because the frame grabber data reside in a computer memory. Note that output pulse widths are equally spaced. They just do not match up with the input signal.

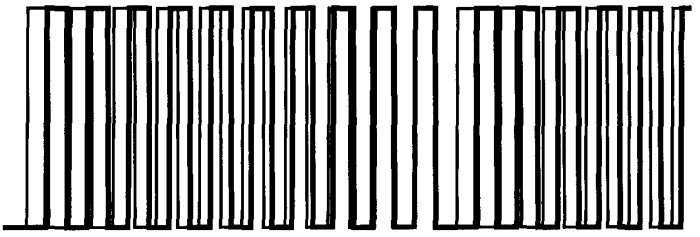


Figure 8-19. Frame grabber output after sample-and-hold circuitry when $u_o/u_{iN} = 0.952$. The light line is the input and the heavy line is the output. Here, u_{iN} is the Nyquist frequency created by the frame grabber clock rate. The output is representative of what is seen on a monitor.

Some frame grabbers have an internal anti-alias filter. This filter ensures that the frame grabber does not produce any additional aliasing. The filter cutoff is linked to the frame grabber clock and is not related to the camera output. Once aliasing has occurred in the camera, the frame grabber anti-alias filter cannot remove it. If the filter does not exist, the frame grabber may create additional aliasing.[15]

Aliasing is rarely discussed with displays, but under certain conditions can occur. Many displays employ digital techniques to match the electronic signal to the CRT characteristics. Ideally, the internal analog-to-digital converter should have a low-pass filter to eliminate possible aliasing. Although good engineering design recommends that the low-pass filter be included, cost considerations may eliminate it. With standard video formats, the bandwidth is limited by the video format and there is no need for the low-pass filter. If a nonstandard bandwidth is used such as that required by a high resolution CCD camera, these monitors may exhibit some aliasing. This becomes noticeable as the array size increases. For example, a high quality monitor can display 1000 $\times$ 1000 pixels. If the array consists of 4000 $\times$ 4000 elements, the display may alias the image as depicted in Figure 8-1 (page 231). There is often a loss of resolution when changing to a lower resolution format.

8.5. REFERENCES

1. S. K. Park and R. A. Schowengerdt, "Image Sampling, Reconstruction and the Effect of Sample-scene Phasing," *Applied Optics*, Vol. 21(17), pp. 3142-3151 (1982).
2. S. E. Reichenbach, S. K. Park, and R. Narayanswamy, "Characterizing Digital Image Acquisition Devices," *Optical Engineering*, Vol. 30(2), pp. 170-177 (1991).
3. W. Wittenstein, J. C. Fontanella, A. R. Newberry, and J. Baars, "The Definition of the OTF and the Measurement of Aliasing for Sampled Imaging Systems," *Optica Acta*, Vol. 29(1), pp. 41-50 (1982).
4. J. C. Felz, "Development of the Modulation Transfer Function and Contrast Transfer Function for Discrete Systems, Particularly Charge Coupled Devices," *Optical Engineering*, Vol. 29(8), pp. 893-904 (1990).
5. S. K. Park, R. A. Schowengerdt, and M. Kaczynski, "Modulation Transfer Function Analysis for Sampled Image Systems," *Applied Optics*, Vol. 23(15), pp. 2572-2582 (1984).
6. L. deLuca and G. Cardone, "Modulation Transfer Function Cascade Model for a Sampled IR Imaging System," *Applied Optics*, Vol. 30(13), pp. 1659-1664 (1991).
7. G. C. Holst, *Sampling, Aliasing, and Data Fidelity*, JCD Publishing, Winter Park, FL (1998).
8. C. E. Shannon, "Communication in the Presence of Noise," *Proceedings of the IRE*, Vol. 37, pp. 10-21 (January 1949).
9. A. B. Jerri, "The Shannon Sampling Theorem - Its Various Extensions and Applications: A Review," *Proceedings of the IEEE*, Vol. 85(11), pp. 1565-1595 (1977).
10. G. C. Holst, *Sampling, Aliasing, and Data Fidelity*, pp. 77-78, JCD Publishing, Winter Park, FL (1998).
11. G. C. Holst, *Sampling, Aliasing, and Data Fidelity*, pp. 293-320, JCD Publishing, Winter Park, FL (1998).
12. J. E. Greivenkamp, "Color Dependent Optical Prefilter for the Suppression of Aliasing Artifacts," *Applied Optics*, Vol. 29(5), pp 676-684 (1990).
13. T. S. Lomheim, L. W. Schumann, R. M. Shima, J. S. Thompson, and W. F. Woodward, "Electro-Optical Hardware Considerations in Measuring the Imaging Capability of Scanned Time-delay-and-integrate Charge-coupled Imagers," *Optical Engineering*, Vol. 29(8), pp. 911-927 (1990).
14. C. M. Webb, "MRTD, How Far Can We Stretch It?" in *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing V*, G. C. Holst, ed., SPIE Proceedings Vol. 1689, pp. 297-307 (1994).
15. G. C. Holst, *Sampling, Aliasing, and Data Fidelity*, pp. 154-156, JCD Publishing, Winter Park, FL (1998).

9

LINEAR SYSTEM THEORY

Linear system theory was developed for electronic circuitry and has been extended to optical, electro-optical, and mechanical systems. It forms an indispensable part of system analysis. For modeling purposes, electronic imaging systems are characterized as linear spatial-temporal systems that are shift-invariant with respect to both time and two spatial dimensions. Although space is three-dimensional, an imaging system displays only two dimensions.

Electrical filters are different from spatial filters in two ways. They are single-sided in time and must satisfy the causality requirement that no change in the output may occur before the application of an input. Optical filters are double-sided in space. Electrical signals may be either positive or negative, whereas optical intensities are always positive. As a result, optical designers and circuit designers often use different terminology.

With a linear-shift-invariant (LSI) system, signal processing is linear. A shift of the input causes a corresponding shift in the output. For electronic circuits, the shift is in time, whereas for optical systems, the shift is in space. In additional, the input-to-output mapping is single-valued. For optical systems, one addition condition exists: the radiation must be incoherent if the system is to be described in terms of radiance.

A sampled-data system may be considered "globally" shift invariant on a macro-scale. For example, with an electronic imaging system, as a target moves from the top of the field-of-view to the bottom, the image also moves from the top to the bottom. On a micro-scale, moving a point source across a single detector does not change the detector output. That is, the system is not shift-invariant on a micro-scale. Similarly, with electronic circuits, large time shifts of the input produce corresponding large time shifts in output. A shift during a sample time, does not change the output. Here also, the system is not shift-invariant on a micro-scale.

Single-valued mapping only occurs with non-noisy and non-quantized systems. No system is truly noiseless but can be approximated as one when the signal has sufficient amplitude. With digital systems, the output is quantized. The smallest output is the least significant bit and the analog-to-digital converter generally limits the largest signal. If the signal level is large compared to the

least significant bit, then the system can be treated as quasi-linear over a restricted region.

In spite of the disclaimers mentioned, systems are treated as quasi-linear and quasi-shift-invariant over a restricted operating region to take advantage of the wealth of mathematical tools available. An LSI system merely modifies the amplitude and phase of the input. These are specified by the modulation transfer function and phase transfer function.

The symbols used in this book are summarized in the *Symbol List* (page xviii) which appears after the *Table of Contents*.

9.1. LINEAR SYSTEM THEORY

A linear system provides single mapping from an input to an output. Superposition allows the addition of individual inputs to create an unique output. Let h{ } be a linear operator that maps one function, f(t), into another function, g(t):

$$h\{f(t)\} = g(t) \ . \tag{9-1}$$

Let the response to two inputs, $f_1(t)$ and $f_2(t)$, be $g_1(t)$ and $g_2(t)$:

$$h\{f_1(t)\} = g_1(t) \quad and \quad h\{f_2(t)\} = g_2(t) \ . \tag{9-2}$$

For a linear system, the response to a sum of inputs is equal to the sum of responses to each input acting separately. For any arbitrary scale factors, the superposition principle states

$$h\{a_1 f_1(t) + a_2 f_2(t)\} = a_1 g_1(t) + a_2 g_2(t) \ . \tag{9-3}$$

When passing the signal through another linear system, the new operator provides:

$$h_2\{g(t)\} = h_2\{h_1\{f(t)\}\} \ . \tag{9-4}$$

9.1.1. TIME-VARYING SIGNALS

An electrical signal can be thought of as the sum of an infinite array of impulses (Dirac delta functions) located inside the signal boundaries. Thus, the signal can be decomposed into a series of weighted Dirac delta functions (Figure 9-1):

$$v_{IN}(t) = \sum_{t'=-\infty}^{\infty} v_{IN}(t')\delta(t - t')\,\Delta t' \,, \tag{9-5}$$

where $\delta(t - t')$ is the Dirac delta or impulse function at $t = t'$ and $v_{IN}(t')$ is the signal evaluated at t'.

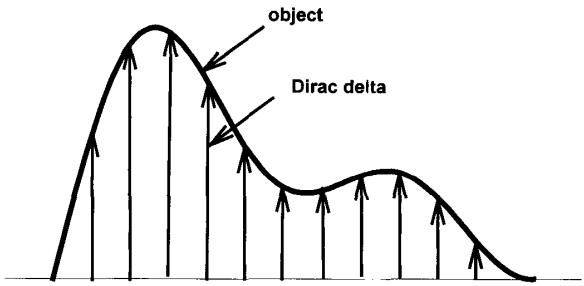


Figure 9-1. A time-varying signal can be decomposed into a series of closely spaced impulses with amplitude equal to the signal at that value. They are shown widely separated for clarity and should not be confused with sampling.

If $v_{IN}(t)$ is applied to a linear circuit, using $h\{\ \}$ will provide the output

$$v_{OUT}(t) = \sum_{t'=-\infty}^{\infty} h\{v_{IN}(t')\delta(t - t')\,\Delta t'\} \quad . \tag{9-6}$$

Comparing Equation 9-5 to Equation 9-3, for each t', $v_{IN}(t')$ replaced a_i and $\delta(t - t')$ replaced $f_i(t)$. That is, the input has been separated into a series of functions $a_1f_1(t) + a_2f_2(t) + \cdots$.

As $\Delta t' \rightarrow 0$, this becomes the convolution integral

$$v_{OUT}(t) = \int_{-\infty}^{\infty} v_{IN}(t')\, h\{\delta(t - t')\}\, dt' \tag{9-7}$$

and is symbolically represented by

$$v_{OUT}(t) = v_{IN}(t) * h(t) , \tag{9-8}$$

where $*$ indicates the convolution operator (other texts may use $\star$ or $\otimes$). The system impulse response is $h\{\delta(t)\}$. In Figure 9-2, each impulse response is given for each t'. The individual impulse responses are added to produce the output. This addition is the definition of superposition.

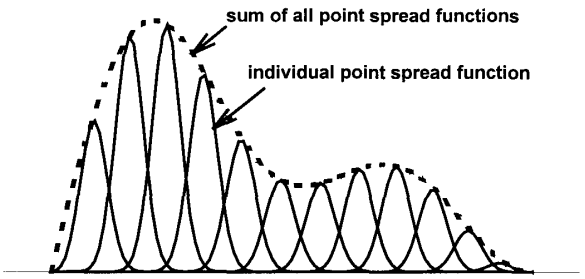


Figure 9-2. The linear operator, h{ }, transforms each input Dirac delta into an output impulse response. The sum of the impulse responses creates the output. Although shown separated for clarity, the impulse responses have infinitesimal separations.

The convolution theorem of Fourier analysis states that the Fourier transform of the convolution of two functions equals the product of the transform of the two functions. That is, multiple convolutions are equivalent to multiplication in the frequency domain. Using frequency domain analysis is convenient because multiplications are easier to perform and easier to visualize than convolutions. Then

$$V_{OUT}(f) = V_{IN}(f)\, H(f) . \tag{9-9}$$

The variable f denotes electrical frequency (Hz). Upper case variables represent the Fourier transform of the signal (lower case variables). Let the input be an impulse. The Fourier transform of an impulse response is

$$V_{IN}(f) = 1 \ . \tag{9-10}$$

That is, all frequencies are present. Because the input to the system contains all frequencies, the output must be a function of the system only. Any variation in phase or amplitude must be due to the system itself. The output is

$$V_{OUT}(f) = H(f) \ . \tag{9-11}$$

9.1.2. SPATIALLY VARYING SIGNALS

The one-dimensional electronic signal decomposition can be easily extended to two-dimensional imagery. An object can be thought of as the sum of an infinite array of impulses located inside the target boundaries. Thus, an object can be decomposed into a two-dimensional array of weighted Dirac delta functions, $\delta(x - x')$, $\delta(y - y')$:

$$o(x,y) = \sum_{x'=-\infty}^{\infty} \sum_{y'=-\infty}^{\infty} o(x',y')\delta(x - x')\,\delta(y - y')\,\Delta x'\,\Delta y' \ . \tag{9-12}$$

An optical system with operator h{ } produces an image

$$i(x,y) = \sum_{x'=-\infty}^{\infty} \sum_{y'=-\infty}^{\infty} h\{o(x',y')\delta(x - x')\,\delta(y - y')\,\Delta x'\,\Delta y'\} \ . \tag{9-13}$$

For small increments, this becomes the convolution integral

$$i(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} o(x',y')\, h\{\delta(x - x')\,\delta(y - y')\}\, dx'\, dy' . \tag{9-14}$$

It is symbolically represented by

$$i(x,y) = o(x,y) ** h(x,y) , \tag{9-15}$$

where $**$ represents the two-dimensional convolution. The function h(x,y) is the optical system's response to an input impulse. The resulting image is the point spread function (PSF). For an ideal circular aperture, the central portion of the PSF is the Airy disk from which the Rayleigh criterion and other resolution metrics are derived. In one dimension h(x) is the line spread function, LSF. The LSF is the resultant image produced by the imaging system when viewing an ideal line. There is no equivalent interpretation for h(t) for electrical circuits. The function h(t) is simply the impulse response.

The transform of the convolution becomes a multiplication in the frequency domain

$$I(u,v) = O(u,v)\, H(u,v) . \tag{9-16}$$

If all frequencies are present in the image (i.e., an impulse), O(u,v) = 1, then I(u,v) = H(u,v). The relationship between spatial frequency (u,v) and electrical frequency (f) is further discussed in Section 10.1., *Frequency Domains*.

9.2. ELECTRONIC IMAGING SYSTEM

The electronic imaging system response consists of both the optical response and the electronic response: h(x,y,t). Time and spatial coordinates are treated separately. For example, optical elements do not generally change with time and therefore are characterized only by spatial coordinates. Similarly, electronic circuitry exhibits only temporal responses. The detector provides the interface between the spatial and temporal components, and its response depends on both temporal and spatial quantities. The conversion of two-dimensional optical information to a one-dimensional electrical response assumes a linear photodetection process. Implicit in the detector response is the conversion from input photon flux to output voltage.

The optical transfer function (OTF) plays a key role in the theoretical evaluation and optimization of an optical system. The modulation transfer function (MTF) is the magnitude and the phase transfer function (PTF) is the phase of the complex-valued OTF. In many applications, the OTF is real-valued and positive so that the OTF and MTF are equal. When focus errors or aberrations are present, the OTF may become negative or even complex valued.

Electronic circuitry also can be described by an MTF and PTF. The combination of the optical MTF and the electronic MTF creates the electronic imaging system MTF. The MTF is the primary parameter used for system design, analysis, and specifications. When coupled with the three-dimensional noise parameters, the MTF and PTF uniquely define system performance (discussed in Chapter 12, *Minimum Resolvable Contrast*).

Symbolically

$$OTF(u,v) = H_{SPATIAL}(u,v) = MTF(u,v)\, e^{jPTF(u,v)} \tag{9-17}$$

and

$$H_{ELECTRONICS}(f) = MTF(f)\, e^{jPTF(f)} . \tag{9-18}$$

With appropriate scaling, the electronic frequencies can be converted into spatial frequencies. This is symbolically represented by $f \rightarrow u$. The electronic circuitry is assumed to modify the horizontal signal only (although this depends on the system design). The combination of spatial and electronic responses is sometimes called the system OTF:

$$OTF_{SYSTEM}(u,v) = MTF(u,v)\, MTF(f \rightarrow u) e^{j[PTF(u,v) + PTF(f \rightarrow u)]} . \tag{9-19}$$

For mathematical convenience, the horizontal and vertical responses are considered separable (usually coincident with the detector array axes). The electrical response is considered to affect only the horizontal response

$$H_{SYSTEM}(u) = H_{SPATIAL}(u) H_{ELECTRONICS}(f \rightarrow u) \tag{9-20}$$

and

$$H_{SYSTEM}(v) = H_{SPATIAL}(v) . \tag{9-21}$$

A system is composed of many components that respond to either spatial or temporal signals. For independent MTFs,

$$MTF_{SYSTEM}(u,v) = \prod_{i=1}^{N} \prod_{j=1}^{M} MTF_i(u,v) MTF_j(f \rightarrow u) . \tag{9-22}$$

While individual lens elements each have their own MTF, the MTF of the lens system is not usually the product of the individual MTFs. This occurs because one lens may minimize the aberrations created by another.

9.3. MTF and PTF INTERPRETATION

The system MTF and PTF alter the image as it passes through the circuitry. For linear-shift-invariant systems, the PTF simply indicates a spatial or temporal shift with respect to an arbitrarily selected origin. An image where the MTF is drastically altered is still recognizable, whereas large PTF nonlinearities can destroy recognizability. Modest PTF nonlinearity may not be noticed visually except those applications where target geometric properties must be preserved (i.e., mapping or photogrammetry). Generally, PTF nonlinearity increases as the spatial frequency increases. Because the MTF is small at high frequencies, the nonlinear-phase-shift effect is diminished.

The MTF is the ratio of output modulation to input modulation normalized to unity at zero frequency. While the modulation changes with system gain, the MTF does not. The input can be as small as desired (assuming a noiseless system with high gain) or it can be as large as desired because the system is assumed not to saturate.

Modulation is the variation of a sinusoidal signal about its average value (Figure 9-3). It can be considered as the AC amplitude divided by the DC level. The modulation is:

$$MODULATION = M = \frac{V_{MAX} - V_{MIN}}{V_{MAX} + V_{MIN}} = \frac{AC}{DC}, \qquad (9\text{-}23)$$

where V_{MAX} and V_{MIN} are the maximum and minimum signal levels, respectively. The modulation transfer function is the output modulation produced by the system divided by the input modulation as a function of frequency:

$$MTF(f) = \frac{M_{OUTPUT}(f)}{M_{INPUT}(f)}. \qquad (9\text{-}24)$$

The concept is presented in Figure 9-4. Three input and output signals are plotted in Figures 9-4a and 9-4b, respectively, and the resultant MTF is shown in Figure 9-4c. As a ratio, the MTF is a relative measure with values ranging from zero to one.

The MTF is a measure of how well a system will faithfully reproduce the input. The highest frequency that can be faithfully reproduced is the system cutoff frequency. If the input frequency is above the cutoff, the output will be proportional to the signal's average value with no modulation.

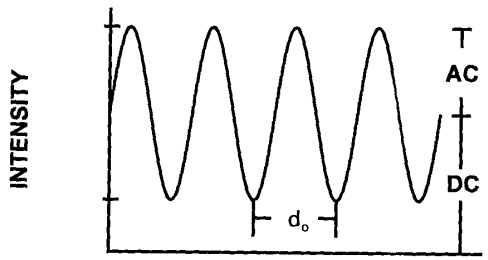


Figure 9-3. Definition of target modulation. The value d_o is the extent of one cycle.

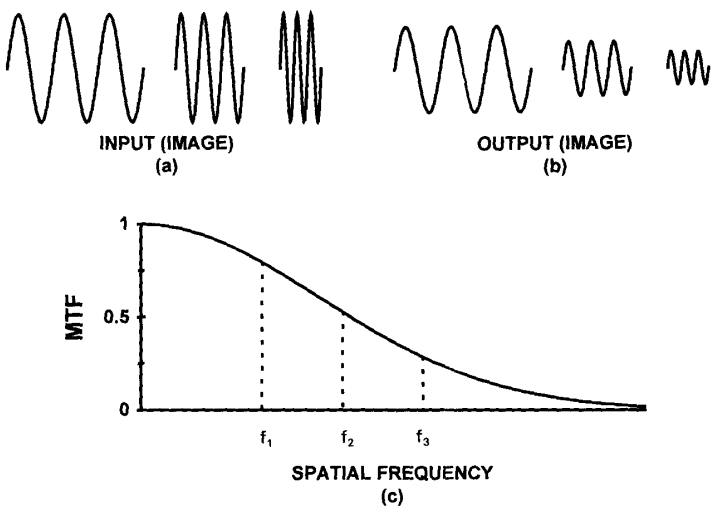


Figure 9-4. Modulation transfer function. (a) Input signal for three different spatial frequencies, (b) output for the three frequencies, and (c) the MTF is the ratio of output-to-input modulation.

9.4. SUPERPOSITION APPLIED to OPTICAL SYSTEMS

If the system MTF is known, the image for any arbitrary object can be computed. First, the object is dissected into its constituent spatial frequencies (i.e., the Fourier transform of the object is obtained). Next, each of these frequencies is multiplied by the system MTF at that frequency. Then, the inverse Fourier transform provides the image.

To illustrate the superposition principle and MTF approach, we will show how an ideal optical system modifies an image. An ideal optical system is, by definition, a linear-phase-shift system. The most popular test target consists of a series of bars: typically three or four bars although more may be used. For illustrative purposes, the periodic bars are assumed of infinite extent. A one-dimensional square wave, when expanded into a Fourier series about the origin, contains only odd harmonics:

$$o(x) = \frac{1}{2} + \frac{2}{\pi}\sum \frac{1}{n}\sin\left(\frac{2\pi nx}{d_o}\right) \quad n = 1, 3, 5, \cdots, \tag{9-25}$$

where d_o is the period of the square wave. The fundamental frequency u_o is $1/d_o$. Note that the peak-to-peak amplitude of the fundamental is $4/\pi$ times the square wave amplitude. In the frequency domain, the square wave provides discrete spatial frequencies of $u_n = 1/d_o, 3/d_o, 5/d_o, \cdots$, with amplitudes $2/\pi$, $2/3\pi$, $2/5\pi$, $\cdots$, respectively.

Let a circular optical system image the square wave. The MTF for a circular, clear aperture, diffraction-limited lens is

$$MTF_{OPTICS}(u_i) = \frac{2}{\pi}\left[\cos^{-1}\left(\frac{u_i}{u_{iC}}\right) - \frac{u_i}{u_{iC}}\sqrt{1-\left(\frac{u_i}{u_{iC}}\right)^2}\right] \quad \textit{when } u_i < u_{iC} \tag{9-26}$$

$$= 0 \quad \textit{elsewhere} .$$

The optical cutoff in image space is $u_{iC} = D_o/(\lambda\, fl)$, where D_o is the aperture diameter, λ is the average wavelength, and fl is the focal length.

By superposition, the diffraction-limited optical MTF and square wave Fourier series amplitudes are multiplied together at each spatial frequency:

$$I(u_n) = MTF_{OPTICS}(u_n)\, O(u_n) \; . \tag{9-27}$$

Taking the inverse Fourier transform provides the resultant image. Equivalently,

$$i(x) = \frac{1}{2} + MTF_{OPTICS}(u_o)\left[\frac{2}{\pi}\sin(2\pi u_o x)\right] + MTF_{OPTICS}(3u_o)\left[\frac{2}{3\pi}\sin(6\pi u_o x)\right] + \cdots \; . \tag{9-28}$$

If u_o is greater than $u_{iC}/3$, only the fundamental of the square wave will be faithfully reproduced by the optical system. Here, the square wave will appear as a sine wave. Note that the optical MTF will reduce the image amplitude. As u_o decreases, the image will look more like a square wave (Figure 9-5). Because the optics does not pass any frequencies above u_{iC}, the resultant wave form is a truncation of the original series modified by the optical MTF. This results in some slight ringing. This ringing is a residual effect of the Gibbs phenomenon. Note that with imaging systems, the intensity is always positive. As the input frequency approaches u_{iC} (Figure 9-5d), the modulation is barely present in the image. However, the average scene intensity approaches 0.5. That is, the scene detail can no longer be perceived at or above u_{iC}. *The scene did not disappear.*

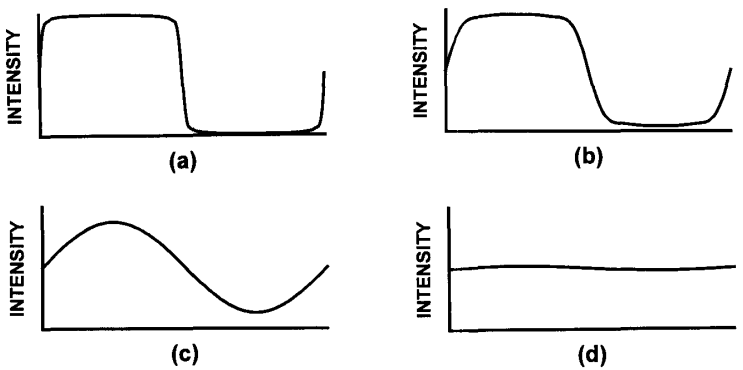


Figure 9-5. A square wave modified by a circular aperture as a function of x/d_o. As the square-wave fundamental frequency increases, the edges become rounded. It will appear as a sinusoid when $u_o \geq u_{iC}/3$. (a) $u_o = u_{iC}/40$, (b) $u_o = u_{iC}/10$, (c) $u_o = u_{iC}/3$, and (d) $u_o = 9u_{iC}/10$.

9.5. CONTRAST TRANSFER FUNCTION

The system response to a square-wave target is the contrast transfer function (CTF) or square-wave response (SWR). The CTF is a convenient measure because of the availability of square-wave targets. It is not a transfer function in the same sense that the MTF is and, as such, subsystem CTFs cannot be cascaded. Figure 9-6 illustrates the relationship between square-wave and sinusoidal amplitudes. The CTF is typically higher than the MTF at all spatial frequencies.

A square wave can be expressed either as an infinite sine (Equation 9-25) or cosine series. The only difference is the arbitrarily selected origin. Coltman[1] selected the cosine expansion. The output amplitude of the square wave, whose fundamental frequency is f_o, is an infinite sum of the input cosine amplitudes modified by the system's MTF:

$$CTF(u_o) = \frac{4}{\pi}\left| MTF(u_o) - \frac{MTF(3u_o)}{3} + \frac{MTF(5u_o)}{5} - \ldots \right| \quad (9\text{-}29)$$

or

$$CTF(u_o) = \frac{4}{\pi}\left| \sum_{k=0}^{\infty} (-1)^k \frac{MTF[(2k+1)u_o]}{2k+1} \right| . \quad (9\text{-}30)$$

By adding and subtracting terms, the MTF at frequency u_o can be expressed as an infinite sum of CTFs.

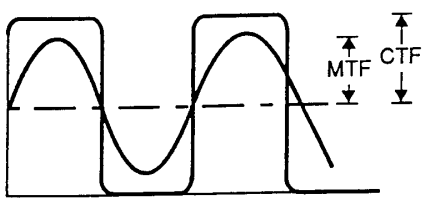


Figure 9-6. AC components of the CTF and MTF. The CTF is usually equal to or greater than the MTF.

Again, assuming a linear-shift-invariant system, the MTF is[1,2]:

$$MTF(u_o) = \frac{\pi}{4}\left|CTF(u_o) + \frac{CTF(3u_o)}{3} - \frac{CTF(5u_o)}{5} + \frac{CTF(7u_o)}{7} + \frac{CTF(11u_o)}{11} + \textit{irregular terms}\right| \tag{9-31}$$

or

$$MTF(u_o) = \frac{\pi}{4}\left|\sum_{k=0}^{\infty} B_k \frac{CTF(ku_o)}{k}\right|, \tag{9-32}$$

where k takes on odd values only: 1, 3, 5, ... and B_k is -1 or 1 according to:

$$B_k = (-1)^m(-1)^{\frac{k-1}{2}} \qquad \textit{for } r = m\ . \tag{9-33}$$

The value r is the number of different prime factors in k. The value m is the total number of primes into which k can be factored. The value $B_k = 0$ for $r < m$.

Theoretically, to obtain the MTF at frequency u_o, an infinite number of square-wave responses must be measured. However, the number required is limited by the spatial cutoff frequency, $u_{MTF=0}$, where the MTF approaches zero and remains zero thereafter. For bar targets with spatial frequency above $\frac{1}{3}u_{MTF=0}$, the MTF is equal to $\pi/4$ times the measured CTF. That is, above $\frac{1}{3}$ $u_{MTF=0}$, only one target is necessary to compute[3] the MTF.

For low spatial frequencies, the output nearly replicates the input (Figure 9-7a). For frequencies near $\frac{1}{3}$ cutoff, the output starts to appear sinusoidal (Figure 9-7b). Above $\frac{1}{3}$ cutoff (Figure 9-7c), the MTF amplitude is $\pi/4$ times the CTF amplitude.

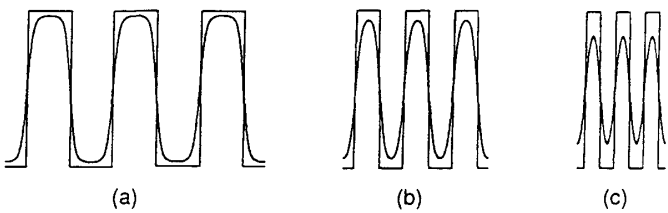


Figure 9-7. Input and output wave forms for a linear system. (a) Very low frequency signals are faithfully reproduced, (b) mid-spatial frequencies tend to look like sinusoids, and (c) input square waves with spatial frequencies above ⅓ $u_{MTF=0}$ appear as sinusoidal outputs.

The relationship between the CTF and MTF was developed for analog optical systems and is appropriate for wet-film photographic cameras and image intensifiers. It can also be applied in the scan direction for analog scanning devices (vidicons and displays).

For sampled data systems, where aliasing can occur, the relationship should be used cautiously. CTF measurements are appropriate tests for system performance verification. Conversion to the MTF that can have significant mathematical complexity when aliasing is present.

Sampling is present in all solid state imaging systems due to the discrete location of the detectors. Phasing effects between the detector and the location of the target introduce problems at nearly all spatial frequencies. The interaction between the bar frequency and the system sampling frequency produces sum and difference frequencies that produce variations in amplitude and bar width (see Section 8.3.3., *Image Distortion*, page 245).

The CTF is well behaved (in the sense that the bar width remains constant) when the bar spatial frequency is proportional to the system Nyquist frequency,

$$u_o = \frac{u_{iN}}{k}, \tag{9-43}$$

where k is an integer. These spatial frequencies provide the in-phase CTF. The CTF can be used to assess system performance at the spatial frequencies selected. It is usually specified only at the Nyquist frequency (k = 1).

9.6. REFERENCES

1. J. W. Coltman, "The Specification of Imaging Properties by Response to a Sine Wave Input," *Journal of the Optical Society of America*, Vol. 44(6), pp. 468-471 (1954).
2. I. Limansky, "A New Resolution Chart for Imaging Systems," *The Electronic Engineer*, Vol. 27(6), pp. 50-55 (1968).
3. G. C. Holst, *Testing and Evaluation of Infrared Imaging Systems*, pp. 238-246, JCD Publishing, Winter Park, FL (1993).

10

SYSTEM MTF

The equations describing array (Chapter 4) and camera (Chapter 6) performance assume that the image is very large and it illuminates many detector elements. That is, the object contained only very low spatial frequencies and that the system MTF is essentially unity over these spatial frequencies.

The optics and detector MTFs modify the signal spatial frequencies. These frequencies may also be affected by electronic filter MTFs although these electronic filter functions are typically designed to pass the entire image frequency spectrum without attenuation.

But noise originates at the detector and within the electronic circuitry. It is modified by the electronic MTFs only. Detector array noise is fixed but analog amplifier noise increases as the electronic bandwidth increases. The camera SNR is maximized when the electronic bandwidth passes the signal with minimal degradation. Wider bandwidth electronics will increase the noise but not the signal and thereby reduce the SNR.

For mathematical convenience, a specific target feature can be associated with a fictitious spatial frequency (Figure 10-1). The SNR for this feature is

$$SNR(u_o) = MTF_{SYS}(u_o)\frac{n_{pe}}{\langle n_{SYS}\rangle}, \qquad (10\text{-}1)$$

where $\langle n_{SYS}\rangle$ includes shot, reset, amplifier, quantization, and fixed pattern noise sources (see Section 4.2., *Noise*, page 123). The total amplifier noise depends upon the electronic bandwidth and that depends on the electronic frequency response associated with the readout signal chain.

A solid state camera consists of optics, an array of pixels, and a signal chain with electronic frequency response characteristics. The complete imaging system includes the display:

$$MTF_{SYS} = MTF_{OPTICS}MTF_{DETECTOR}MTF_{SIGNAL\text{-}CHAIN}MTF_{DISPLAY}. \qquad (10\text{-}2)$$

The observer processes this information to create a perceived MTF:

$$MTF_{PERCEIVED} = MTF_{SYS}\, MTF_{HVS} \,. \qquad (10\text{-}3)$$

MTF_{HVS} is the human visual system MTF which includes both the eye optical MTF and neural processing. When $MTF_{PERCEIVED}$ is coupled to system noise, system responsivity, and the eye integration capabilities, it is possible to predict the minimum resolvable contrast (described in Chapter 12, *Minimum Resolvable Contrast*). MRC is used by the military to predict target detection and recognition ranges for a given probability of success.

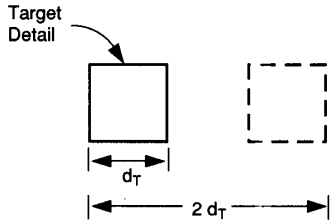


Figure 10-1. Fictitious spatial frequency associated with a target feature. If the image detail size is d_T mm at the detector plane, then the fictitious horizontal frequency is $u_o = 1/(2d_T)$ cycles/mm.

Because solid state arrays are undersampled, the highest frequency that can be reproduced is the array Nyquist frequency, u_{iN}. As a result, it is common practice to evaluate Equation 10-1 at $u_o = u_{iN}$. This dictates the smallest target feature that can be faithfully reproduced. The value u_{iN} should be added to all MTF curves (see Section 8.3.2., *Detector Array Nyquist Frequency*, page 242).

The MTFs presented in this chapter apply to all staring arrays. Only the optics and detector array are sensitive to the scene spectral and spatial content. The remaining subsystem MTF components only modify the electrical signals. These signal chain MTFs apply to all imaging systems independent of the system spectral response.

Optical spatial frequency is two-dimensional with the frequency ranging from $-\infty$ to $+\infty$. Although not truly separable, for convenience, the horizontal MTF, MTF(u), and vertical MTF, MTF(v), are analyzed separately. Electrical filters are causal, one-dimensional, and are considered to operate in the direction of the readout clocking only. Thus, the horizontal and vertical MTFs may be different.

The symbols used in this book are summarized in the *Symbol List* (page xviii) which appears after the *Table of Contents*.

10.1. FREQUENCY DOMAINS

There are four different locations where spatial frequency domain analysis is appropriate. They are object space (before the camera optics), image space (after the camera optics), display, and observer (at the eye) spatial frequencies. Simple equations relate the spatial frequencies in all these domains.

Figure 10-2 illustrates the spatial frequency associated with a bar target. Bar patterns are the most common test targets and are characterized by their fundamental spatial frequency. Using the small angle approximation, the angle subtended by one cycle (one bar and one space) is d_o/R_1, where d_o is the spatial extent of one cycle and R_1 is the distance from the imaging system entrance aperture to the target. When using a collimator to project the targets at apparently long ranges, the collimator focal length replaces R_1. Targets placed in the collimator's focal plane can be described in object space.

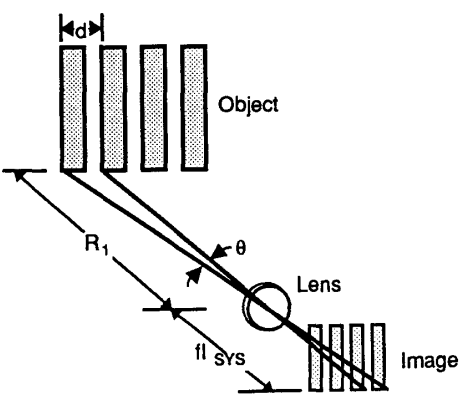


Figure 10-2. Correspondence of spatial frequencies in object and image space. Although the MTF is defined for sinusoidal signals, the bar target is the most popular test target (shown here). Bar targets are characterized by their fundamental frequency.

The horizontal object-space spatial frequency, u_{ob}, is the inverse of the horizontal target angular subtense and is usually expressed in cycles/mrad:

$$u_{ob} = \frac{1}{1000}\left(\frac{R_1}{d_o}\right) \quad \frac{cycles}{mrad} . \tag{10-4}$$

A similar equation exists for the vertical object-space spatial frequency. The object-space domain is used by the military for describing system performance.

Optical designers typically quote spatial frequencies in image space to specify the resolving capability of lens systems. Photographic cameras and CCD cameras are typically specified with spatial frequencies in image space which is the object-space spatial frequency divided by the system focal length:

$$u_i = \frac{u_{ob}}{fl} \quad \frac{line\text{-}pairs}{mm} \quad or \quad \frac{cycles}{mm} , \tag{10-5}$$

where u_i is the inverse of one cycle in the focal plane of the lens system. Although used interchangeably, line-pairs suggest square wave targets and cycles suggest sinusoidal targets. To maintain dimensionality, if u_{ob} is measured in cycles/mrad then the focal length must be measured in meters to obtain cycles/mm.

Analog electronic filters are one-dimensional and modify a serial data stream. The electrical frequency is related to the horizontal field-of-view (HFOV) and the time it takes to read out one line. For analog filters immediately following the detector array, where the readout time is $t_{H\text{-}LINE}$,

$$f_e = \frac{HFOV \cdot fl}{t_{H-LINE}} u_i \quad Hz . \tag{10-6}$$

For staring arrays with N_H horizontal detectors,

$$f_e = \frac{(N_H - 1)d_{CCH} + d_H}{t_{H-LINE}} u_i \quad Hz . \tag{10-7}$$

While digital filters operate on a data array, the effective sampling rate is related to the timing of the readout clock:

$$f_{eS} = \frac{1}{T_{CLOCK}} \quad Hz \ . \tag{10-8}$$

After the digital-to-analog converter, the serial stream data rate and therefore the filter design is linked to the video standard. The active line time, $t_{VIDEO\text{-}LINE}$, (Table 10-1) and array size provide the link between image space and video frequencies:

$$f_v = \frac{HFOV \cdot fl}{t_{VIDEO\text{-}LINE}} u_i \quad Hz \ . \tag{10-9}$$

Video standards were discussed in Section 5.2.1., *Video Timing* (page 151). For staring arrays,

$$f_v = \frac{(N_H - 1) d_{CCH} + d_H}{t_{VIDEO\text{-}LINE}} u_i \quad Hz \ . \tag{10-10}$$

Table 10-1
STANDARD VIDEO TIMING

FORMAT	TOTAL LINE TIME (μs)	MINIMUM ACTIVE LINE TIME (μs) $t_{VIDEO\text{-}LINE}$
EIA 170	63.492	52.092
NTSC	63.555	52.456
PAL	64.0	51.7
SECAM	64.0	51.7

The field-of-view and monitor size link image spatial frequency to the horizontal and vertical display spatial frequencies:

$$u_d = \frac{HFOV \cdot fl}{W_{MONITOR}} u_i \quad \frac{cycles}{mm} \tag{10-11}$$

and

$$v_d = \frac{VFOV \cdot fl}{H_{MONITOR}} v_i \quad \frac{cycles}{mm} . \tag{10-12}$$

Both the monitor width, $W_{MONITOR}$, and height, $H_{MONITOR}$, are usually measured in millimeters. When the monitor aspect ratio is the same as the FOV ratio (the usual case),

$$\frac{HFOV}{W_{MONITOR}} = \frac{VFOV}{H_{MONITOR}} . \tag{10-13}$$

The spatial frequency presented to the observer depends on the observer's viewing distance and the image size on the display (Figure 10-3). Observer response is assumed to be rotationally symmetric so that the horizontal and vertical responses are the same. The usual units[1] are cycles/deg:

$$u_{eye} = \frac{1}{2 \tan^{-1}\left(\frac{1}{2} \frac{1}{D u_d}\right)} \quad \frac{cycles}{deg} . \tag{10-14}$$

Here, the arc tangent is expressed in degrees. Because image detail is important, the small angle approximation provides

$$u_{eye} = 0.01745 D u_d \quad \frac{cycles}{deg} . \tag{10-15}$$

Table 10-2 summarizes the various spatial frequencies.

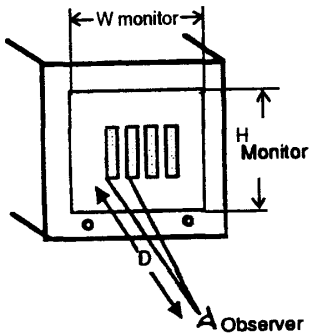


Figure 10-3. Correspondence of spatial frequencies for the display and for the eye.

Table 10-2
TRANSFORM PAIR UNITS

SUBSYSTEM RESPONSE	SIGNAL VARIABLES	FREQUENCY VARIABLES
Optics, detector	x denotes distance (mm) y denotes distance (mm)	u_i expressed as cycles/mm v_i expressed as cycles/mm
Optics, detector	θ_x denotes angle (mrad) θ_y denotes angle (mrad)	u_{ob} expressed as cycles/mrad v_{ob} expressed as cycles/mrad
Analog electronics	t denotes time (sec)	f_e expressed as cycles/s (Hz)
Video electronics	t denotes time (sec)	f_v expressed as cycles/s (Hz)
Display	x denotes distance (mm) y denotes distance (mm)	u_d expressed as cycles/mm v_d expressed as cycles/mm
Observer	θ_x denotes angle (deg) θ_y denotes angle (deg)	u_{eye} expressed as cycles/deg v_{eye} expressed as cycles/deg

10.2. OPTICS MTF

Optical systems consist of several lenses or mirror elements with varying focal lengths and varying indices of refraction. Multiple elements are used to minimize lens aberrations. While individual element MTFs are used for design and fabrication, these individual MTFs cannot be cascaded to obtain the optical system MTF. For modeling purposes, the optical system is treated as a single lens that has the same effective focal length and aberrations as the lens system.

The optical anti-aliasing filter, which is used to blur the image on a single chip color filter array, is also part of the lens system. Its performance is best described by the way it affects the detector MTF. Therefore, its description is included in the detector section.

Optical spatial frequency is two-dimensional with the frequency ranging from $-\infty$ to $+\infty$. The diffraction-limited MTF for a circular aperture was first introduced in Section 9.4., *Superposition Applied to Optical Systems* (page 261). For an aberration free and radially symmetric optical system, MTF_{OPTICS} is the same in the horizontal and vertical directions. In the horizontal direction,

$$MTF_{OPTICS}(u_i) = \frac{2}{\pi}\left[\cos^{-1}\left(\frac{u_i}{u_{iC}}\right) - \frac{u_i}{u_{iC}}\sqrt{1-\left(\frac{u_i}{u_{iC}}\right)^2}\right] \quad when\ u_i < u_{iC} \qquad (10\text{-}16)$$

$$= 0 \quad elsewhere\ .$$

The image-space optics cutoff is $u_{iC} = D_O/(\lambda\ fl) = 1/(F\lambda)$, where D_O is the aperture diameter and fl is the effective focal length. The f-number, F, is equal to fl/D_O.

Figure 10-4 illustrates MTF_{OPTICS} as a function of u_i/u_{iC}. Because the cutoff frequency is wavelength dependent, Equation 10-16 and Figure 10-4 are only valid for noncoherent monochromatic light. The extension to polychromatic light is lens specific. Most lens systems are color corrected (achromatized) and therefore there is no simple way to apply this simple formula to predict the MTF. As an approximation to the polychromatic MTF, the average wavelength is used to calculate the cutoff frequency:

$$\lambda_{ave} \approx \frac{\lambda_{MAX} + \lambda_{MIN}}{2}\ . \qquad (10\text{-}17)$$

For example, if the system spectral response ranges from 0.4 to 0.7 μm, then $\lambda_{AVE} = 0.55\ \mu$m.

For most systems operating in the visible spectral region, the optical system may be considered near diffraction limited and in focus. Approximations for aberrated and defocused optical systems are in Reference 2. The MTFs for optical systems with rectangular apertures[3] or telescopes with a central obscuration (Cassegrainian optics) are also available.[4]

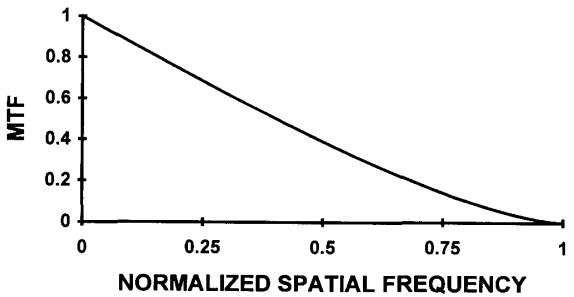


Figure 10-4. MTF_{OPTICS} for a circular aperture normalized to u_i/u_{iC}.

10.3. DETECTORS

The two-dimensional spatial response of a rectangular detector is

$$MTF_{DETECTOR}(u_i, v_i) = \left|sinc(d_H u_i)\right| \left|sinc(d_V v_i)\right| , \qquad (10\text{-}18)$$

where $\text{sinc}(x) = \sin(\pi x)/(\pi x)$ and d_H and d_V are the photosensitive detector sizes in the horizontal and vertical directions, respectively. Note that arrays are often specified by the pixel size. With 100% fill factor arrays, the pixel size is equal to the detector size. With finite fill factor arrays, the photosensitive detector dimensions are less than the pixel dimensions. A microlens or an anti-alias filter may optically modify the effective size (discussed in Section 10.3.4., *Optical Anti-Alias Filter*).

Although the detector MTF is valid for all spatial frequencies from $-\infty$ to $+\infty$, it is typically plotted up to the first zero (called the detector cutoff) which occurs at $u_{iD} = 1/d_H$ (Figure 10-5). Because d_H may be different than d_V, the horizontal and vertical MTF can be different.

A phase shift occurs when $u_i > 1/d_H$ and $v_i > 1/d_V$. Figure 10-5 illustrates the MTF without regard to the array Nyquist frequency. The analyst must add the array Nyquist frequency to the curve (see Section 8.3.2, *Detector Array Nyquist Frequency*, page 242). The MTF of virtual phase detectors is a complex expression.[5]

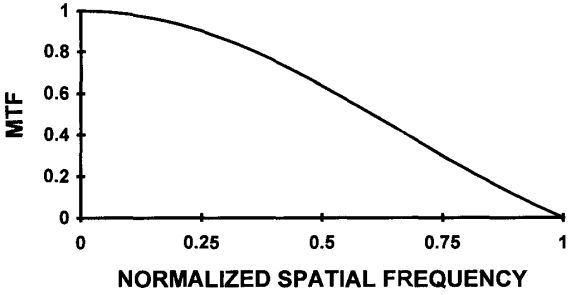


Figure 10-5. Horizontal detector MTF as a function of normalized spatial frequency $d_H u_i$. The MTF is usually plotted up to the first zero ($d_H u_i = 1$). For staring arrays, the array Nyquist frequency must be added: $u_{iN} = 1/(2d_{CCH})$.

10.3.1. DIFFUSION MTF

As the wavelength increases, photon absorption occurs at increasing depths in the detector material (see Section 4.1.1., *Spectral Response*, page 104). A photoelectron generated deep within the substrate will experience a three-dimensional random walk until it recombines or reaches the edge of a depletion region where the pixel electric field exists. A photoelectron generated under one well may eventually land in an adjoining well. This blurs the image and its effect is described by an MTF. Beyond 0.8 μm, diffusion (random walk) affects the detector MTF and the response is expressed by $MTF_{DETECTOR}MTF_{DIFFUSION}$. Imagery produced by near-IR sources will appear slightly blurry compared to imagery produced by a visible source.

The diffusion MTF[6] for a front-illuminated device is

$$MTF_{DIFFUSION}(u_i) = \frac{1 - \dfrac{\exp(-\alpha_{ABS} L_D)}{1 + \alpha_{ABS} L(u_i)}}{1 - \dfrac{\exp(-\alpha_{ABS} L_D)}{1 + \alpha_{ABS} L_{DIFF}}}, \qquad (10\text{-}19)$$

where α_{ABS} is the spectral absorption coefficient. Figure 4-2 (page 105) illustrated the absorption coefficient in bulk p-type silicon. The factor L_D is the depletion width and $L(u_i)$ is the frequency-dependent component of the diffusion length (L_{DIFF}) given by

$$L(u_i) = \frac{L_{DIFF}}{\sqrt{1 + (2\pi L_{DIFF} u_i)^2}}. \qquad (10\text{-}20)$$

The diffusion length typically ranges between 50 μm and 200 μm. The depletion width is approximately equal to the gate width. For short wavelengths ($\lambda < 0.6$ μm), α_{ABS} is large and lateral diffusion is negligible. Here, $MTF_{DIFFUSION}$ approaches unity. For the near IR ($\lambda > 0.8$ μm), the diffusion MTF may dominate the detector MTF (Figure 10-6). By symmetry, $MTF_{DIFFUSION}(u_i) = MTF_{DIFFUSION}(v_i)$ and $L(u_i) = L(v_i)$.

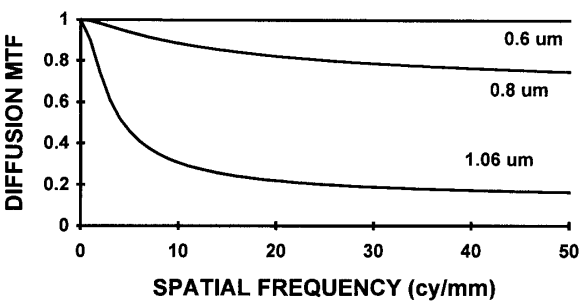


Figure 10-6. $MTF_{DIFFUSION}$ as a function of wavelength. The value L_D = 10 μm and L_{DIFF} = 100 μm. As the diffusion length increases, the MTF decreases. For comparison, a 20-μm detector will have a cutoff at 50 cycles/mm.

To account for the effects of lateral diffusion, Schumann and Lomheim[7] introduced a trapezoidal response approximation. This MTF only applies in the direction perpendicular to charge transport:

$$MTF_{DIFFUSION}(u_i) \approx sinc(d_{CCH}u_i)\, sinc[(d_{CCH} - \beta)u_i] \ , \qquad (10\text{-}21)$$

where β is the region over which the detector responsivity is flat (Figure 10-7). At the boundary between pixels, the trapezoidal response is 0.5 indicating equal probability of an electron going to either pixel.

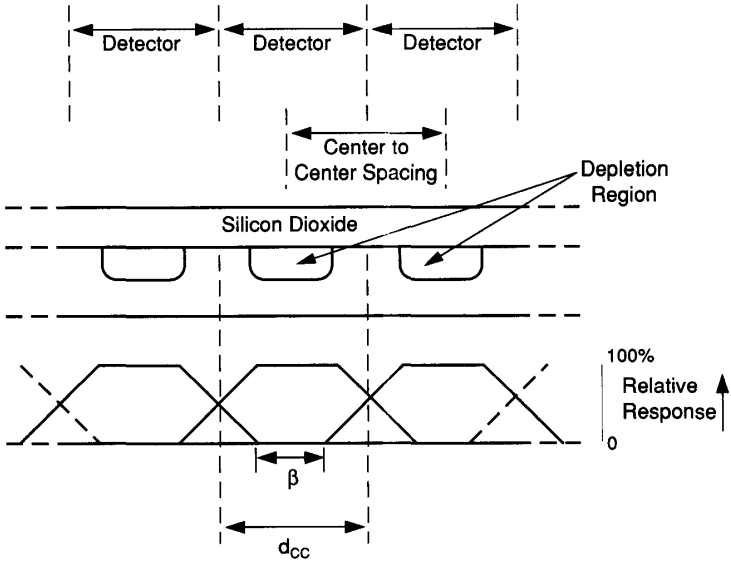


Figure 10-7. Trapezoidal approximation for long-wavelength photon absorption. Most of the charge goes to the charge well nearest to the photon absorption site.

10.3.2. CHARGE TRANSFER EFFICIENCY

CCD performance is characterized by the transfer efficiency and number of transfers (see Section 3.4., *Charge Transfer Efficiency*, page 74). Assuming vertical transfer, an MTF[8] that accounts for the incomplete transfer of electrons is

$$MTF_{CTE}(\nu_i) = \exp\left(-N_{TRANS}(1-\epsilon)\left[1-\cos\left(\frac{\pi\,\nu_i}{\nu_\epsilon}\right)\right]\right), \qquad (10\text{-}22)$$

where N_{TRANS} is the total number of charge transfers from a detector to the output amplifier and ϵ is the charge transfer efficiency for each transfer. It is the number of detectors multiplied by the number of CCD phases. When $N_{TRANS}(\epsilon - 1)$ is small, then $MTF(u_{iN}) \approx 1 - 2N_{TRANS}(1-\epsilon)$.

This MTF depends upon the number of charge transfers. If the image is located next to the readout, the number of transfers is small. But if the image is at the extreme end of the array, the charge must be transferred across the entire array and through the serial readout register. For an average response, N_{TRANS} should be one-half the maximum number of transfers. This is mathematically equivalent to $(MTF_{CTE})^{½}$.

Transfers may be either in the vertical direction (column readout) or in the horizontal direction (row readout). Horizontal readout affects $MTF(u_i)$ and vertical readout affects $MTF(v_i)$. MTF_{CTE} only applies to the transfer readout direction. Figure 10-8 illustrates MTF_{CTE} at Nyquist frequency for several values of transfer efficiency. For consumer devices, CTE > 0.9999 and N_{TRANS} > 1500. For scientific applications, CTE > 0.999999 and N_{TRANS} can exceed 5000.

Example 10-1
Number of Transfers

What is the maximum number of transfers that exist in an array that contains 1000 × 1000 detector elements?

For a charge packet that is furthest from the sense node, the charge must be transferred down one column (1000 pixels) and across 1000 pixels for a total of 2000 pixels. The number of transfers is 4000, 6000, and 8000 for two-phase, three-phase, and four-phase devices, respectively.

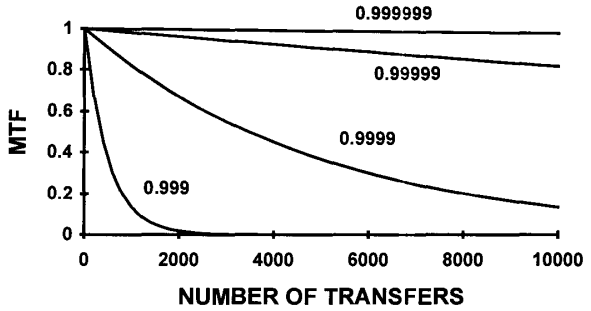


Figure 10-8. MTF_{CTE} at Nyquist frequency for several values of transfer efficiency. As the number of transfers increases, the efficiency must also increase to ensure that the MTF is not adversely affected.

10.3.3. TDI

TDI operation was described in Section 3.3.6., *Time Delay and Integration*, page 70. It is essential that the clock rate match the image velocity so that the photogenerated charge packet is always in sync with the image. Any mismatch will blur the image. Image rotation with respect to the array axes affects both the horizontal (TDI direction) and vertical (readout direction) MTFs simultaneously. For simplicity they are treated as approximately separable - $MTF(u_i,v_i) \approx MTF(u_i)MTF(v_i)$.

If the TDI direction is horizontal, velocity errors degrade[9] the detector MTF by

$$MTF_{TDI}(u_i) = \frac{\sin(\pi\, N_{TDI}\, d_{ERROR}\, u_i)}{N_{TDI} \sin(\pi\, d_{ERROR}\, u_i)} = \frac{sinc(N_{TDI} d_{ERROR}\, u_i)}{sinc(d_{ERROR}\, u_i)}, \quad (10\text{-}23)$$

where N_{TDI} is the number of TDI elements. The value d_{ERROR} is the difference between the expected image location and the actual location at the first pixel (Figure 10-9). After N_{TDI} elements, the image is displaced $N_{TDI}d_{ERROR}$ from the desired location. MTF_{TDI} is simply the MTF of an averaging filter for N_{TDI} samples displaced d_{ERROR} one relative to the next. As an averaging filter, the effective sampling rate is $1/d_{ERROR}$.

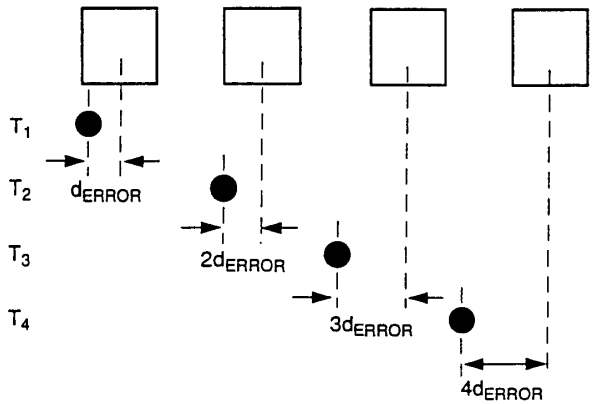


Figure 10-9. Definition of d_{ERROR}. With a velocity mismatch, the target falls further and further away from the desired location (on the center of the detector). After four TDI elements, the target is displaced from the detector center by $4d_{ERROR}$.

This is a one-dimensional MTF that apples in the TDI direction only. In image space

$$d_{ERROR} = |\Delta V| t_{INT} \quad mm \ , \tag{10-24}$$

where ΔV is the relative velocity error between the image motion and the charge packet "motion" or "velocity." The charge packet "velocity" is d_{CCH}/t_{INT} where t_{INT} is the integration time for each pixel.

Figure 10-10 portrays the MTF for two different relative velocity errors for 32 and 96 TDI stages. As the number of TDI stages increases, d_{ERROR} must decrease to maintain the MTF. Equivalently, the accuracy of knowing the target velocity must increase as N_{TDI} increases. Because the image is constantly moving, a slight image smear occurs[10] because the photogenerated charge is created at discrete times (e.g., at intervals of t_{CLOCK}). This linear motion MTF (discussed in Section 10.5.1., *Linear Motion*) is often small and therefore is usually neglected.

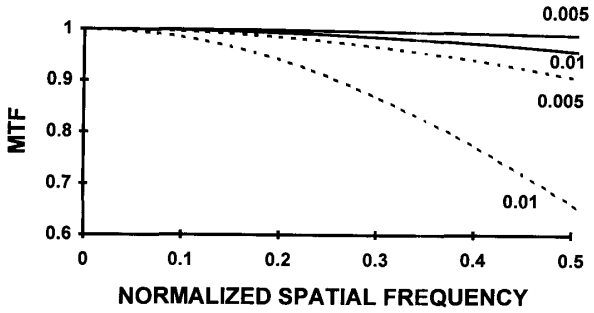


Figure 10-10. MTF degradation due to mismatch between the image velocity and the charge packet "velocity." The value d_{ERROR}/d_{CCH} equals 0.005 and 0.01. The spatial frequency axis is normalized to $d_{CCH}u_i$ and the MTF scale is expanded from 0.6 to 1.0. The solid lines represent $N_{TDI} = 32$ and the dashed lines are for $N_{TDI} = 96$.

TDI is also sensitive to image rotation. If the image motion subtends an angle θ with respect to the TDI rows (TDI direction), the image moves vertically $N_{TDI}d_{CCH}\tan(\theta)$ across the array. The MTF degradation in the vertical direction is

$$MTF_{TDI}(v_i) = \frac{\sin\left[\pi N_{TDI} d_{CCH} \tan(\theta)\, v_i\right]}{N_{TDI} \sin\left[\pi d_{CCH} \tan(\theta) v_i\right]}, \qquad (10\text{-}25)$$

where $MTF_{TDI}(v_i)$ is simply the MTF of an averaging filter for N_{TDI} samples displaced $d_{CCH}\tan(\theta)$ one relative to the next. Figure 10-11 illustrates the vertical MTF degradation due to image rotation for 32 and 96 stages. The spatial frequency is normalized to $d_{CCH}v_i$ where d_{CCH} is the *horizontal* detector pitch.

TDI offers the advantage that the signal-to-noise ratio increases by the square root of the number of TDI stages. This suggests that many TDI detectors should be used. The disadvantage is that the scan velocity must be tightly controlled as the number of TDI stages increases.

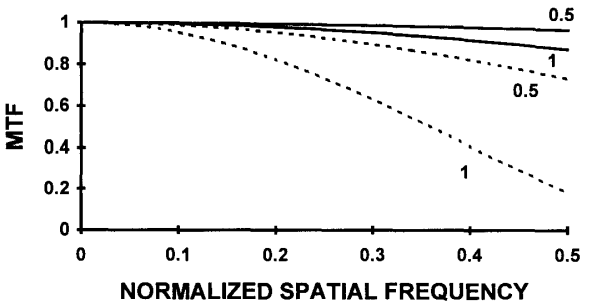


Figure 10-11. Vertical MTF degradation in the readout direction due to image rotation of 0.5 and 1 deg with respect to the array axis. The spatial frequency axis is normalized to $d_{CCH}v_i$. The solid lines represent $N_{TDI} = 32$ and the dashed lines are for $N_{TDI} = 96$.

10.3.4. OPTICAL ANTI-ALIAS FILTER

In color filter arrays, like-colored sensitive detector center-to-center spacings may be much larger than the detector pitch. This potentially increases aliasing (see Section 8.3.2., *Detector Array Nyquist Frequency*, page 242). The optical anti-alias filter blurs the image and effectively increases the optical area[11] of the detector. The filter, which consists of one or more birefringent crystals, is placed between the lens assembly and the detector array. Birefringent crystals break a beam into two components: the ordinary and the extraordinary. As illustrated in Figure 10-12, the detector collects light from a larger beam. Rays that would have missed the detector are refracted onto the detector. This makes the detector appear optically larger. While this changes the system MTF, the system sensitivity does not change. The light that is lost to adjoining areas is replaced by light refracted onto the detector.

The filter design depends upon the CFA design. If detectors sensitive to like colors are next to each other, the filter may not be required. If the detectors alternate, then an appropriately designed filter will make the detectors appear twice as large. Similarly, if the detectors are spaced three apart, then the filter will make the detector appear three times greater. This larger value (the effective detector size) is used in Equation 10-18 (page 275).

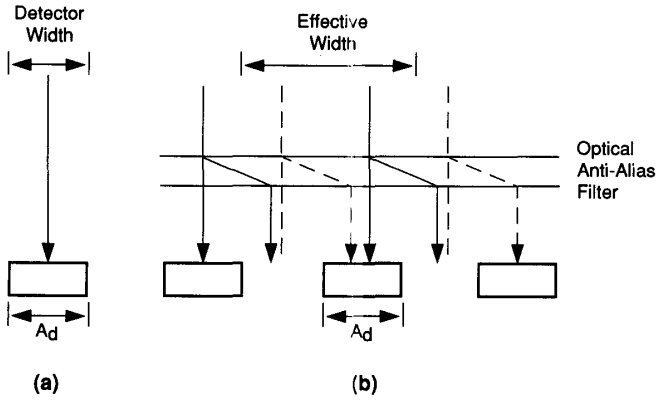


Figure 10-12. The birefringent filter increases the effective optical area of the detector. (a) No filter and (b) filter. Rays that would fall between detectors are refracted onto the detector.

Multiple crystals are used to change the effective detector size both vertically and horizontally. Because the spacing depends upon the CFA design, there is no unique anti-alias filter design. Clearly, if the detectors are unequally spaced, the Nyquist frequency for each color is different. Similarly, with different effective sizes, the MTFs are also different for each color. Complex algorithms are required to avoid color aliasing (see Example 5-1, page 167) and to equalize the MTFs. The anti-alias filter is essential for CFAs to reduce color moiré patterns and can be used in any camera to reduce aliasing.

Figure 10-13 illustrates how the detector MTF is modified by the birefringent crystal when the detector pitch is twice as large as the element size. While the filter reduces aliasing, it also reduces the overall system MTF. For scientific applications, this reduction may be too severe.

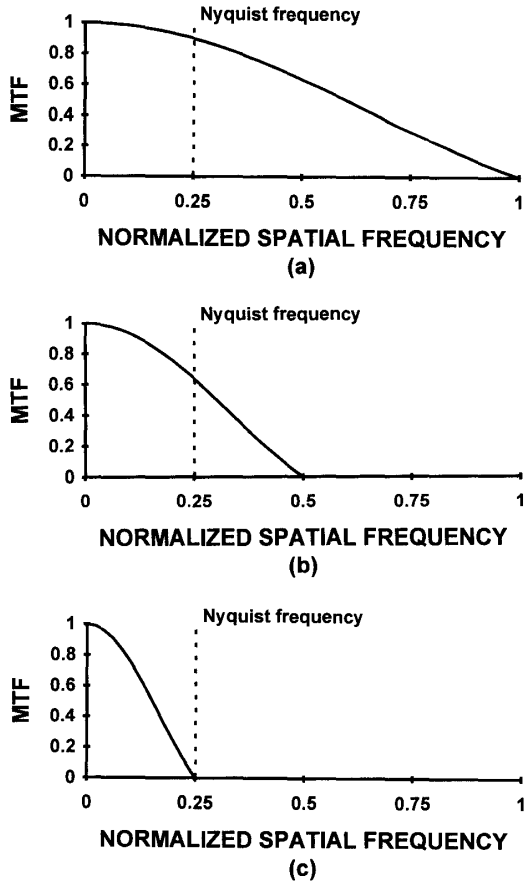


Figure 10-13. $MTF_{DETECTOR}$ modifications due to the anti-alias filter as a function of $d_H u_i$. The detector pitch is twice the detector element size. (a) No filter, (b) a filter that doubles the detector optical size, and (c) a filter that quadruples the effective detector size. The filter does not affect the array Nyquist frequency.

10.4. OPTICS-DETECTOR SUBSYSTEM

For many solid state cameras, the system MTF is dominated by the optics and detector MTFs. Here, the electronic MTFs maximize the signal. They are approximately unity over the spatial frequencies of interest. With electronic zoom, the effects of the display and observer MTFs can be minimized. Therefore, it is worthwhile to treat the optics and detector MTFs as a subsystem assembly and to examine this subsystem MTF.

When viewing a point source, diffraction-limited optics produces a diffraction pattern that consists of a bull's-eye pattern with concentric rings. The center is the Airy disk and its diameter is

$$d_{AIRY} = 2.44\frac{\lambda fl}{D_o} = 2.44\lambda F . \qquad (10\text{-}26)$$

Assume that the detector array has square detectors with 100% fill factor ($d_H = d_V = d_{CCV} = d_{CCH} = d$). When d_{AIRY} is larger than the detector size, the system is said to be optics-limited. Here, changes in the optics MTF (i.e., D_o or fl) significantly affect MTF_{SYSTEM}. If $d_{AIRY} < d$, the system is detector-limited and changes in the detector size affect MTF_{SYSTEM} at u_{iN}. Figure 10-14 illustrates $MTF_{OPTICS}MTF_{DETECTOR}$ at Nyquist frequency as a function of d/Fλ. It is the ratio of the optical cutoff to the detector cutoff: $u_{iC}/u_{iD} = d/F\lambda$.

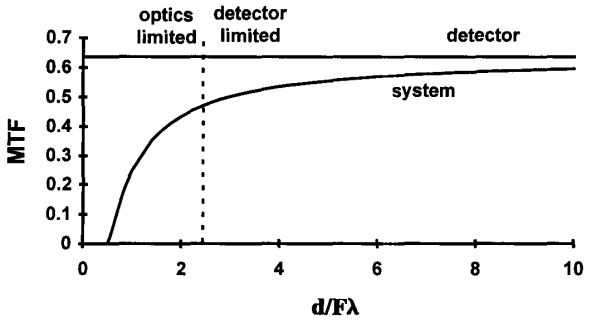


Figure 10-14. $MTF_{OPTICS}MTF_{DETECTOR}$ at Nyquist frequency as a function of d/Fλ. The fill factor is 100% ($d = d_H = d_V = d_{CCH} = d_{CCV}$). The vertical line indicates where the Airy disk diameter is equal to the detector size. The best MTF that can be achieved occurs when MTF_{OPTICS} is negligible compared to $MTF_{DETECTOR}$. Equivalently, $MTF_{OPTICS}MTF_{DETECTOR} \approx MTF_{DETECTOR} = 0.637$.

The MTF at Nyquist frequency is often used as a measure of performance. As the MTF increases, image quality should increase. Unavoidably, as the MTF increases, aliasing also increases[12] and image quality suffers. To avoid aliasing, the MTF at the Nyquist frequency must be zero. This requires an infinite f-number lens or an optical anti-alias filter (discussed in the previous section).

Smaller chips make smaller cameras (we live in a world were smaller is presumed more desirable). However, to maintain the MTF the f-number must decrease. For visible systems ($\lambda \approx 0.5\ \mu$m), the transition from the optics-limited to detector-limited case occurs when $F = d/1.22$. If $d = 10\ \mu$m, then F must be greater than 8.2 to enter the optics-limited condition (Figure 10-15).

For a small format chip, modest f-numbers provide an optics-limited camera system. For low illumination level imagery, the user will probably switch to a low f-number lens. If the camera enters the detector-limited region, pixelation becomes evident and image quality suffers. Large format chips are detector-limited for the more common f-number lenses. The transition from detector-limited to optics-limited was arbitrarily selected as $d = d_{AIRY}$. The transition is actually gradual.

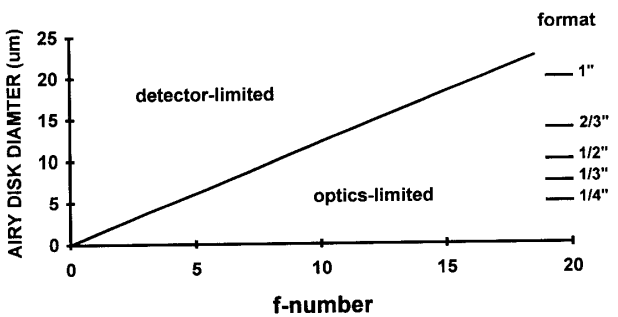


Figure 10-15. Airy disk diameter as a function of f-number when $\lambda = 0.5\ \mu$m. The nominal pixel sizes for the various video formats are from Table 3-3 (page 99).

With wet-film cameras, in most cases, more scene detail can be discerned by enlarging the print. This occurs because the silver granules in the film are very small compared to the Airy disk. That is, the camera system is almost always optics-limited. Enlarging the image provides more detail. With solid state cameras, this is not true. Electronic zoom (which is analogous to print enlargement) provides more detail only up to the point where individual pixels can be seen. There is no more detail available.

The output voltage (for square detectors) is given by the simple camera formula:

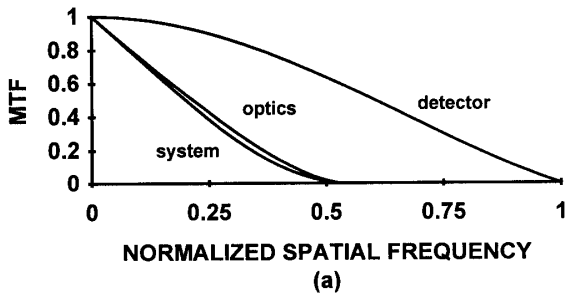
$$v_{CAMERA} = \frac{1}{4}\left(\frac{dD_o}{fl}\right)^2 \int_{\lambda_1}^{\lambda_2} R(\lambda)\, M_{illum}(\lambda)\, d\lambda \ . \qquad (10\text{-}27)$$

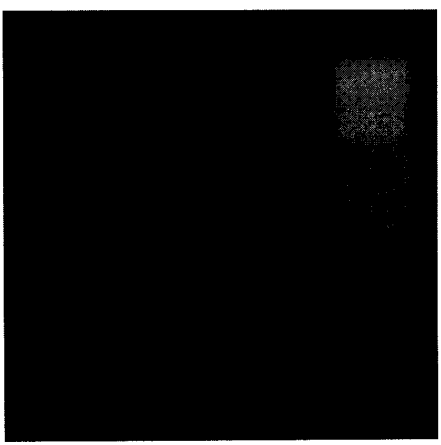
Decreasing the focal length, increases the MTF at u_{iN}. It also increases the camera output (improves sensitivity). But decreasing the focal length increases the scene spatial frequencies relative to the detector response. That is, the system's response moves from being detector-limited to being optics-limited. Thus, a tradeoff exists between sensitivity and image quality (as defined by the MTF).

While parametric analyses are convenient, absolute values are lost. The MTF at Nyquist frequency is important. However, it is even more important to know the relationship between the scene spatial content and the actual value of the Nyquist frequency. In object space, as the focal length decreases, the absolute value of the Nyquist frequency also decreases. It is instructive to see the effects of d/Fλ on computer-generated imagery. Details of the simulation are provided in Reference 13. Figure 10-16 illustrates the original scene.

Figure 10-16. Image created by the *System Image Analyzer* software.[14]

When $d/F\lambda = 0.5$, $MTF(u_N) = 0$ and the optics have low-pass filtered the scene. Here, no aliasing can occur. Figure 10-17a provides the MTFs and Figure 10-17b illustrates the resultant imagery which includes a Butterworth post-reconstruction filter where N = 10 (discussed in Section 10.6.4., *Post-Reconstruction Filter*). While the low-pass filtering action of the optics removed all aliasing, it also significantly attenuated the in-band frequency amplitudes.

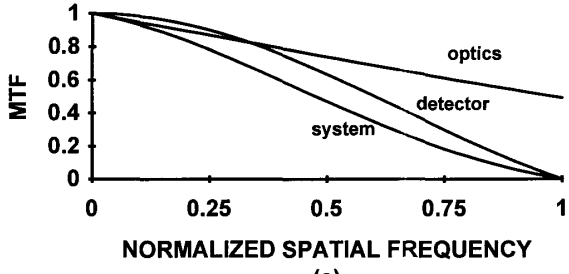




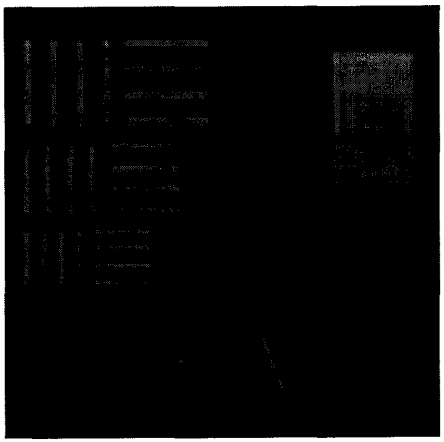
(b)

Figure 10-17. (a) MTF_{OPTICS} and $MTF_{DETECTOR}$ when $d/F\lambda = 0.5$ and (b) reconstructed image including a Butterworth filter where N = 10. The optics has significantly blurred the image and no aliasing exists. Created by the *System Image Analyzer* software.[14]

Figures 10-18 and 10-19 illustrate the cases where $d/F\lambda = 2.44$ and 10, respectively. Note that the MTFs illustrated in Figures 10-18a and 10-19a do not include the reconstruction filter whereas the imagery includes a 10^{th}-order Butterworth reconstruction filter.



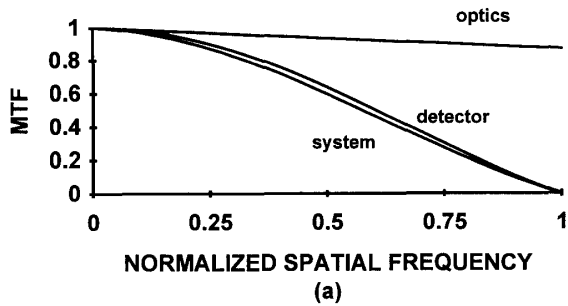
(a)



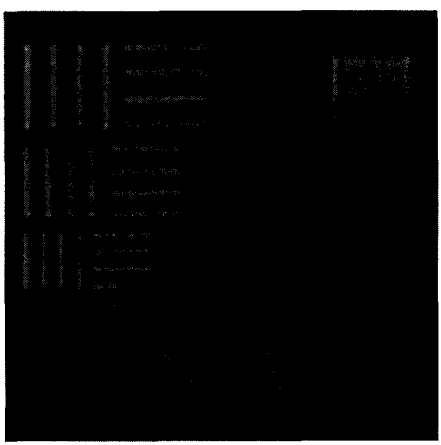
(b)

Figure 10-18. (a) MTF_{OPTICS} and $MTF_{DETECTOR}$ when $d/F\lambda = 2.44$ and (b) reconstructed image including a Butterworth filter where N = 10.

Careful examination of Figure 10-19b will reveal that the edges are sharper compared to Figure 10-18b but aliasing has increased. However, the effect is not dramatic when using a high order Butterworth reconstruction filter. At Nyquist frequency, $MTF_{OPTICS}MTF_{DETECTOR}$ increased from 0.472 to 0.592. In terms of imagery, whether $d/F\lambda = 10$ is better than $d/F\lambda = 2.44$ depends on the scene spectral content and the reconstruction filter used. The amount of aliased signal is scene-dependent and the amount of aliasing accepted depends on personal preference.



(a)



(b)

Figure 10-19. (a) MTF_{OPTICS} and $MTF_{DETECTOR}$ when $d/F\lambda = 10$ and (b) reconstructed image including a Butterworth filter where N = 10.

10.5. MOTION

Motion blurs imagery. When motion is fast compared to the detector integration time, details become blurred in a reproduced image. In real-time imagery, edges may appear fuzzy to the observer even though any one frame may provide a sharp image. The eye integration time blends many frames. The effects of motion during the entire integration and human interpretation process must be considered.

Linear motion includes both target movement (relative to the imaging system) and motion of the imaging system (stationary target). High frequency random motion is simply called jitter. Typically, an imaging system is subjected to linear and random motion simultaneously:

$$MTF_{MOTION} = MTF_{LINEAR}\, MTF_{RANDOM} \,. \tag{10-28}$$

In a laboratory, the system and test targets may be mounted on a vibration-isolated stabilized table. Here, no motion is expected so that $MTF_{MOTION} = 1$.

10.5.1. LINEAR MOTION

The OTF degradation due to horizontal linear motion is:

$$OTF_{LINEAR}(u) = sinc(a_L u) \,, \tag{10-29}$$

where a_L is the distance the target edge moved across the image plane. It is equal to $v_R \Delta t$ where v_R is the relative image velocity between the sensor and the target. For scientific and machine vision applications, the exposure time is equal to the detector integration time. But observer's eye blends many frames of data and the eye integration time should be used. Although the exact eye integration time is debatable, the values most often cited are between 0.1 and 0.2 s.

Figure 10-20 illustrates the OTF due to linear motion as a fraction of the detector size, d. Recall that the MTF is the magnitude of the OTF. Linear motion only affects the MTF in the direction of the motion. As a rule-of-thumb, when the linear motion causes an image shift less than about 20% of the detector size, it has minimal effect on system performance.

As the motion increases, a phase reversal occurs which is represented by a negative OTF. Here, periodic black bars appear white and periodic white bars appear black. Figure 10-21 illustrates a test pattern smeared by linear motion. At cutoff (where the system MTF approaches the first zero), the bar pattern cannot be resolved. Phase reversal occurs above cutoff.

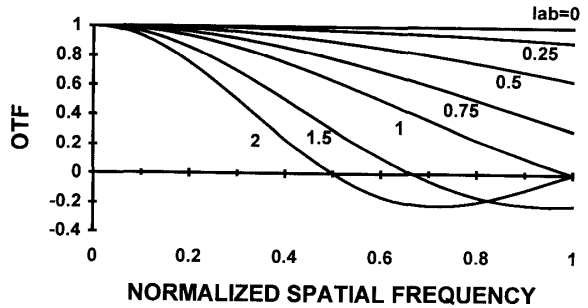
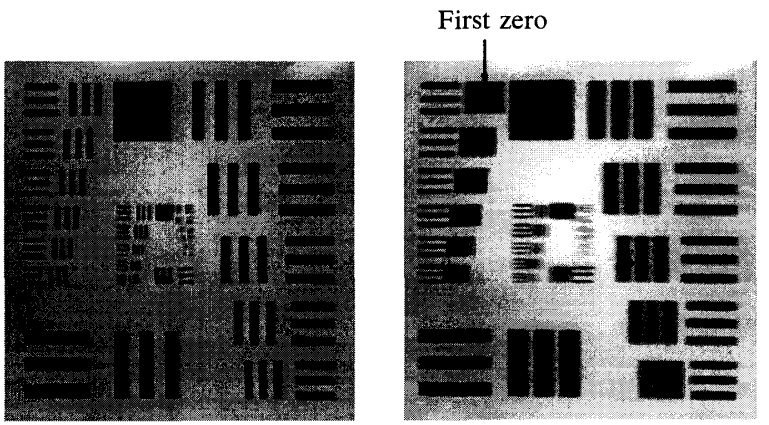


Figure 10-20. OTF_{LINEAR} as a function of normalized spatial frequency du_i. The curves represent values of a_L/d. In the laboratory a_L is usually zero. The negative OTF represents phase reversal. The MTF is always positive.



ORIGINAL IMAGE LINEAR MOTION

Figure 10-21. Linear motion smears imagery. Because the motion is horizontal, the vertical resolution is maintained. When the bar target frequency is equal to the first zero of OTF_{LINEAR}, the bars all blend together. Higher spatial frequency bars illustrate phase reversal: There appear to be two black bars on a gray background. (From Reference 15: by courtesy of G. L. Conrad).

10.5.2. RANDOM MOTION (JITTER)

With high frequency motion, it is assumed that the image has moved often during the integration time so that the central limit theorem is valid. The central limit theorem says that many random movements can be described by a Gaussian distribution. The Gaussian MTF is

$$MTF_{RANDOM}(u_i) = e^{-2\pi^2\sigma_R^2 u_i^2}, \tag{10-30}$$

where σ_R is the rms random displacement in millimeter. Figure 10-22 illustrates the random motion MTF. As a rule-of-thumb, when the rms value of the random motion is less than about 10% of d (detector size), system performance is not significantly affected. Jitter is considered equal in all directions and is included in both the horizontal and vertical system MTFs. High frequency motion blurs imagery (Figure 10-23).

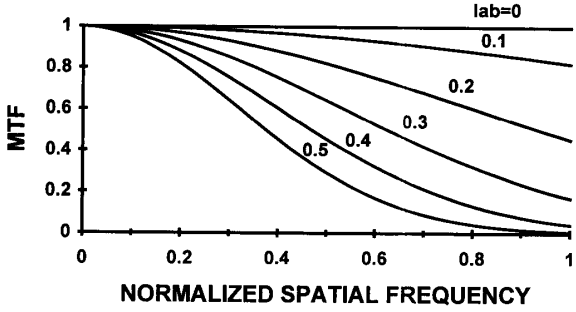
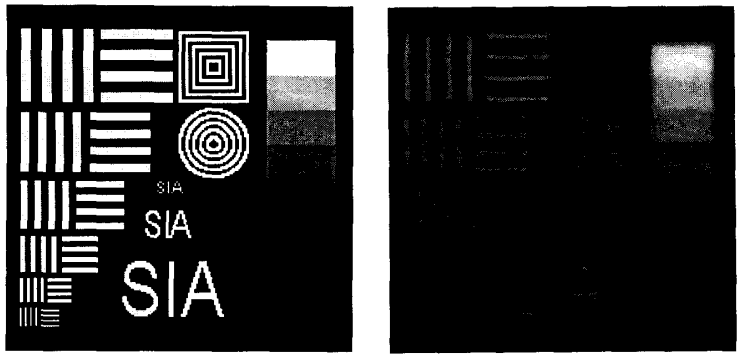


Figure 10-22. MTF degradation due to high frequency random (Gaussian) movement as a function of $\sigma_R u_i$. In the laboratory, σ_R is usually zero.



(a) **(b)**

Figure 10-23. Random motion blurs imagery. (a) Original image, (b) two-dimensional random motion. Generated[14] by SIA.

10.6. ELECTRONIC FILTERS

When filter response is provided as a function of electrical frequency, it is denoted by $H(f_e)$ or $H(f_v)$. When the electrical frequency is expressed in object space or image space units, the filter impulse response is labeled as an MTF. Boost filters may be used to enhance the signal at selected spatial frequencies. Both analog and digital filters can be used.

10.6.1. ANTI-ALIAS ANALOG FILTER

Ideally, $H(f_e)$ is a filter with MTF of unity up to the array Nyquist frequency and then drops to zero. This filter maximizes the SNR by attenuating out-of-band amplifier noise and passing the signal without attenuation. The ideal filter is unrealizable but can be approximated by N^{th}-order Butterworth filters:

$$H_{LOWPASS}(f_e) = \frac{1}{\sqrt{1 + \left(\frac{f_e}{f_{e3dB}}\right)^{2N}}}, \tag{10-31}$$

where f_{e3dB} is the frequency at which the power is one-half or the amplitude is 0.707. As $N \rightarrow \infty$, $H_{LOWPASS}(f_e)$ approaches the ideal filter response with the cutoff frequency of f_{e3dB} (Figure 10-24).

When f_{e3dB} is approximately equivalent to the array Nyquist frequency, the sampling theorem can be satisfied. However, square waves (bar targets) will appear as sinusoids. Thus, f_{e3dB} must be much higher to preserve the square-wave response.

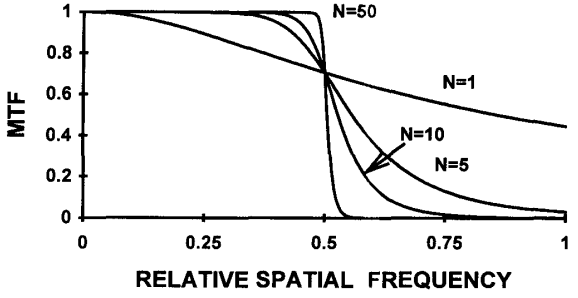


Figure 10-24. Butterworth filters for N = 1, 5, 10, and 50. The value $f_{e3dB} = 0.5$.

10.6.2. DIGITAL FILTERS

There are many available image processing algorithms. Only a few of these can be described mathematically in closed-form and only these can be included in an end-to-end system performance model. The performance of the remaining algorithms can be inferred only by viewing the system output for a few representative inputs. Because the system is not spatially invariant, predicting system performance when viewing any other target is pure conjecture.

The sampling process replicates the frequency spectrum at nf_{eS} (see Section 8.1., *Sampling Theorem*, page 234). Similarly, digital filter response is symmetrical about the Nyquist frequency and repeats at multiples of the sampling frequency. The highest frequency of interest is the Nyquist frequency.

Digital filters process data that reside in a memory. The units assigned to the filter match the units assigned to the data arrays. With one-to-one mapping of pixels to datels, the filter sampling frequency is the same as the array sampling frequency. With this mapping, each filter coefficient processes one pixel value. Digital filters can be two-dimensional.

There are two general classes of digital filters[16]: infinite impulse response (IIR) and finite impulse response (FIR). Both have advantages and disadvantages. The FIR has a linear phase shift whereas the IIR does not. IIR filters tend to have excellent amplitude response whereas FIR filters tend to have more ripple. FIR filters are typically symmetrical in that the weightings are symmetrical about the center sample. They are also the easiest to implement in hardware or software.

Figure 10-25 illustrates two FIR filters. The digital filter design software provides the coefficients, A_i. The central data point is replaced by the digital filter coefficients as they operate on the neighboring data points. The filter is then moved one data point and the process is repeated until the entire data set has been operated on. Edge effects exist with any digital filter. The filter illustrated in Figure 10-25a requires seven inputs before a valid output can be achieved. At the very beginning of the data set, there are insufficient data points to have a valid output at data point 1, 2, or 3. The user must be aware of edge effects at both the beginning and the end of his data record. In effect, this states that edges cannot be filtered.

The following MTF equations are only valid where there are no edge effects. For FIR filters where the multiplicative factors (weightings) are symmetrical about the center, the filter is mathematically represented by a cosine series (sometimes called a cosine filter). With its frequency response transposed to electrical frequency, an FIR filter with an odd number of samples (also called taps) provides

$$H_{DFILTER}(f_e) = \left| \sum_{k=0}^{\frac{N-1}{2}} A_k \cos\left(\frac{2\pi k f_e}{f_{eS}} \right) \right| . \qquad (10\text{-}32)$$

For an even number of samples,

$$H_{DFILTER}(f_e) = \left| A_o + \sum_{k=1}^{\frac{N}{2}} A_k \cos\left(\frac{2\pi (k-1) f_e}{f_{eS}} \right) \right| . \qquad (10\text{-}33)$$

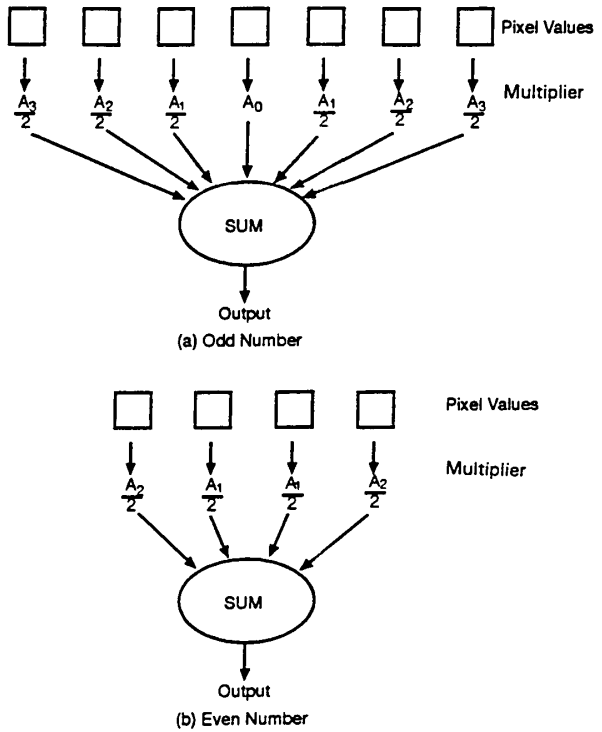


Figure 10-25. Symmetrical digital filters. (a) 7-tap (odd number) filter and (b) 4-tap (even number) filter.

The sum of the coefficients should equal unity so that the $H_{DFILTER}(f_e = 0)$ is one:

$$\sum A_k = 1 \, . \tag{10-34}$$

Although the above equations provide the filter response in closed form, the response of a real filter is limited by the ability to implement the coefficients. The smallest increment is the LSB.

With an averaging filter,[9] all the multipliers shown in Figure 10-25 are equal. When N_{AVE} datels are averaged, the equivalent MTF is

$$H_{AVE}(f_e) = \left| \frac{\sin\left(N_{AVE}\pi \frac{f_e}{f_{eS}}\right)}{N_{AVE}\sin\left(\pi \frac{f_e}{f_{eS}}\right)} \right| . \tag{10-35}$$

The first zero of this function occurs at f_{eS}/N_{AVE}.

Averaging samples together is called binning or super pixeling. It provides the same MTF as if the detector elements were proportionally larger. That is, averaging two samples together provides the same MTF as a single detector element that is twice as large. Cameras that operate in either the pseudo-interlace or the progressive scan modes will have a different vertical MTF for each mode. In the pseudo-interlace mode, two pixels are averaged together. If $d_H = d_{CCH}$, then

$$\begin{aligned} MTF_{DETECTOR}MTF_{DFILTER} &= sinc(d_H f_e)\frac{sinc\left(N_{AVE}\frac{f_e}{f_{eS}}\right)}{sinc\left(\frac{f_e}{f_{eS}}\right)} \\ &= sinc(N_{AVE}d_H f_e) . \end{aligned} \tag{10-36}$$

where the effective detector width is $N_{AVE}d_H$.

10.6.3. SAMPLE-AND-HOLD

After the digital-to-analog converter (DAC), the analog signal changes only at discrete times. The sample-and-hold circuitry within the DAC extends the data into a continuous analog signal. The MTF of a zero-order sample-and-hold when transposed into image space is:

$$MTF_{S\&H} = sinc\left(\frac{u_i}{u_{iS}}\right) , \tag{10-37}$$

where $u_{iS} = 1/d_{CCH}$. Figure 10-26 illustrates the sample-and-hold MTF as a function of normalized spatial frequency.

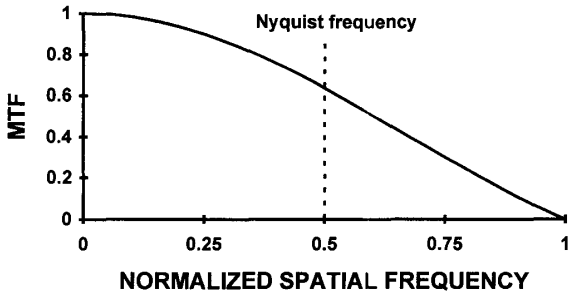


Figure 10-26. Zero-order sample-and-hold MTF normalized to $d_{CCH}u_i$. The sample-and-hold acts as a low-pass filter.

10.6.4. POST-RECONSTRUCTION FILTER

After the digital-to-analog conversion, the image still appears blocky due to the discrete nature of the conversion. The ideal post-reconstruction filter removes all the higher order frequencies (Figure 10-27) such that only the original smooth signal remains. The output is delayed but this is not noticeable on most imagery.

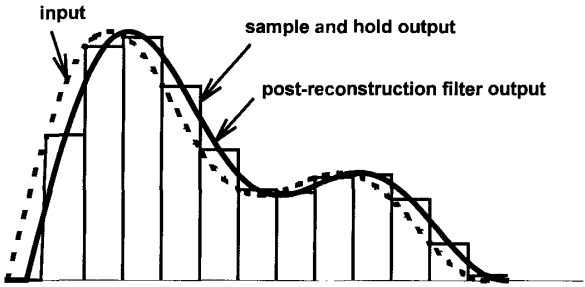


Figure 10-27. The post-reconstruction filter removes the blocky (stair-step) effect created by the sample-and-hold circuitry within the DAC.

Ideally, $H(f_v)$ is a filter with MTF of unity up to the array Nyquist frequency and then drops to zero. The ideal filter is unrealizable but can be approximated by N^{th}-order Butterworth filters:

$$H_{LOWPASS}(f_v) = \frac{1}{\sqrt{1 + \left(\frac{f_v}{f_{v3dB}}\right)^{2N}}}, \tag{10-38}$$

where f_{v3dB} is the frequency at which the power is one-half or the amplitude is 0.707. As $N \rightarrow \infty$, $H_{LOWPASS}(f_v)$ approaches the ideal filter response with the cutoff frequency of f_{v3dB}. Real filters will have some roll-off and f_{v3dB} should be sufficiently large so that it does not affect the in-band MTF. The post-reconstruction filter can be of the same functional form as the anti-alias filter (see Section 10.6.1., page 295) but serves a different purpose. It removes the spurious frequencies created by the sampling process.

For analog video output, A low-pass filter is used as a post-reconstruction filter. This filter affects only the horizontal signal. While the video bandwidth has been standardized (NTSC, PAL, and SECAM are 4.2, 5.5, and 6 MHz, respectively), the filter design has not. Thus, different cameras may have different post-reconstruction filters and therefore different horizontal MTFs even though they are built to the same video standard. No reconstruction filter is used in the vertical direction. Here, the display and human visual response provide the reconstruction. As a result of different detector sizes (d_V and d_H) and horizontal analog filters, the MTFs in the vertical and horizontal directions can be quite different.

☞

Example 10-2
POST-RECONSTRUCTION FILTER

What is the required spatial frequency cutoff for an ideal post-reconstruction filter if the desired output is EIA 170 and there are 640 horizontal detectors in the array?

For EIA 170, t_{LINE} is 52.09 μs. The Nyquist frequency in the video domain is

$$f_{VN} = \frac{1}{2}\frac{640}{52.09\times10^{-6}} = 6.14\ Mhz \tag{10-39}$$

The bandwidth necessary to transmit the detector data must exceed 6.14 MHz to avoid roll-off effects. This is a greater bandwidth than allowed by the EIA 170 standard. However, the camera output only has to conform to EIA 170 timing so that the information can be displayed on a conventional monitor.

10.6.5. BOOST

MTF degradation caused by the various subsystems can be partially compensated with electronic boost filters. Boost filters also amplify noise so that the signal-to-noise ratio may degrade. For systems that are contrast-limited, boost may improve image quality. However, for noisy images, the advantages of boost are less obvious. As a result, these filters are used only in high contrast situations (typical of consumer applications) and are not used in scientific applications where low signal-to-noise situations are often encountered. Historically, each subsystem was called an "aperture." Therefore, boost provided "aperture correction." This should not be confused with the optics MTF where the optical diameter is the entrance aperture. "Aperture correction" has been used in video for years.

The boost amplifier can either be an analog or a digital circuit with peaking that compensates for any specified MTF roll-off. The MTF of a boost filter, by definition, exceeds one over a limited spatial frequency range. When used with all the other subsystem MTFs, the resultant MTF_{SYSTEM} is typically less than one for all spatial frequencies. Excessive boost can cause ringing at sharp edges. The boost algorithm should enhance signal amplitudes with frequencies below the Nyquist because reconstruction usually limits the response to the Nyquist frequency.

A variety of boost circuits are available.[17] One example is the tuned circuit with response

$$H_{BOOST}(f_v) = \frac{1}{\sqrt{\left(1 - \left(\frac{f_v}{f_{BOOST}}\right)^2\right)^2 + \left(\frac{f_v}{Qf_{BOOST}}\right)^2}}, \quad (10\text{-}40)$$

where Q is the quality factor and equal to boost amplitude when $f_v = f_{BOOST}$.

Although illustrated in the video domain, boost can be placed in any part of the circuit where a serial stream of data exists. By appropriate selection of coefficients, digital filters can also provide boost. MTF_{BOOST} can also approximate the inverse of MTF_{SYSTEM} (without boost). The combination of MTF_{BOOST} and MTF_{SYSTEM} can provide unity MTF over spatial frequencies of interest. This suggests that the reproduced image will precisely match the scene in every spatial detail.

10.7. CRT-BASED DISPLAY

For most electronic imaging systems, the display medium acts as a reconstruction filter. Displays do not completely attenuate the replicated frequency amplitudes created by the sampling process. The HVS further reduces the sampling remnants so that the perceived image appears continuous with minimal sampling artifacts. Therefore, displays cannot be analyzed in isolation but must be considered as part of the display/observer system.

The display MTF is a composite that includes both the internal amplifier and the CRT responses. Implicit in the MTF is the conversion from input voltage to output display brightness. Although not explicitly stated, the equation implies radial symmetry and the MTFs are the same in both the vertical and horizontal directions. That is, the MTF is considered separable and $MTF_{CRT}(u_d, v_d) = MTF_{CRT}(u_d)MTF_{CRT}(v_d)$ and $MTF_{CRT}(v_d) = MTF_{CRT}(u_d)$. While this may not be precisely true, the assumption is adequate for most calculations.

The amplifier bandwidth is not a number that is often useful to the user. We must assume that the manufacturer selected an adequate bandwidth to provide the display resolution. For example, a 1600 × 1200-disel display must have a wider bandwidth than a 1280 × 1024-disel display by a factor of approximately $1600/1280 = 1.25$. MTF_{CRT} includes both the electronics MTF and spot diameter. For an infinitesimal input, the spot diameter is affected by the video bandwidth. That is, the minimum spot diameter is the equivalent point spread function.

It is reasonable to assume that the resultant spot is Gaussian (see Section 7.3., *Spot Size*, page 210). Then,

$$MTF_{CRT}(u_d) \approx e^{-2\pi^2 \sigma_{SPOT}^2 u_d^2} = e^{-2\pi^2 \left(\frac{S}{2.35} u_d\right)^2}, \tag{10-41}$$

where σ_{SPOT} is the standard deviation of the Gaussian beam profile and S is the spot diameter at FWHM. These parameters must have units of millimeters because the frequency has units of cycles/mm. The spot diameter is

$$S = \sqrt{8\ln(2)}\, \sigma_{SPOT} = 2.35\, \sigma_{SPOT}. \tag{10-42}$$

If the spot size is not specified, then σ_{SPOT} can be estimated from the TV limiting resolution. With the TV limiting resolution test, an observer views a wedge pattern and selects that spatial frequency at which the converging bar pattern can no longer be seen. The display industry has standardized the bar spacing to $2.35\sigma_{SPOT}$.

If N_{TV}/PH is the TV limiting resolution and N_{TV} lines are displayed on a monitor, then $H_{MONITOR} \approx 2.35\sigma_{SPOT}N_{TV}$. When transposed to image space,

$$MTF_{CRT}(u_i) \approx e^{-2\pi^2 \left(\frac{VFOV\,fl}{2.35\, N_{TV}} u_i\right)^2}. \tag{10-43}$$

For staring arrays with N_V displayed vertical lines,

$$MTF_{CRT}(u_i) = e^{-2\pi^2 \left(\frac{(N_V - 1) d_{CCV} + d_V}{2.35\, N_{TV}} u_i\right)^2}. \tag{10-44}$$

Figure 10-28 illustrates $MTF_{DETECTOR}$ and $MTF_{DISPLAY}$. The display spot size is independent of the detector size and the display resolution is independent of the number of detector elements. Matching the display resolution (number of pixels) to the number of detector elements does not uniquely specify the display MTF as suggested by Equation 10-44. For example, suppose an array contained 1000 elements horizontally. A monitor that can resolve 1000 pixels may not provide good imagery if the resolution/addressability ratio (RAR) is large. Equation 10-44 is only valid when RAR $\approx$ 1. The RAR was described in Section 7.6., *Addressability* (page 222).

Many consumer CCD cameras match the number of vertical detectors to the number of active video lines. With consumer televisions, $N_{TV} \approx 0.7N_V$. Thus, the video standard uniquely defines the CRT MTF. For 100% fill factor arrays,

$$MTF_{CRT}(u_i) \approx e^{-2\pi^2 \left(\frac{d_{CCV}}{1.645} u_i\right)^2}. \tag{10-45}$$

At Nyquist frequency, $MTF_{CRT} = 0.161$. That is, the "standard" design effectively reduces sampling effects by attenuating the amplitudes above u_{iN}. However, the in-band amplitudes are unavoidably attenuated also.

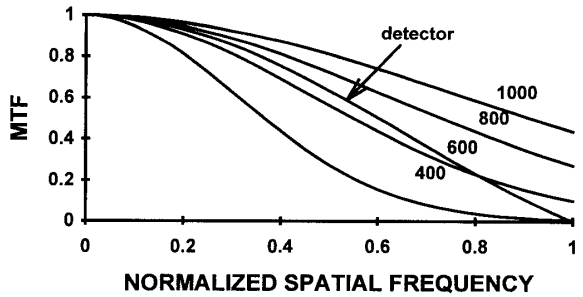


Figure 10-28. $MTF_{DISPLAY}$ as a function of $d_{CCV}v_i$ for four different TV limiting resolutions when $N_V = 480$. The display MTF limits most systems. The Nyquist frequency (not shown) is at $d_{CCV}v_i = 0.5$.

☞

Example 10-3
DISPLAY MTF

What should the display resolution be to maximize the system MTF?

Equations 10-43 or 10-44 are used to calculate image MTF when the input is a sinusoid. The Nyquist frequency is used most often when specifying system response. For EIA 170, $N_V = 485$. Inserting $d_V u_i = 0.5$ into Equation 10-45 provides

$$MTF_{CRT}(u_{iN}) = e^{-2\pi^2\left(\frac{0.5}{2.35}\frac{485}{N_{TV}}\right)^2} \quad (10\text{-}46)$$

For a 100% fill factor array, the detector MTF is 0.63 at u_{iN}. Figure 10-29 illustrates the detector and display MTF at Nyquist frequency. The display resolution must increase to 1400 TVL/PH for display MTF to be 0.90 at the Nyquist frequency. This suggests that most imaging systems are display resolution limited. This result is based upon the assumption that the beam profile is Gaussian. A different shaped beam (i.e., flat-topped), will provide a different relationship between TV limiting resolution and MTF.

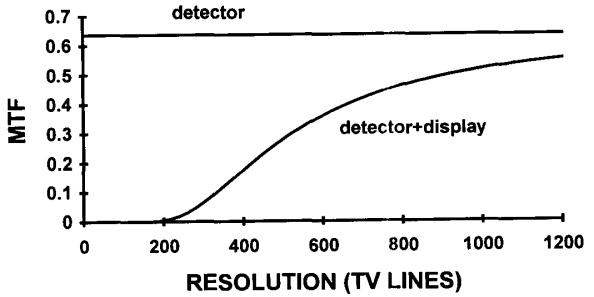


Figure 10-29. $MTF_{DETECTOR}MTF_{DISPLAY}$ at Nyquist frequency as a function of TV liming resolution. The detectors are contiguous so that the center-to-center spacing is equal to the detector element size. The value $N_V = 485$. All other MTFs are assumed to be unity at u_{iN}.

10.8. HUMAN VISUAL RESPONSE

The observer processes imagery to create a perceived MTF:

$$MTF_{PERCEIVED} = MTF_{SYS}\, MTF_{HVS} \,. \tag{10-47}$$

The HVS MTF should only be used to understand the relationship between the imaging system MTF and the eye response. It is not used to describe sensor performance. When $MTF_{PERCEIVED}$ is combined with system noise, system responsivity, and the eye integration capabilities, it is possible to predict the minimum resolvable contrast (described in Chapter 12, *Minimum Resolvable Contrast*). The eye's contrast sensitivity is also included in all image quality metrics (discussed in Chapter 11, *Image Quality*).

The sine wave response is used as an approximation to the eye-brain MTF. It is sometimes called the human visual system MTF or HVS-MTF. The sine wave response depends on diffraction by the pupil, aberrations of the lens, finite size of the photoreceptors, ocular tremor, and neural interconnections within the retina and brain. Diffraction and aberrations vary with ambient luminance, monitor brightness, and chromatic composition of the light. Because the retina is composed of rods and cones of varying densities, the location and size of the object with respect to the fovea significantly affect the sine wave response.

The sine wave response ignores spatial noise, background luminance, angular orientation and exposure time. Each of these parameters significantly affects the interpretation of image visibility. The MTF in the purest sense is noise independent but the human eye's response is very sensitive to spatial and temporal noise. Therefore, the sine wave response is only an approximation to the true response. Furthermore, the overall population exhibits large variations in response. Any MTF approximation used for the HVS therefore is only a crude approximation and probably represents the largest uncertainty in the overall MTF analysis approach. Factors that affect observer performance include monitor viewing ratio (usually selected to reduce raster effects), target size, and observation time. Visual psychophysical data can be found in Farrell and Booth's[18] and Biberman's[19] books.

The usual measurement is to determine the minimum modulation required to just perceive a sinusoidal target. The modulation is

$$M_t = \frac{L_{MAX} - L_{MIN}}{L_{MAX} + L_{MIN}}, \tag{10-48}$$

where L_{MAX} and L_{MIN} are the maximum and minimum luminances, respectively. The eye's detection capability depends upon the visual angle subtended by the target size at the observer. As shown in Figure 10-30, in the absence of noise, the eye's modulation threshold (minimum perceivable modulation) is characteristically J-shaped. It is labeled as M_t in the visual psychophysical literature. The eye is most sensitive to spatial frequencies[20] that range between 3 and 5 cycles/deg at typical ambient lighting levels.

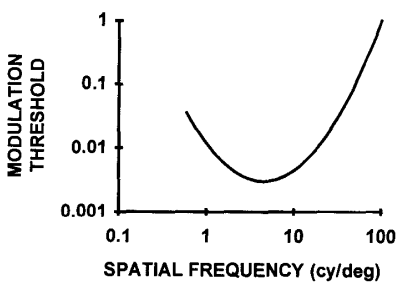


Figure 10-30. A representative modulation threshold function.

The increase in threshold at low frequencies is due to the eye's inhibitory signal processing component. The contrast sensitivity is the inverse of the modulation threshold curve. When normalized to one, the contrast sensitivity or sine wave response is called the HVS MTF. By convention, the MTF is defined to be one at zero spatial frequency and to be a monotonically decreasing function. The way to include the inhibitory response as an MTF component is not clear.

Nevertheless, various researchers modeled the normalized contrast sensitivity and labeled it as the HVS MTF (Figure 10-31). Nill[21] recommended

$$MTF_{HVS}(f_{eye}) = 2.71\left[0.19 + 0.81\left(\frac{f_{eye}}{f_{PEAK}}\right)\right] e^{-\left(\frac{f_{eye}}{f_{PEAK}}\right)} . \qquad (10\text{-}49)$$

Schulze[22] recommended

$$MTF_{HVS}(f_{eye}) = 2.71\left(e^{-0.1138 f_{eye}} - e^{-0.325 f_{eye}}\right) . \qquad (10\text{-}50)$$

Campbell and Robson's high luminance data[23] can be approximated by

$$MTF_{HVS}(f_{eye}) = 10^{-1.4 \log_{10}\left(\frac{f_{eye}}{f_{PEAK}}\right)^2} . \qquad (10\text{-}51)$$

deJong and Bakker[24] used

$$MTF_{HVS}(f_{eye}) = \sin^2\left[\frac{\pi}{2}\left(\frac{f_{eye}}{f_{PEAK}}\right)^{1/3}\right] . \qquad (10\text{-}52)$$

Neglecting the low frequency inhibitory response, Kornfeld and Lawson[25] suggested that

$$MTF_{HVS}(f_{eye}) = e^{-\Gamma \frac{f_{eye}}{17.45}} \qquad (10\text{-}53)$$

where Γ is a light-level-dependent eye response factor that was presented in tabular form[26] by Ratches et al. A third order polynomial fit provides

$$\Gamma = 1.444 - 0.344 \log(B) + 0.0395 \log^2(B) + 0.00197 \log^3(B) , \qquad (10\text{-}54)$$

where B is the monitor brightness in foot-Lamberts.

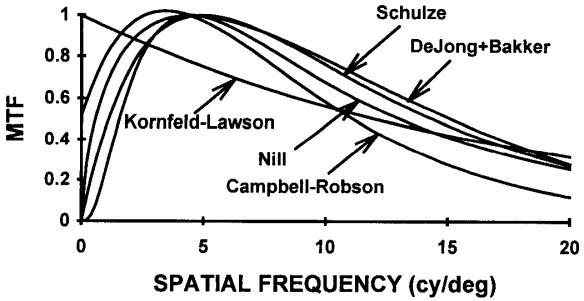


Figure 10-31. Various mathematical equations that approximate the HVS MTF.

Considering the large variation in observer responses, these curves have a similar shape when $f_{eye} > f_{PEAK}$. The original U.S. Army Night Vision Laboratory's static performance model[26] and FLIR92[27] use the Kornfeld-Lawson eye model. These models were validated by the U.S. Army when detecting tank-sized targets at modest ranges. This translates into modest spatial frequencies. The Kornfeld-Lawson model appears to follow all other models (to within a multiplicative factor) for mid-range spatial frequencies. This multiplicative factor becomes a normalization issue for the MRC (discussed in Chapter 12, *Minimum Resolvable Contrast*).

Insufficient data exist to say with certainty which eye model is best. Because of this uncertainty, *all analyses, no matter which eye model is used, must only be used for comparative performance purposes.*

Two scenarios are possible: (1) the observer is allowed to move his head and (2) the head is fixed in space. Because the eye's detection capability depends upon the angular subtense of the target, head movement may provide different results than if the head is fixed. In laboratory testing, the distance to the monitor is not usually specified nor limited in any way. To maximize detection capability (stay on the minimum of the contrast threshold curve), an observer subconsciously moves toward the monitor to perceive small targets and further away to see larger targets.

By allowing the observer to adjust the viewing distance, several interrelated detection criteria are optimized that include striving for apparent edge sharpness and maximizing perceived signal-to-noise ratio. This apparently results in an

equal detection capability for all spatial frequencies such that the eye's contrast sensitivity approaches a constant. This results in a nearly constant MTF that is called the "non-limiting eye MTF" or $MTF_{HVS} = 1$. Equation 10-53 is used in those situations where the head is fixed, such as when a pilot is strapped into a cockpit seat.

10.9. INTENSIFIED CCD

The intensified CCD (ICCD) is more difficult to analyze. The MTF of an image intensifier tube (IIT) varies with tube type,[28] manufacturing precision,[29] and materials used. Figure 10-32 illustrates a possible range of MTFs. The MTF should be obtained from the manufacturer before proceeding with system analysis. It can be approximated by a simple Gaussian expression:

$$MTF_{INTENSIFIER}(u) = e^{-2\pi^2 \sigma_{IIT}^2 u^2} . \tag{10-55}$$

MTF theory will provide an optimistic result of fiber optic coupled ICCDs. It does not include the CCD mismatch or the chicken wire effects. The MTF associated with lens coupling is probably more accurate. The MTF of an ICCD is $MTF_{INTENSIFIER}$ muitiplied by the appropriate MTFs presented in this chapter.

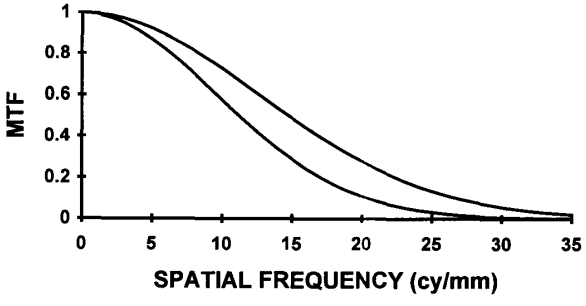


Figure 10-32. Range of MTFs for GenII image intensifier (From Reference 29).

10.10. SAMPLING EFFECTS

For a sampled-data system, a wide range of MTF values is possible[30-34] for any given spatial frequency (see Section 8.3.3., *Image Distortion*, page 245). To account for this variation, a sample-scene MTF may be included. The detector spatial frequency response is represented by $MTF_{DETECTOR}MTF_{PHASE}$. This MTF provides an average performance response that may be used to calculate the system response general imagery.

In general, the "MTF" is[35]

$$MTF_{PHASE}(u) \approx \cos\left(\frac{u_i}{u_{iN}}\,\theta\right), \tag{10-56}$$

where θ is the phase angle between the target and the sampling lattice. For example, when $u_i = u_{iN}$, the MTF is a maximum when $\theta = 0$ (in-phase) and a zero when $\theta = \pi/2$ (out-of-phase).

To approximate a median value for phasing, θ is set to $\pi/4$. Here, approximately one-half of the time the MTF will be higher and one-half of the time the MTF will be lower. The median sampling MTF is

$$MTF_{MEDIAN}(u) \approx \cos\left(\frac{\pi}{2}\,\frac{u_i}{u_{iS}}\right). \tag{10-57}$$

At Nyquist frequency, MTF_{PHASE} is 0.707. This is the Kell factor so often reported when specifying the resolution of monitors. An MTF averaged over all phases is represented by

$$MTF_{AVERAGE}(u_i) \approx sinc\left(\frac{u_i}{u_{iS}}\right). \tag{10-58}$$

Figure 10-33 illustrates the difference between the two equations. Because these are approximations, they may be considered roughly equal over the range of interest (zero to the Nyquist frequency). MTF_{PHASE} can be applied to both the vertical and horizontal MTFs.

Sampling effects are phase-dependent and, as a result, have no unique MTF. For system modeling purposes, average MTFs were assigned for some more common image processing algorithms. Because average MTFs are used, only an average system performance can be predicted. Performance can vary dramatically depending upon the target phase with respect to the sampling lattice. For laboratory measurements, it is common practice to "peak-up" the target (in-phase relationship) and then $MTF_{PHASE} = 1$. This is done to get repeatable results.

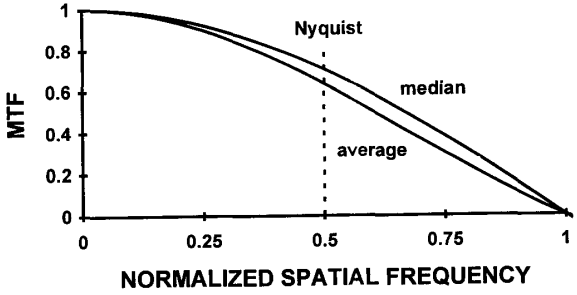


Figure 10-33. Average and median scene-sample phase MTFs normalized to u/u_S. The MTF is defined only up to the Nyquist frequency.

Frequencies above u_{iN} are aliased down to lower frequencies. Signal aliasing is ignored in most analyses and has been ignored here. A complete end-to-end analysis should include both aliased noise and aliased signal[36]. These aliased components limit[37] the extent to which a sampled image can be sharpened. Image reconstruction may include data removal (decimation) or data expansion (interpolation). The effective sampling frequency and aliasing can appear to change using different image processing algorithms.[38-40]

10.11. REFERENCES

1. See, for example, R. J. Farrell and J. M. Booth, *Design Handbook for Imagery Interpretation Equipment*, Reprinted with corrections, Report D180-19063-1, Boeing Aerospace Company, Seattle, WA, February 1984.
2. G. C. Holst, *Testing and Evaluation of Infrared Imaging Systems*, pp. 106-108, JCD Publishing, Winter Park, Fl (1993).
3. G. C. Holst, *Testing and Evaluation of Infrared Imaging Systems*, pp. 195-200, JCD Publishing, Winter Park, Fl (1993).
4. G. C. Holst, *Testing and Evaluation of Infrared Imaging Systems*, pp. 104-105, JCD Publishing, Winter Park, Fl (1993).

5. H. H. Hosack, "Aperture Response and Optical Performance of Patterned-Electrode Virtual-Phase Imagers," *IEEE Transactions on Electron Devices*, Vol. ED-28(1) pp. 53-63 (1981).
6. D. H. Sieb, "Carrier Diffusion Degradation of Modulation Transfer Function in Charge Coupled Imagers," *IEEE Transactions on Electron Devices*, Vol. ED-21(5), pp. 210-217 (1974).
7. L. W. Schumann and T. S. Lomheim, "Modulation Transfer Function and Quantum Efficiency Correlation at Long Wavelengths (Greater Than 800 nm) in Linear Charge Coupled Imagers," *Applied Optics*, Vol. 28(9), pp. 1701-1709 (1989).
8. E. L. Dereniak and D. G. Crowe, *Optical Radiation Detectors*, pp. 199-203, John Wiley & Sons, New York (1984).
9. H. V. Kennedy, "Miscellaneous Modulation Transfer Function (MTF) Effects Relating to Sampling Summing," in *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing*, G. C. Holst, ed., SPIE Proceedings Vol. 1488, pp. 165-176 (1991).
10. H.-S. Wong. Y. L. Yao, and E. S. Schlig, "TDI Charge-coupled Devices: Design and Applications," *IBM Journal of Research and Development*, Vol. 36(1), pp. 83-105 (1992).
11. J. E. Greivenkamp, "Color Dependent Optical Prefilter for the Suppression of Aliasing Artifacts," *Applied Optics*, Vol. 29(5), pp 676-684 (1990).
12. G. C. Holst, *Sampling, Aliasing, and Data Fidelity*, pp. 308-311, JCD Publishing, Winter Park, FL (1998).
13. G. C. Holst, *Sampling, Aliasing, and Data Fidelity*, pp. 259-268, JCD Publishing, Winter Park, FL (1998).
14. *System Image Analyzer* is available from JCD Publishing, Winter Park, FL (1995).
15. G. L. Conrad, "Reconnaissance System Performance Predictions Through Image Processing," in *Airborne Reconnaissance XIV*, P A. Henkel, F. R. LaGesse, and W. W. Schurter, eds., SPIE Proceedings Vol. 1342, pp. 138-145 (1990).
16. A variety of texts on digital filter design are available. See, for example, *Digital Signal Processing*, A. V. Oppenheim and R. W. Schafer, Prentice-Hall, New Jersey (1975).
17. G. C. Holst, *Electro-Optical System Performance*, pp. 146-148, JCD Publishing, Winter Park, FL (1995).
18. R. J. Farrell and J. M. Booth, *Design Handbook for Imagery Interpretation Equipment*, Reprinted with corrections, Report D180-19063-1, Boeing Aerospace Company, Seattle Wash, February 1984.
19. L. M. Biberman, ed., *Perception of Displayed Information*, Plenum Press, New York (1973).
20. B. O. Hultgren, "Subjective Quality Factor Revisited," in *Human Vision and Electronic Imaging: Models, Methods, and Applications*, B. E. Rogowitz and J. P. Allebach, eds., SPIE Proceedings Vol. 1249, pp. 12-22 (1990).
21. N. Nill, "A Visual Model Weighted Cosine Transform for Image Compression and Quality Measurements," *IEEE Transactions on Communications*, Vol. 33(6), pp. 551-557 (1985).
22. T. J. Schulze, "A Procedure for Calculating the Resolution of Electro-Optical Systems," in *Airborne Reconnaissance XIV*, P A. Henkel, F. R. LaGesse, and W. W. Schurter, eds., SPIE Proceedings Vol. 1342, pp. 317-327 (1990).
23. F. W. Campbell and J. G. Robson, "Application of Fourier Analysis to the Visibility of Gratings," *Journal of Physiology*, Vol. 197, pp. 551-566 (1968).
24. A. N. deJong and S. J. M. Bakker, "Fast and Objective MRTD Measurements," in *Infrared Systems - Design and Testing*, P. R. Hall and J. S. Seeley, eds., SPIE Proceedings Vol. 916, pp. 127-143 (1988).
25. G. H. Kornfeld and W. R. Lawson, "Visual Perception Model," *Journal of the Optical Society of America*, Vol. 61(6), pp. 811-820 (1971).
26. J. Ratches, W. R. Lawson, L. P. Obert, R. J. Bergemann, T. W. Cassidy, and J. M. Swenson, "Night Vision Laboratory Static Performance Model for Thermal Viewing Systems," U.S. Army Electronics Command Report ECOM Report 7043, p. 11, Ft. Monmouth, NJ (1975).

27. *FLIR92 Thermal Imaging Systems Performance Model*, Analyst's Reference Guide, Document RG5008993, Ft. Belvoir, VA, January 1993.
28. G. M. Williams, Jr., "A High Performance LLLTV CCD Camera for Nighttime Pilotage," in *Electron Tubes and Image Intensifiers*, C. B. Johnson and B. N. Laprade, eds., SPIE Proceedings Vol. 1655, pp. 14-32 (1992).
29. I. P. Csorba, *Image Tubes*, pp. 79-103, Howard W. Sams, Indianapolis, IN (1985).
30. S. K. Park and R. A. Schowengerdt, "Image Sampling, Reconstruction and the Effect of Sample-scene Phasing," *Applied Optics*, Vol. 21(17). pp. 3142-3151 (1982).
31. J. C. Feltz and M. A. Karim, "Modulation Transfer Function of Charge-coupled Devices," *Applied Optics*, Vol. 29(5), pp. 717-722 (1990).
32. J. C. Feltz, "Development of the Modulation Transfer Function and Contrast Transfer Function for Discrete Systems, Particularly Charge-coupled Devices," *Optical Engineering*, Vol. 29(8), pp. 893-904 (1990).
34. L. de Luca and G. Cardone, "Modulation Transfer Function Cascade Model for a Sampled IR Imaging System," *Applied Optics*, Vol. 30(13), pp. 1659-1664 (1991).
35. F. A. Rosell, "Effects of Image Sampling," in *The Fundamentals of Thermal Imaging Systems*, F. Rosell and G. Harvey, eds., p. 217, NRL Report 8311, Naval Research Laboratory, Washington D.C. (1979).
36. S. K. Park, "Image Gathering, Interpolation, and Restoration: A Fidelity Analysis," in *Visual Information Processing*, F. O. Huck and R. D. Juday, eds., SPIE Proceedings Vol. 1705, pp. 134-144 (1992).
37. S. K. Park and R. Hazra, "Image Restoration Versus Aliased Noise Enhancement," in *Visual Information Processing III*, F. O. Huck and R. D. Juday, eds., SPIE Proceedings Vol. 2239, pp. 52-62 (1994).
38. I. Ajewole, "A Comparative Analysis of Digital Filters for Image Decimation and Interpolation," in *Applications of Digital Image Processing VIII*, A. G. Tescher, ed., SPIE Proceedings Vol. 575, pp. 2-12 (1985).
38. W. F. Schrieber and D. E. Troxel, "Transformation Between Continuous and Discrete Representations of Images: A Perceptual Approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMA-7(2), pp. 178-186 (1985).
39. R. W. Schafer and L. R. Rabiner, "A Digital Signal Processing Approach to Interpolation," *Proceedings of the IEEE*, Vol. 61(8), pp. 692-702 (1973).
40. A. H. Lettington and Q. H. Hong, "Interpolator for Infrared Images," *Optical Engineering*, Vol. 33(3), pp. 725-729 (1994).

11

IMAGE QUALITY

Our perception of good image quality is based upon the real-world experiences of seeing all colors, all intensities, and textures. An imaging system has a limited field-of-view, limited temporal and spatial resolutions, and presents a two-dimensional view of a three-dimensional world. In the real-world our eyes scan the entire scene. Not all of the available information is captured by an imaging system. Furthermore, an imaging system introduces noise and the loss of image quality due to noise can only be estimated. Cameras sensitive to wavelengths less than 0.4 μm and greater than 0.7 μm provide imagery that we cannot directly perceive. The quality of the ultraviolet or infrared imagery can only be estimated because we do not know how it really appears.

Image quality is a subjective impression ranging from poor to excellent. It is a somewhat learned ability. It is a perceptual one, accomplished by the brain, affected by and incorporating inputs from other sensory systems, emotions, learning, and memory. The relationships are many and are still not well understood. Perceptual quality of the same scene varies between individuals and temporally for the same individual. Large variations exist in an observer's judgment as to the correct rank ordering of image quality from poor to best and therefore image quality cannot be placed on an absolute scale. Visual psychophysical investigations have not measured all the properties relevant to imaging systems.

Many formulas exist for predicting image quality. Each is appropriate under a particular set of viewing conditions. These expressions are typically obtained from empirical data in which multiple observers view many images with a known amount of degradation. The observers rank the imagery from worst to best and then an equation is derived which relates the ranking scale to the amount of degradation.

If the only metric for image quality was resolution, then we would attempt to maximize resolution in our system design. Many tests have provided insight into image quality metrics. In general, images with higher MTFs and less noise are judged as having *better* image quality. There is no single *ideal* MTF shape that provides best image quality. For example, Kusaka[1] showed that the MTF that produced the most aesthetically pleasing images depended on the scene content.

The metrics suggested by Granger and Cupery, Shade, and Barten offer[2-6] additional insight on how to optimize an imaging system. Granger and Cupery developed the Subjective Quality Factor (SQF): an empirically derived relationship using individuals' responses when viewing many photographs. Shade used photographs and included high quality TV images. Barten's approach is more comprehensive in that it includes a variety of display parameters. It now includes[7] contrast, luminance, viewing ratio, number of scan lines, and noise.

While Barten has incorporated the sampling effects of flat panel displays,[7] no model includes the sampling (and associated aliasing) that takes place at the detector. Until this aliasing is quantified, no metric will predict image quality for sampled data systems. Only an end-to-end system approach can infer overall image quality.

There are potentially two different system design requirements: (1) good image quality and (2) performing a specific task. Sometimes these are equivalent, other times they are not. All image quality metrics incorporate some form of the system MTF. The underlying assumption is that the image spectrum is limited by the system MTF. Equivalently, it is assumed that the scene contains all spatial frequencies and that the displayed image is limited by system MTF - a reasonable assumption for general imagery. While good image quality is always desired, a military system is designed to detect and recognize specific targets (discussed in Chapter 12, *Minimum Resolvable Contrast*). Optimized military systems will have high MTF at the (assumed) target frequencies and other spatial frequencies are considered less important. Computer monitors are usually designed for alphanumeric legibility.

Often, the display is the limiting factor in terms of image quality and resolution. No matter how good the electronic imaging system is, if the display resolution is poor, then the overall system resolution is poor. If the system is optics-limited, then a high resolution display does not offer any "extra" system resolution. A high resolution display just ensures that all the information available is displayed. System resolution may be limited by the HVS. If the observer is too far from the screen, not all of the image detail can be discerned.

The HVS appears to operate as a tuned spatial-temporal filter where the tuning varies according to the task at hand. Because the HVS approximates an optimum filter, no system performance improvement is expected by *precisely* matching the image spectrum to the HVS preferred spectrum. Rather, if the displayed spectrum is within the limits of the eye spectrum, the HVS will automatically tune to the image. Clearly, the overall system magnification should be set such that the frequency of the maximum interest coincides with the peak

frequency of the HVS's contrast sensitivity. Although this implies a specific frequency, the range of optimization is broad.

A large number of metrics are related to image quality. Most are based on monochrome imagery such as resolution, MTF, and minimum resolvable contrast. CRT resolution was presented in Section 7.5., *Resolution* (page 218). Color reproduction and tonal transfer issues, which are very important to color cameras, are not covered here.

The symbols used in this book are summarized in the *Symbol List* (page xviii) which appears after the *Table of Contents*.

11.1. RESOLUTION METRICS

An overwhelming majority of image quality discussions center on resolution. Resolution has been in use so long that it is thought to be fundamental and that it uniquely determines system performance. There are four different types of resolution: (1) temporal, which is the ability to separate events in time; (2) gray scale, which is determined by the analog-to-digital converter design, noise floor; or the monitor capability, (3) spectral; and (4) spatial. An imaging system operating at 30 Hz frame rate has a temporal resolution of 1/30 sec. Gray-scale resolution is a measure of the dynamic range. The spectral resolution is simply the spectral band pass (e.g., UV, visible, or near infrared) of the system. This section covers spatial resolution.

Resolution provides valuable information regarding the finest spatial detail that can be discerned. A large variety of resolution measures exist[8] and the various definitions may not be interchangeable. An electronic imaging system is composed of many subsystems and each has its own metric for resolution (Table 11-1). These can also be divided into analog (Table 11-2) and sampled-data resolution metrics (Table 11-3).

A single measure of spatial resolution cannot be satisfactorily used to compare all sensor systems. Spatial resolution does not provide information about total imaging capability or contrast sensitivity. The minimum resolvable contrast simultaneously provides contrast sensitivity information and resolution information for monochrome systems. Resolution does not include the effects of system noise and is not related to sensitivity.

Table 11-1
SUBSYSTEM MEASURES of RESOLUTION

SUBSYSTEM	RESOLUTION METRIC
Optics	Rayleigh criterion Airy disk diameter Blur diameter
Detectors	Detector-angular-subtense Instantaneous-field-of-view Effective-instantaneous-field-of-view Detector pitch Pixel-angular-subtense
Electronics	Bandwidth
Electronic imaging system (MTF approach)	Limiting resolution Nyquist frequency
Displays	TV limiting resolution

Table 11-2
RESOLUTION MEASURES for ANALOG SYSTEMS

RESOLUTION	DESCRIPTION
Rayleigh criterion	Ability to distinguish two adjacent point sources
Airy disk	Diffraction-limited diameter produced by a point source
Blur diameter	Actual minimum diameter produced by a point source
Limiting Resolution	Spatial frequency at which MTF= 0.02 to 0.10
TV limiting resolution	Number of resolved lines per picture height
Ground resolved distance	The smallest test target (1 cycle) that a photointerpreter can distinguish
Ground resolution	An estimate of the limiting feature size seen by a photointerpreter
Resel	Smallest region that contains unique information

Table 11-3
RESOLUTION MEASURES for SAMPLED DATA SYSTEMS

RESOLUTION	DESCRIPTION
Detector angular subtense	Angle subtended by one detector element
Instantaneous field-of-view	Angular region over which the detector senses radiation
Nyquist frequency	One-half of the sampling frequency
Detector pitch	Center-to-center spacing
Effective-instantaneous-field-of-view	One-half of the reciprocal of the object-space spatial frequency at which the MTF is 0.5
Pixels, datels, and disels	Number of detector elements or number of digital data points

When resolution is defined as the inverse of the system cutoff in angular space (typical for military applications), the back-of-the-envelope approximation provides the range at which the target can be "detected":

$$Range \sim \frac{target\ size}{resolution} . \qquad (11\text{-}1)$$

This is the maximum range at which a periodic target can be faithfully reproduced.

Analog measures of spatial resolution may be determined by: (1) the width of a point source image, (2) the minimum detectable separation of two point sources, (3) the spatial frequency at which the MTF drops to a defined level, or (4) the smallest detail resolved by an observer. These measures assume that the system output image is a replica of the input scene (a linear-shift-invariant system).

As imaging systems are incorporated into new disciplines, resolution must be specified in terms used by those industries. For example, photointerpreters use the ground resolved distance (GRD) when evaluating reconnaissance imagery captured on film. But as solid state imaging systems replace the wet film process, the GRD, by default, becomes an imaging system resolution measure.

When viewing aerial imagery that contains a test pattern (such as the U.S. Air Force 1951 standard 3-bar target), an image analyst determines the smallest discernible cycle on the ground. This cycle width (bar plus space) is the GRD. It includes the system MTF and possible degradation by the atmosphere and is

a function of altitude. The GRD is related[9,10] to the 10-point Imagery Interpretability Rating Scale (IIRS). The National Imagery Interpretability Rating Scale (NIIRS) has replaced the IIRS.[11-14]

The smallest elements that can be created by a digital subsystem are the pixel, datel, and disel. These are created by the detector array, ADC, and the display media, respectively. The resel is the smallest element that can be created by an analog system. The optics, analog electronics, and video standard (or transmission link) each have their own resel. System resolution is a convolution of all "-els."

Complex systems cannot be characterized by a single number (for example, an "-el"). Nevertheless, this chapter discusses various measures because they have been in existence so long. They were developed to describe the resolution of specific subsystems and do not provide an end-to-end performance metric. Nor do they include aliasing. Aliasing effects are discussed in Reference 15.

11.2. OPTICAL RESOLUTION METRICS

Table 11-4 provides various optical resolutions. As analog metrics, these are various definitions of a resel. Diffraction measures include the Rayleigh criterion and the Airy disk diameter. The Airy disk is the bright center of the diffraction pattern produced by an ideal optical system. In the focal plane of the lens, the Airy disk diameter is

$$d_{AIRY} = 2.44\frac{\lambda}{D_o}fl = 2.44\lambda F \; . \qquad (11\text{-}2)$$

The Rayleigh criterion is a measure of the ability to distinguish two closely spaced objects when the objects are point sources. Optical aberrations and focus limitations increase the diffraction-limited spot diameter to the blur diameter. Optical designers use ray tracing programs to calculate the blur diameter. The blur diameter size is dependent on how it is specified (i.e., the fraction of encircled energy[16]). den Dekker and van den Bos provide[17] an in-depth review of optical resolution metrics.

Considerable literature[18] has been written on the image forming capability of lens systems. Image quality metrics include aberrations, Strehl ratio, and blur or spot diagrams. These metrics, in one form or another, compare the actual blur diagram to diffraction-limited spot size. As the blur diameter increases, image quality decreases and edges become fuzzy. The ability to see this

degradation requires a high resolution sensor. Both the human eye and photographic film visually have this resolution. However, with electronic imaging systems, the detectors are often too large to see the degradation because the detectors are often larger than the blur diameter.

Table 11-4
OPTICAL RESOLUTION METRICS (in image space)

RESOLUTION	DESCRIPTION	DEFINITION (usual units)
Rayleigh criterion	Ability to distinguish two adjacent point sources	$x_{RES} = 1.22 \lambda F$ (Calculated: mm)
Airy disk	Diffraction-limited diameter produced by a point source	$x_{RES} = 2.44 \lambda F$ (Calculated: mm)
Blur diameter	Actual minimum diameter produced by a point source	Calculated from ray tracing (mm)

11.3. DETECTOR RESOLUTION

Considering the Airy disk, the wavelength dependence suggests that the resolution increases as the wavelength decreases (smaller is better). However, this is only true when the detector size is much smaller than the Airy disk size. For most electronic imaging systems the reverse is true. The detector is the same size or larger than the disk diameter (see Section 10.4., *Optics-Detector Subsystem*, page 286). This means that the Airy disk is not adequately sampled and therefore the resolution afforded by the optics cannot be exploited. Systems designed for star detection are usually optics-limited (optical resel larger than a pixel), whereas for general imagery the systems are usually detector-limited (pixel larger than the optical resel)

Detector arrays are often specified by the number of pixels and detector pitch. These are not meaningful until an optical system is placed in front of the array. Table 11-5 provides the most common resolution metrics expressed in object-space units. The detector-angular-subtense (DAS) is often used by the military to describe the resolution of systems when the detector is the limiting subsystem. If the detector's element horizontal and vertical dimensions are different, then the DAS in the two directions is different. Note that the PAS is different than the DAS (see Section 1.5., *Pixels, Datels, Disels, and Resels*, page 14). Only with 100% fill factor arrays are they the same. Spatial sampling rates are determined by the pitch in image space or the PAS in object space.

Table 11-5
DETECTOR ELEMENT RESOLUTION MEASURES (in object space)

RESOLUTION	DESCRIPTION	DEFINITION (usual units)
Detector-angular-subtense	Angle subtended by one detector element	d/fl (Calculated: mrad)
Instantaneous-field-of-view	Angular region over which the detector senses radiation	Measured width at 50% amplitude (mrad)
Pixel-angular-subtense	Angle subtended by one pixel	d_{cc}/fl (Calculated: mrad)
Effective-instantaneous-field-of-view	One-half of the reciprocal of the object-space spatial frequency at which the MTF is 0.5	Measured or calculated (mrad)
Ground sampled distance	Projection of detector pitch onto the ground	d_{cc}R/fl (calculated: inches)
Nyquist frequency	One-half of the angle subtended by detector pitch	d_{cc}/2fl (calculated: mrad)

If resolution is defined by the DAS, then a vanishingly small detector is desired. Simultaneously, the f-number must approach zero so the system is not optics-limited (see Figure 10-15, page 287). Although any detector size, optical aperture, and focal length can be chosen to select the resolution limit, the same parameters affect the sensitivity (see Equation 10-27, page 288). With practical optical systems, the minimum f-number is about 1.0 with a theoretical limit of 0.5. Resolution and sensitivity factors are coupled into a single figure of merit: the minimum resolvable contrast (discussed in Chapter 12, *Minimum Resolvable Contrast*).

The instantaneous-field-of-view (IFOV) is the linear angle from which a single detector element senses radiation. It is a summary measure that includes both the optical and detector responses. The IFOV is typically a measured quantity. Here, a point source transverses the detector element and the detector output is graphed as a function of angle. The IFOV is the full-width, one-half maximum amplitude of the resultant signal. For detector-limited systems (large $d/F\lambda$), the IFOV is approximately equal to the DAS. Equivalently, if the optical blur diameter is small compared to the DAS, then they are approximately equal.

The detector horizontal cutoff is $u_{oD} = fl/d_H = 1/DAS$ in object space. With detector arrays, signal fidelity is limited to the Nyquist frequency. Here, the effective-instantaneous-field-of-view (EIFOV) offers an alternate measure of resolution (Figure 11-1). The value u_{EIFOV} is the spatial frequency at which the MTF is 0.5. Calling 1/DAS the cutoff provides a misleading representation. Recall that the detector can respond to frequencies above 1/DAS (see Figure 8-9, page 241).

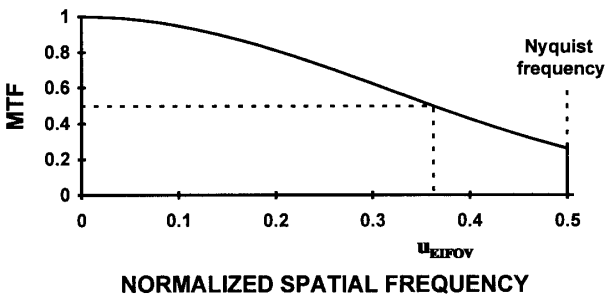


Figure 11-1. Definition of effective-instantaneous-field-of-view for an undersampled system. The spatial frequency has been normalized to the Nyquist frequency, f_N. EIFOV = $1/(2\,u_{EIFOV})$.

11.4. ELECTRICAL RESOLUTION METRIC

For high data rate systems such as line-scanners, the electronic bandwidth may limit system response. For electronic circuits, resolution is implied by its bandwidth. The minimum pulse width is approximately

$$\tau_{MINIMUM} \approx \frac{1}{2\,BW}, \qquad (11\text{-}3)$$

where $\tau_{MINIMUM}$ is measured in seconds and BW is measured in Hertz. The 3-dB point (half-power frequency) is often called the bandwidth. With this minimum value the output signal *looks* like the input. Other expressions (e.g., $\tau_{MINIMUM}$ = 1/BW) simply imply there is a measurable output without any inference about signal fidelity.

11.5. MTF-BASED RESOLUTION

Usually systems are designed to have high MTF, and Figure 10-14 (page 286) illustrated the MTF at Nyquist frequency. However, the value of Nyquist frequency is a measure of resolution; 100% fill factor staring arrays have the highest Nyquist frequency.

11.5.1. LIMITING RESOLUTION

Other measures also exist such as the limiting resolution. This is the spatial frequency at which the MTF is, say, 5%. Figure 11-2 illustrates two systems that have the same limiting resolution. Which system is selected depends on the specific application. System A is better at high spatial frequencies, and system B is better at low spatial frequencies. Equivalently, if the scene contains mostly low frequencies (large objects), then system B should be used. If edge detection is important (high spatial frequencies), then system A should be considered. This clearly indicates the difficulty encountered when using resolution exclusively as a measure of system performance.

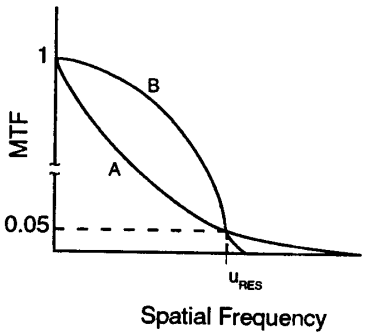


Figure 11-2. The MTFs of two different imaging systems with the same limiting resolution, u_{RES}.

11.5.2. MTF/TM RESOLUTION

Early resolution measures were based on wet-film camera characteristics. The intersection of the system MTF and film threshold modulation (TM) provided a measure of camera system performance (Figure 11-3). TM is the minimum modulation required at a given spatial frequency to discern a difference. The value u_{RES} is the limiting spatial frequency or intersection frequency.

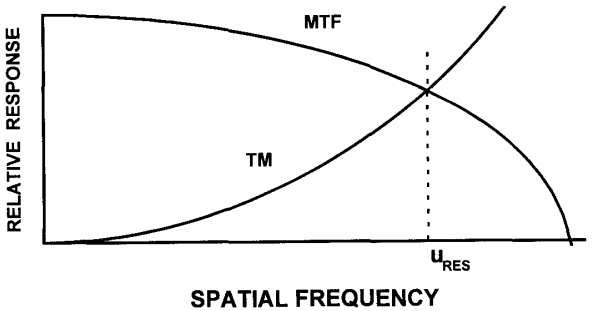


Figure 11-3. The spatial frequency of the intersection of the film threshold modulation and MTF_{SYSTEM} is a measure of resolution.

Conceptually, for spatial frequencies above u_{RES}, the film requires more modulation than that provided by the system. Below u_{RES} the film can easily discern the scene modulation. Because many measures were in support of aerial imagery, the intersection was also called the aerial image modulation (AIM). With the advent of electronic imaging systems, the film TM was replaced with the HVS's threshold modulation (discussed in Section 11.7., *MTFA*).

11.5.3. SHADE'S EQUIVALENT RESOLUTION

Shade[3] equated apparent image sharpness of a television picture to an equivalent pass band:

$$N_e = \int_0^\infty [MTF(f)]^2 \, df \, . \qquad (11\text{-}4)$$

As reported by Lloyd,[19] Sendall modified Shade's equivalent resolution such that

$$R_{EQ} = \frac{1}{2N_e} = \frac{1}{2\int_0^{\infty} [MTF_{SYS}(u)]^2 \, du} . \tag{11-5}$$

The value R_{EQ} cannot be directly measured and is a mathematical construct that expresses overall performance. As R_{EQ} decreases, the resolution "improves" (smaller is better). As an approximation, the system resolution, $R_{EQ\text{-}SYS}$, may be estimated from the component equivalent resolutions, R_i, by

$$R_{EQ\text{-}SYS} \approx \sqrt{R_1^2 + R_2^2 + \ldots + R_n^2} . \tag{11-6}$$

Shade's approach using the square of the MTF emphasized those spatial frequencies at which the MTF is relatively high. It appears to be a good measure for classical systems in which the MTF is decreasing (such as a Gaussian distribution). The equivalent resolution approach assumes that the system is completely analog and it ignores sampling effects. Therefore, R_{EQ} becomes a resel. It probably should not be used (nor should any other image quality metric) to compare systems with significantly different MTFs. Nor should it be used with systems that can create significant aliasing.

As a summary metric, $R_{EQ\text{-}SYS}$ provides a better indication of system performance than just a single metric such as the DAS. $R_{EQ\text{-}SYS}$ probably should be used in Equation 11-1 (page 319) to obtain a more realistic measure of range performance. The value R_{EQ} cannot be evaluated in closed form for all subsystems such as boost, defocused optics, or Chebyshev filters. For these subsystems, Equation 11-6 must be evaluated numerically. If a subsystem MTF is essentially unity over the spatial frequency region of interest, then that subsystem R_{EQ} can be ignored (e.g., $R_{EQ} = 0$). Table 11-6 provides the equivalent resolution for several common MTFs.

Table 11-6
ONE-DIMENSIONAL EQUIVALENT RESOLUTIONS

SUBSYSTEM	MTF	R_{EQ}
Optics	$\frac{2}{\pi}\left(\cos^{-1}\left(\frac{u_i}{u_{iC}}\right)-\frac{u_i}{u_{iC}}\sqrt{1-\left(\frac{u_i}{u_{iC}}\right)^2}\right)$	$1.845\,\lambda F$
Rectangular detector	$sinc(d\,u_i)$	d
N^{th}-order low-pass filter	$\frac{1}{\sqrt{1+\left(\frac{f_e}{f_{e3dB}}\right)^{2N}}}$	$\frac{1}{\pi f_{e3dB}}$ *when* $N = 1$ $\frac{1}{2f_{e3dB}}$ *when* $N \to \infty$
CRT-based display	$e^{-2\pi^2\sigma_{SPOT}^2 u_d^2}$	$2\sqrt{\pi}\,\sigma_{SPOT}$

If the electronics provides adequate bandwidth, it is instructive to see the relationship between the Airy disk diameter and the detector size. Using the same parameters as Figure 10-14 (page 286), $x = d/(F\lambda)$, for a system with resolution limited by the optics/detector combination,

$$R_{EQ\text{-}SYS} = d\sqrt{\left(\frac{1.845}{x}\right)^2+1}\ . \qquad (11\text{-}7)$$

The equation is graphed in Figure 11-4. As x increases, $R_{EQ\text{-}SYS}$ approaches d. For small values of x, the system becomes optics-limited and the equivalent resolution increases. When $x = 2.44$, $R_{EQ\text{-}SYS} = 1.25d$ and when $x = 4$, $R_{EQ\text{-}SYS} = 1.10d$. Matching the Airy disk to the detector size creates an equivalent resolution 25% larger than the actual detector size.

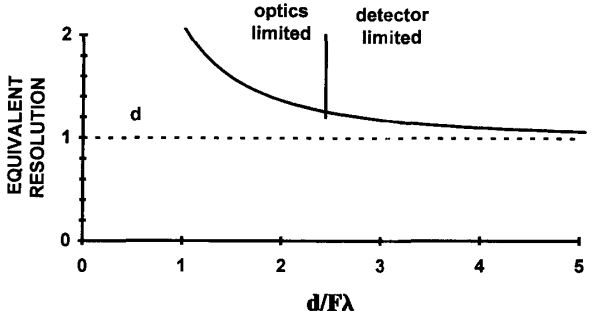


Figure 11-4. Equivalent resolution normalized to d as a function of d/Fλ. The vertical line indicates where the Airy disk diameter is equal to the detector size. The best resolution occurs when d/Fλ approaches infinity. Here the optics MTF approaches unity.

11.6. SYSTEM RESOLUTION EXAMPLES

Example 11-1
PIXELS and DATELS

An imaging system consists of 512 × 512 detector elements. Assume that the fill factor is 100% and the camera's analog output is digitized at either 1024 samples/line or 2048 samples/line. What is the resolution?

The external analog-to-digital converter creates either two or four datels for each pixel. In both cases, the pixel size remains the same. Increasing the sampling rate does not change the system's resolution but improves the repeatability of all measurements by reducing phasing effects that might occur in the analog-to-digital converter. Any image processing algorithm that operates on this higher number must consider the sensor resolution.

Cameras with a standard video output may not have the resolution suggested by the video standard. For example, a camera containing an array of 320 × 240 pixels may have its output formatted into EIA 170 timing. The signal may be digitized by a frame grabber that creates 640 × 480 datels. The image size is still 320 × 240 pixels.

Example 11-2
DISELS and PIXELS

A staring array consists of 512 × 512 detectors. Using 4× electronic zoom it is presented on a digital monitor that displays 1024 × 1024 disels. What is the system resolution?

Each pixel is mapped onto four datels that are then mapped one-to-one onto disels. Here, the imaging system determines the system resolution, not the number of datels or disels. High quality monitors only ensure that the image quality is not degraded. Electronic zoom cannot increase resolution. But decimation or minifying may reduce resolution. Because image processing is performed on datels, the image analyst must be made aware of the system resolution.

Example 11-3
SYSTEM EQUIVALENT RESOLUTION

A CCD camera has an analog output digitized by a frame grabber. What is the apparent system resolution for image processing and what is the resolution displayed on a CRT-based monitor?

This CCD camera contains 480 horizontal detectors. Each detector is 15 μm on 15 μm centers (100% fill factor). The optical aperture is 1 inch and the focal length is 5.6 inches (F = 5.6). The analog electronics is wideband and it does not significantly affect the resolution. Camera output is in the standard EIA 170 format with a bandwidth of 4.2 MHz and $t_{VIDEO\text{-}LINE}$ = 52.45 μs. The frame grabber provides 640 horizontal samples and a standard consumer CRT-based monitor is used.

Assuming $\lambda_{AVE} \approx 0.5\,\mu m$, R_{EQ} for the optics is $1.845\lambda_{AVE}/D_o = 0.0363$ mrad. The DAS is d/fl or 0.105 mrad. Here, $d/(F\lambda) = \alpha D_o/\lambda = 5.33$. That is, the optical/detector subsystem is detector-limited. The HFOV is

$$HFOV = \frac{N_H d_H}{fl} = \frac{480\,(15\times10^{-6}\,m)}{(5.6\,in)\left(25.4\times10^{-3}\frac{m}{in}\right)} = 50.6\,mrad\,. \quad (11\text{-}8)$$

With an active line time of 52.45 μs, the video bandwidth in object space is

$$f_{o\text{-}VIDEO} = \frac{t_{VIDEO\text{-}LINE}}{HFOV}BW = \frac{52.45\times10^{-6}}{50.6\times10^{-3}}4.2\times10^{6} = 4.35\,\frac{cycles}{mrad}\,. \quad (11\text{-}9)$$

Assuming a sharp cutoff, $R_{EQ} \approx 1/2f_{3dB} = 0.115$ mrad. The system equivalent resolution is

$$R_{EQ\text{-}SYS} = \sqrt{0.0363^2 + 0.105^2 + 0.115^2} = 0.160\,mrad\,. \quad (11\text{-}10)$$

Here, the EIA 170 bandwidth significantly affects the resolution. Neglecting phasing effects, each frame grabber datel represents approximately (0.159)(480)/640 = 0.119 mrad. These are overlapping resolution elements where the center-to-center spacing is (0.105)(480)/640 = 0.0787 mrad.

The CRT equivalent resolution is

$$R_{EQ} = 2\sqrt{\pi}\,\sigma_o = 2\sqrt{\pi}\left(\frac{d_{CC}}{1.645\cdot fl}\right) = 0.227\,mrad\,. \quad (11\text{-}11)$$

The display significantly degrades the resolution. But this is necessary to suppress the raster and sampling artifacts. Each disel represents overlapping resolution elements with apparent size

$$R_{EQ\text{-}SYS} = \sqrt{0.159^2 + 0.227^2} = 0.277\,mrad\,. \quad (11\text{-}12)$$

Numerical integration of Equation 11-5 provides $R_{EQ\text{-}SYS} = 0.263$. For this example, Equations 11-5 and 11-6 provided similar results.

11.7. MTFA

The area bounded by the system MTF and the HVS's modulation threshold (M_t) is called the modulation transfer function area (MTFA) (Figure 11-5a). As the MTFA increases, the perceived image quality appears to increase. The spatial frequency at which the MTF and M_t intersect is a measure of the overall resolution. However, the resolution is not unique because M_t depends on the display viewing distance. The HVS's inhibitory response (see Figure 10-31, page 309) is omitted from the MTFA approach. Mathematically,

$$MTFA = \int_0^{f_o} [MTF(f) - M_t(f)]\, df \,. \qquad (11\text{-}13)$$

where f_o is the limiting spatial frequency or intersection frequency. The value M_t is also called the aerial image modulation and demand modulation function. It is the minimum modulation required at a given spatial frequency to discern a difference.

According to Snyder,[20] the MTFA appears to correspond well with performance in military detection tasks where the targets are embedded in noise. The noise elevates the HVS's modulation threshold so that the MTFA decreases with increasing noise (Figure 11-5b).

For low noise, general imagery, the HVS's contrast sensitivity has little effect on the MTFA. The integrand in Equation 11-13 can be replaced with just the system MTF. Here, the high spatial frequencies are weighted too heavily compared to the low spatial frequencies and Shade's approach appears better.

From an analyst's viewpoint, the convolution theorem states that MTFs should be multiplied. The contrast sensitivity is proportional to the HVS MTF. Therefore, the system MTF should be multiplied by the contrast sensitivity or divided by the inverse (modulation threshold). However, the intersection frequency can be used as a measure of limiting resolution of the system. It can be used, for example, to calculate the GRD.

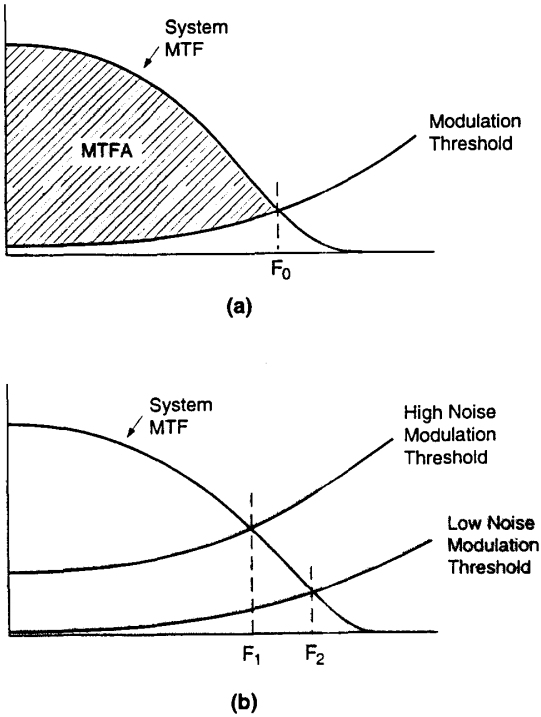


Figure 11-5. MTFA. The area between system MTF and M_t up to the intersection is a measure of image quality. The spatial frequency at the intersection is a measure of limiting resolution. The area and limiting resolution depend on the noise level within the imagery. (a) Low noise and (b) elevated noise.

11.8. SUBJECTIVE QUALITY FACTOR

According to Granger and Cupery,[2] the spatial frequency range important to image quality is in the region from just below the peak to three times the peak sensitivity of the eye. The Campbell-Robson high luminance data, and the HVS models of Nill, Schultz, and deJong and Bakker support Granger and Cupery's approach whereas the Kornfeld-Lawson HVS model does not (see Section 10.8., *Human Visual Response*, page 306). Recall that the Kornfeld-Lawson model is considered only applicable for spatial frequencies greater than f_{PEAK}.

In logarithmic units, the normalized subjective quality factor (SQF) is:

$$SQF = K \int_{\log(3)}^{\log(12)} MTF_{SYS}(f_{eye})\, d[\log(f_{eye})] \,, \qquad (11\text{-}14)$$

where K is a normalization constant. In the SQF approach, only those frequencies that are very important to the eye are included. The spatial frequency presented to the eye depends on the image size on the display, the distance to the display, and electronic zoom. Table 11-7 provides Granger and Cupery's interpretation of the SQF. These results are based on many observers viewing noiseless photographs with known MTF degradation.

The SQF is an adequate image quality metric for low noise imagery when the illumination is fixed at a moderate level. As with the MTFA and Shade's equivalent pass band, the SQF is intended for general imagery only. It cannot be used for specific applications such as the legibility of alphanumeric characters.

Table 11-7
SUBJECTIVE QUALITY FACTOR

SQF	SUBJECTIVE IMAGE QUALITY
0.92	Excellent
0.80	Good
0.75	Acceptable
0.50	Unsatisfactory
0.25	Unusable

11.9. SQUARE-ROOT INTEGRAL

Barten[4-7] introduced the square-root integral (SQRI) as a measure of image quality. The SQRI is

$$SQRI = \frac{1}{\ln(2)} \int_{f_{min}}^{f_{max}} \sqrt{MTF_{SYS}(f_{eye})\, CSF_{EYE}(f_{eye})}\, \frac{df_{eye}}{f_{eye}} . \quad (11\text{-}15)$$

The multiplication overcomes the theoretical objection raised with the MTFA where the CSF was subtracted from the MTF. MTF_{SYS} includes all the subsystem MTFs up to and including the display. Barten modeled the contrast sensitivity function as

$$CSF(f_{eye}) = \frac{1}{M_t} = a f_{eye} e^{-bf_{eye}} \sqrt{1 + c e^{bf_{eye}}} , \quad (11\text{-}16)$$

where

$$a = \frac{540\left(1 + \frac{0.7}{L_{AVE}}\right)^{-0.2}}{1 + \frac{12}{w\left(1 + \frac{f_{eye}}{3}\right)^2}} , \quad (11\text{-}17a)$$

$$b = 0.3\left(1 + \frac{100}{L_{AVE}}\right)^{0.15} , \quad (11\text{-}17b)$$

and

$$c = 0.06 , \quad (11\text{-}17c)$$

where f_{eye} has units of cycles/degree, L_{AVE} is the average luminance in candelas per square meter, and w is the angular display size as seen by the observer. The factor 1/ln(2) allows SQRI to be expressed in just-noticeable-difference (JND) units. Both average luminance level and display size affect the CSF. This CSF is a good approximation over five orders of magnitude of luminance levels and for angular display sizes ranging from 0.5° to 60°.

This model includes the effects of various display parameters such as viewing ratio, luminance, resolution, contrast, addressability, and noise. Display viewing ratio and average luminance are incorporated in the HVS's CSF (via w and L_{AVE} in the CSF equation). Display resolution (display MTF) is included in MTF_{SYS}.

Contrast loss due to reflected ambient lighting effectively reduces the display MTF. If ΔL is the additional luminance due to reflected ambient light, then the modulation depth is reduced by the factor:

$$\eta_{DISPLAY} = \frac{L_{AVE}}{L_{AVE} + \Delta L} . \qquad (11\text{-}18)$$

This factor is the same for all spatial frequencies of interest and the perceived effect on image quality is incorporated into the SQRI by multiplying MTF_{SYS} by $\sqrt{\eta_{DISPLAY}}$.

The display parameters dictate the limits of integration. The value f_{MIN} is that spatial frequency that provides one-half cycle across the display and f_{MAX} is the maximum spatial frequency of the display and this is related to display addressability (see Section 7.6., *Addressability*, page 222). For example, if the number of active lines is N_{ACTIVE}, then

$$f_{MAX} = \frac{(Kell\,factor)\ N_{ACTIVE}}{2} D_{FOV} . \qquad (11\text{-}19)$$

The small angle approximation was used to calculate field-of-view, D_{FOV}, subtended by the observer. It is the display height, $H_{MONITOR}$, divided by the viewing distance, D.

Noise increases the contrast sensitivity function. Because the SQRI must decrease with noise, Barten added noise as an additional factor to the HVS MTF:

$$MTF_{t\text{-}noise}(f_{eye}) = \sqrt{MTF^2_{eye}(f_{eye}) + (k_{noise} M_n)^2} . \qquad (11\text{-}20)$$

The value M_n can be calculated from the noise power spectral density presented to the eye.[5] Barten predicted the quality of imagery when both static (pattern noise) and dynamic (temporally varying) noise are present. He considered both one-dimensional and two-dimensional noise sources. His approach encompasses actual sensor operation where different noise components may exist in the horizontal and vertical directions. As such, he considered some of the three-dimensional noise components used for military applications (discussed in Section 12.2., *Three-Dimensional Noise Model*).

11.10. REFERENCES

1. H. Kusaka, "Consideration of Vision and Picture Quality - Psychological Effects Induced by Picture Sharpness," in *Human Vision, Visual Processing, and Digital Display*, B. E. Rogowitz, ed., SPIE Proceedings Vol. 1077, pp. 50-55 (1989).
2. E. M. Granger and K. N. Cupery, "An Optical Merit Function (SQF) which Correlates With Subjective Image Judgments," *Photographic Science and Engineering*, Vol. 16, pp. 221-230 (1972).
3. O. H. Shade, Sr., "Image Gradation, Graininess, and Sharpness in Television and Motion Picture Systems," published in four parts in *SMPTE Journal*: "Part I: Image Structure and Transfer Characteristics," Vol. 56(2), pp. 137-171 (1951); "Part II: The Grain Structure of Motion Pictures - An Analysis of Deviations and Fluctuations of the Sample Number," Vol. 58(2), pp. 181-222 (1952); "Part III: The Grain Structure of Television Images," Vol. 61(2), pp. 97-164 (1953); "Part IV: Image Analysis in Photographic and Television Systems," Vol. 64(11), pp. 593-617 (1955).
4. P. G. Barten, "Evaluation of Subjective Image Quality with the Square-root Integral Method," *Journal of the Optical Society of America. A*, Vol. 17(10), pp. 2024-2031 (1990).
5. P. G. Barten, "Evaluation of the Effect of Noise on Subjective Image Quality," in *Human Vision, Visual Processing, and Digital Display II*, J. P. Allenbach, M. H. Brill, and B. E. Rogowitz, eds., SPIE Proceedings Vol. 1453, pp. 2-15 (1991).
6. P. G. Barten, "Physical Model for the Contrast Sensitivity of the Human Eye," in *Human Vision, Visual Processing, and Digital Display III*, B. E. Rogowitz, SPIE Proceedings Vol. 1666, pp. 57-72 (1992).
7. P. G. Barten presents short courses at numerous symposia. See, for example, "Display Image Quality Evaluation," *Application Seminar Notes*, SID International Symposium held in Orlando, FL (May 1995). Published by the Society for Information Display, Santa Ana, CA, or "MTF, CSF, and SQRI for Image Quality," IS&T/SPIE's Symposium on Electronic Imaging: Science and Technology, San Jose, CA (February 1995).
8. G. C. Holst, *Electro-Optical Imaging System Performance*, pp. 218-228, JCD Publishing, Winter Park, FL (1996).
9. Air Standardization Agreement: "Minimum Ground Object Sizes for Imaging Interpretation," Air Standardization Co-ordinating Committee report AIR STD 101/11 (31 December 1976).
10. Air Standardization Agreement: "Imagery Interpretability Rating Scale," Air Standardization Co-ordinating Committee report AIR STD 101/11 (10 July 1978).
11. K. Riehl and L. Maver, "A Comparison of Two Common Aerial Reconnaissance Image Quality Measures," in *Airborne Reconnaissance XX*, W. G. Fishell, A. A. Andraitis, A. C. Crane, Jr., and M. S. Fagan, eds., SPIE Proceedings Vol. 2829, pp. 242-254 (1996).
12. J. M. Irvine, "National Imagery Interpretability Rating Scales (NIIRS): Overview and Methodology," in *Airborne Reconnaissance XXI*, W. C. Fishell, ed., SPIE Proceedings Vol. 3128, pp. 93-103 (1997).

13. G. C. Holst, *Sampling, Aliasing, and Data Fidelity*, pp. 311-317, JCD Publishing, Winter Park, FL (1998).
14. J. C. Leachtenauer, W. Malila, J. Irvine, L. Colburn, and N. Salvaggio, General Image-Quality Equation: GIQE," *Applied Optics*, Vol. 36(32), pp. 8322-8328 (1997).
15. G. C. Holst, *Sampling, Aliasing, and Data Fidelity*, pp. 308-311, JCD Publishing, Winter Park, FL (1998).
16. L. M. Beyer, S. H. Cobb, and L. C. Clune, "Ensquared Power for Obscured Circular Pupils with Off-Center Imaging," *Applied Optics*, Vol. 30(25), pp. 3569-3574 (1991).
17. A. J. den Dekker and A. van den Bos, "Resolution: A Survey," *Journal of the Optical Society A*, Vol. 14(3), pp. 547-557 (1997).
18. See, for example, W. J. Smith, *Modern Optical Engineering*, 2nd Edition, McGraw-Hill, New York (1990).
19. J. M. Lloyd, *Thermal Imaging*, p. 109, Plenum Press, New York (1975).
20. H. L. Snyder, "Image Quality and Observer Performance," in *Perception of Displayed Information*, L. M. Biberman, ed., pp. 87-118, Plenum Press, New York, NY (1973).

12

MINIMUM RESOLVABLE CONTRAST

Most image quality metrics were developed for high contrast imagery. Here, image quality is closely linked with the MTF. High MTF cameras provide better image quality and therefore provide better resolution under optimum conditions. While most resolution metrics are applicable to high contrast targets, the relative rating of systems may change for low contrast scenes. A high MTF system with excessive noise will not permit detection of low contrast targets.

The military is interested in detecting and recognizing targets at the greatest possible range. This is achieved by increasing the system gain until noise is apparent and the imagery is "snowy." Because the distinction between target and background is measured by contrast, the signal threshold is expressed as the minimum resolvable contrast (MRC). MRC applies to monochrome systems only. It is a measure of an observer's ability to detect a 3-bar target embedded in noise. The Johnson criterion relates real targets to equivalent bar chart targets. Thus, the link between real targets and the MRC is the spatial frequency of the equivalent 3-bar target.

By providing an angular resolution, it is simple to determine the range: R = (target size)/(resolution). Therefore, for MRC calculations, object-space spatial frequency is used. The MTFs provided in Chapter 10 were specified in the image-space spatial frequency domain. The conversion is $u_{ob} = fl\, u_i$ and $v_{ob} = fl\, v_i$, where fl is the effective optical focal length.

In the 1950s, Shade[1] modeled the resolution of photographic film and TV sensors as a function of light level. His approach is the framework of all models used today. In the 1970s, Rosell and Willson[2] applied Shade's results to thermal imaging systems and low light level televisions. Since then, most modeling efforts have concentrated on infrared imaging systems. However, the same modeling equations apply to all imaging systems. Different forms exist to accommodate the terminology associated with different technologies (scanning versus staring, visible versus thermal imaging systems, etc.).

MRC was originally developed for vidicons and low light level televisions. The approach taken here extends the thermal imaging system performance model (popularly called FLIR92) methodology[3] to visible systems. FLIR92 calculates the minimum resolvable temperature (MRT) where the target-background

radiance difference is characterized by a differential temperature. FLIR92 methodology and the three-dimensional noise model[4] are reformatted for solid state cameras. The 3-D noise model was developed to describe the various noise sources present in thermal imaging systems. These same noises exist in lesser amounts within solid state cameras.

The MRC approach uses terminology that characterizes human perceptual capability. The observer does not see the same contrast at the display as that presented by the target to the camera. Both the displayed contrast and brightness can be adjusted on the display so the observed contrast may be significantly different from what is present at the entrance aperture of the imaging system. Observer threshold is expressed as the perceived signal-to-noise ratio rather than the perceived contrast.

Within the real world, there is a probability associated with every parameter. The target contrast is not one number but a range of values that follow a diurnal cycle. The atmospheric transmittance is not fixed but can change in minutes. There appear to be an overwhelming set of combinations and permutations. Therefore, only a few representative target contrasts and a few representative atmospheric conditions are selected and the performance range is calculated for these conditions.

The Nyquist frequency limits the highest frequency that can be reproduced. The specific manner that the Nyquist frequency is included in the model greatly affects range performance. This is a result of the assumptions used and is not necessarily representative of actual system performance. While no model is perfect, models are still excellent for comparative analyses.

The symbols used in this book are summarized in the *Symbol List* (page xviii) which appears after the *Table of Contents*.

12.1. THE OBSERVER

The original Night Vision Laboratory static performance model[5] and FLIR92 use the Kornfeld-Lawson eye model (see Section 10.8., *Human Visual Response*, page 306). These models were validated by the U.S. Army when detecting tank-sized targets at modest ranges. Insufficient data exist to say with certainty which of the eye models available is best. Because of this and other model uncertainties, range predictions cannot be placed on an absolute scale. *All analyses, no matter which eye model is used, must be used only for comparative performance purposes.*

12.1.1. PERCEIVED SIGNAL-TO-NOISE RATIO

The human visual system "filters" enhance the displayed SNR. The perceived signal-to-noise ratio, referred back to the array output, is

$$SNR_P = k\frac{MTF_{SYS}\,\Delta n_{pe}}{\langle n_{SYS}\rangle}\,\frac{1}{(eye\ spatial\ filter)(eye\ temporal\ filter)}\ . \quad (12\text{-}1)$$

MTF_{SYS} includes the camera, electronics, display, and HVS MTFs (see Chapter 10, *System MTF*, page 267). All calculations are performed with object-space units (e.g., cycles/mrad). The value Δn_{pe} is the difference in the number of photoelectrons generated by the target and its immediate background. The value k depends on the optical aperture diameter, focal length, detector spectral quantum efficiency, etc.

This approach differs from the camera signal-to-noise ratio presented in Section 6.1.2. (page 186). There, the target was considered very large so the system MTF did not affect the signal strength. The rationale for using the MTF here, is that observer metrics (i.e., MRC) are based on the visibility of high frequency bar targets.

Although the HVS's modulation threshold is reported as a J-shaped curve (see Figure 10-30, page 307), the actual shape depends on the noise power spectral density.[6-8] If the noise is restricted to certain spatial frequencies, then the detection of targets of comparable spatial frequencies becomes more difficult. The observer's ability to see a specific spatial frequency target depends on the noise content in the neighborhood of that spatial frequency (Figure 12-1). Low spatial frequency noise components will interfere with detecting low frequency targets (large objects). Mid-spatial frequency noise increases the modulation threshold curve at mid-frequencies and so on. These noise factors are included in the MRC model via the three-dimensional noise model. As such, they are not included in the HVS MTF.

For nearly all imaging systems, square waves (bar targets) are used for system characterization. Using the CTF_{SYS} (See Section 9.5., *Contrast Transfer Function*, page 263), the perceived signal-to-noise ratio is

$$SNR_P = k'\,\frac{CTF_{SYS}\,\Delta n_{pe}}{\langle n_{SYS}\rangle (eye\ spatial\ filter)(eye\ temporal\ filter)}\ , \quad (12\text{-}2)$$

where k′ is a constant that includes k and the conversion from MTF to CTF.

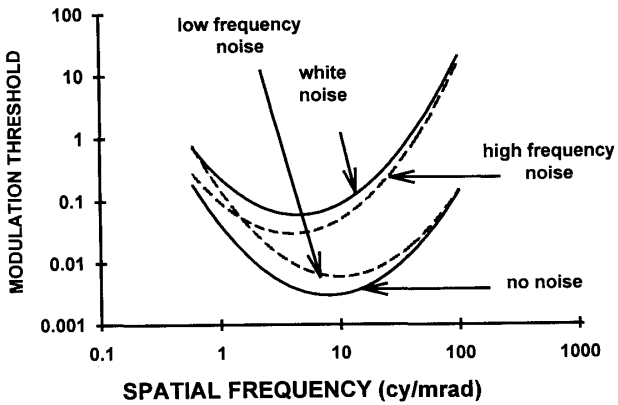


Figure 12-1. Effects of spectral noise on the HVS's modulation threshold. The MRC should approximately follow the shape of the modulation threshold curve.

Selecting a threshold value that is required to just perceive a target and inverting Equation 12-2 provide a minimum detectable target to background difference:

$$\Delta n_{MINIMUM} = SNR_{TH} \frac{\langle n_{SYS} \rangle (eye\,spatial\,filter)(eye\,temporal\,filter)}{k' CTF_{SYS}} \quad . \quad (12\text{-}3)$$

The system noise may consist of many components and the HVS may integrate each component differently. The function β_i is the "filter" that interprets noise source $\langle n_i \rangle$:

$$\Delta n_{MINIMUM} = SNR_{TH} \frac{\sqrt{\langle n_1^2 \rangle \beta_1 + ... + \langle n_m^2 \rangle \beta_m}}{k' CTF_{SYS}} \quad (12\text{-}4)$$

or

$$\Delta n_{MINIMUM} = SNR_{TH} \frac{\langle n_1 \rangle}{k' CTF_{SYS}} \sqrt{\beta_1 + ... + \frac{\langle n_m^2 \rangle}{\langle n_1^2 \rangle} \beta_m} \quad . \quad (12\text{-}5)$$

A test bar target is approximated by a square wave of infinite extent. If the system is band-limited, then only the fundamental of the square wave has sufficient amplitude to contribute to the response (valid for high frequencies). The fundamental's amplitude is $4/\pi$ times the square wave amplitude. The HVS is sensitive to the average value of the first harmonic and the average value of a half-cycle sine wave is $2/\pi$. Therefore, the conversion from a square wave (CTF_{SYS}) to a sinusoid (MTF_{SYS}) requires a factor of $8/\pi^2$ when viewed by an observer:

$$\Delta n_{MINIMUM} = SNR_{TH} \frac{\pi^2}{8} \frac{\langle n_1 \rangle}{k\, MTF_{SYS}} \sqrt{\beta_1^2 + \ldots + \frac{\langle n_m^2 \rangle}{\langle n_1^2 \rangle} \beta_m^2} \quad . \tag{12-6}$$

Following FLIR92 terminology,[3] $\langle n_1 \rangle$ is the shot noise. The other noise terms represent all other noise sources. The three-dimensional noise model quantifies each $\langle n_i^2 \rangle$.

In Section 6.2.1., *Camera SNR* (page 186), a target SNR was calculated. The background light level was assumed to be zero so that the photon shot noise was associated with the target only. For target detection, the target exists on a positive background. Both the target and the background regions provide photons that the camera converts to photoelectrons. Let $n_{pe\text{-}T}$ and $n_{pe\text{-}B}$ be the number of photoelectrons associated with the target and background, respectively. There is shot noise associated with both. The "average" shot noise variance is

$$\langle n_1^2 \rangle = \langle n_{SHOT}^2 \rangle \approx \frac{n_{pe\text{-}T} + n_{pe\text{-}B}}{2} \quad . \tag{12-7}$$

A variety of definitions exist for contrast. The contrast, as defined in the original MRC derivation, is the target luminance minus the background luminance divided by the background luminance. Because both the target and background are assumed to be illuminated by the same source (sun, moon, etc.), the contrast depends on reflectances:

$$C_o = \frac{|L_T - L_B|}{L_B} = \frac{|\rho_T - \rho_B|}{\rho_B} \quad . \tag{12-8}$$

Historically, the target was always considered darker than the background. However, the target can be brighter than the background and there does not appear to be any difference between the detectability of objects that are of negative or positive contrast.

Substituting Equations 12-7 and 12-8 into Equation 12-6 and calling the minimum detectable contrast the MRC, provides

$$MRC = SNR_{TH}\frac{\pi^2}{8}\frac{1}{k\,MTF_{SYS}}\sqrt{\frac{MRC+2}{2n_{pe-B}}}\sqrt{\beta_1+\ldots+\frac{\langle n_m^2\rangle}{\langle n_{SHOT}^2\rangle}\beta_m}\,. \quad (12\text{-}9)$$

Because MRC appears on both sides of the equation, it can be solved either iteratively or as a quadratic equation. The MRC is a family of curves dependent on the background level (Figure 12-2). MRC approaches a maximum value of unity as the system MTF approaches zero. Thus, MRC is linked to MTF but the resolution (i.e., the spatial frequency where the contrast approaches one) depends on the background.

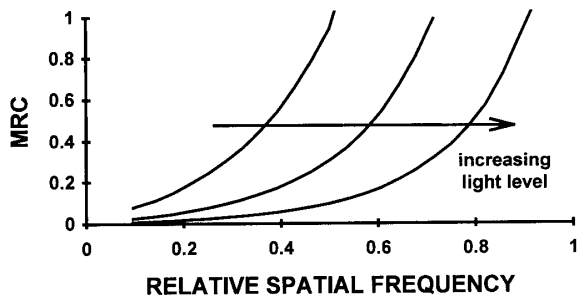


Figure 12-2. MRC is a family of curves that depends on the background light level.

12.1.2. PERCEIVED RESOLUTION

Perceived resolution depends on the size of the target, ambient lighting, and target-to-background contrast. These issues were not considered when discussing display resolution because they are relatively constant.

For signals embedded in noise, the situation is different. The HVS is then very sensitive to the signal-to-noise ratio. Slight variations affect target visibility. As illustrated in Figure 12-3, the signal increases faster than the accompanying photon shot noise. Given a threshold of perceptibility, as the SNR increases, more detail can be seen because more signals rise above the threshold. Thus, camera "resolution" depends on background light level. This "resolution" is strictly due to the HVS's response and is not related to hardware performance. In contrast, hardware resolution is typically specified as that spatial frequency where the hardware MTF is, say, 10% and is measured under relatively high signal-to-noise ratio conditions.

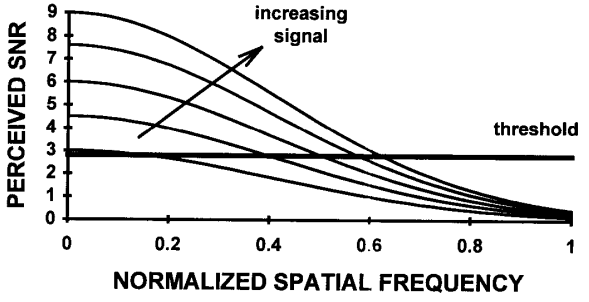


Figure 12-3. As the target signal increases, the SNR increases. For a fixed threshold, increasing the signal appears to increase the resolution.

Figure 12-4 illustrates the perceived resolution as a function of light level. The apparent resolution increases with light level. Because the SNR is proportional to contrast, the "resolution" is also a function of contrast. High contrast targets produce higher resolution. Again, this is strictly due to the HVS's response and is not related to hardware limitations.

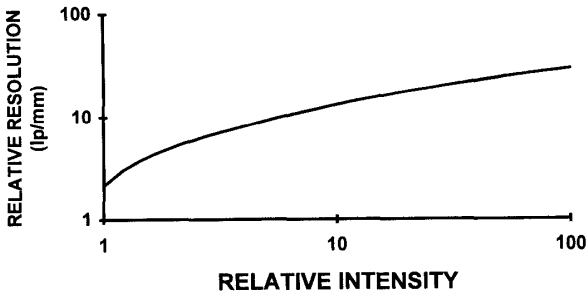


Figure 12-4. Perceived resolution as a function of light level.

12.1.3. THE HVS "FILTER"

The HVS system is probably the most difficult system to model. Although many models exist, the two most often employed are the matched filter and the synchronous integrator model. With the matched filter, the HVS acts like a tuned circuit. Here, the perceived SNR is affected only by spatial frequencies that are near the target's spatial frequency. All other spatial frequencies are ignored. This is not a filter in the usual sense because the signal and noise are "filtered" by it. Rather, visual psychophysical data suggest that the HVS acts as if it were a filter that can be described mathematically by a filter function. With the synchronous integrator model, the HVS integrates over an angular region defined by the target edges.

Scott and D'Agostino state[9] *The most frequently encountered methods...describe the eye/brain spatial integration using either a matched filter or a synchronous integrator model ... they yield virtually identical ... predictions ... Any potential difference ... will be lost in the inherent error of the .... measurements. For these reasons, and because ... a synchronous integrator is somewhat simpler ...., FLIR92 has been written as a synchronous integrator model.* Russell and Willson[2] used the synchronous integrator model for television systems and their approach is presented in this chapter.

12.2. THREE-DIMENSIONAL NOISE MODEL

The three-dimensional noise model[4] provides the basic framework for analyzing various noise sources. It was developed to describe the noise in thermal imaging systems. However, the methodology also applies to solid state cameras.

The noise is divided into a set of eight components that relate temporal and spatial noise to a three-dimensional coordinate system (Figure 12-5). This approach allows full characterization of all noise sources including random noise, pattern noise, streaks, rain, 1/f noise, and any other artifact. Analyzing the noise in this manner has the advantage of simplifying the understanding of complex phenomena by breaking it down into a manageable set of components. The method simplifies the incorporation of complex noise factors into model formulations.

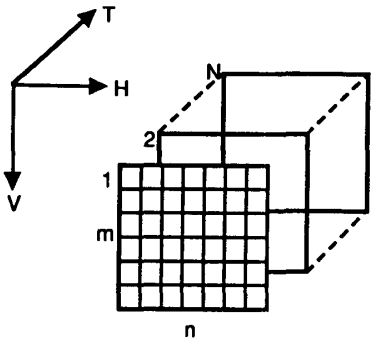


Figure 12-5. Three-dimensional noise model coordinate system illustrating data set N_{TVH}.

The T-dimension is the temporal dimension representing the video framing sequence. The other two dimensions (horizontal and vertical) provide spatial information. For a staring array, n and m indicate detector element locations in the horizontal and vertical directions, respectively.

Table 12-1 describes the three-dimensional noise components. The subscripts describe the noise "direction." The value σ_{TVH} represents the rms noise of the three-dimensional data set $m \times n \times N$. The value σ_{VH} is the rms noise after averaging in the T-direction. Its data set contains $m \times n$ elements. And so on.

Table 12-1
THREE-DIMENSIONAL NOISE DESCRIPTORS

3-D NOISE COMPONENT	DESCRIPTION
σ_{TVH}	Random 3-D noise
σ_{VH}	Spatial noise that does not change from frame-to-frame
σ_{TH}	Variations in column averages that change from frame-to-frame (rain)
σ_{TV}	Variations in row averages that change from frame-to-frame (streaking)
σ_{V}	Variations in row averages that are fixed in time (horizontal lines or bands)
σ_{H}	Variations in column averages that are fixed in time (vertical lines)
σ_{T}	Frame-to-frame intensity variations (flicker)

Depending on the system design and operation, any one of these noise components can dominate. The origins of these components are significantly different and the existence and manifestation depend on the specific design of the imaging system. Not all of the noise components may be present in every imaging system.

Assuming the noise components are independent, the total system noise is

$$\sigma_{SYS} = \sqrt{\sigma^2_{TVH} + \sigma^2_{TH} + \sigma^2_{TV} + \sigma^2_{VH} + \sigma^2_{H} + \sigma^2_{V} + \sigma^2_{T}} \ . \qquad (12\text{-}10)$$

Independence was a fundamental assumption during the development of the three-dimensional noise model. For nearly all systems, σ_T is considered negligible compared to σ_{TVH} and therefore is omitted from the data set. The value σ_{TVH} consists of the shot noise and the noise floor. Pattern noise is incorporated though σ_{VH}, σ_V, and σ_H. Currently, only σ_{TVH} is predicted and the remaining noise components must be determined from measurements or estimates. Figures 12-6 through 12-8, generated by the *System Image Analyzer*,[10] illustrate the effect of these noise sources. Figure 12-6 is the basic image. In Figure 12-7, FPN, PRNU, and random noise appear similar in a single frame. Random noise changes from frame-to-frame whereas pattern noise does not. Figure 12-8 illustrates how σ_{TV} or σ_V affect the visibility of horizontal bars. Scanning systems often exhibit this type of noise. Pixels will always have different responsivities. When the row-average responsivities are different in TDI sensors, banding occurs.[11]

Figure 12-6. Ideal image with $\sigma_{SYS} = 0$. Generated by *System Image Analyzer* software.[10]

Figure 12-7. Image with random, FPN, or PRNU noise.

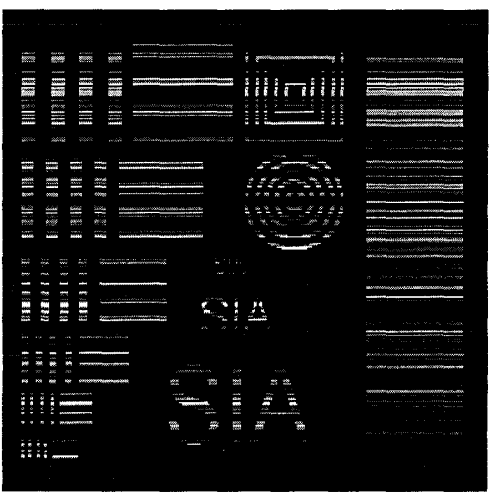


Figure 12-8. Image with dominant horizontal banding (high σ_{TV} or σ_V).

While the three-dimensional noise model was created to describe noise sources in thermal imaging systems, many of the same noises exist in all imaging systems. The FLIR92 default values should apply to all imaging systems built to similar designs.

With all systems, the most prevalent noise source is assumed to be temporal noise (σ_{TVH}). For staring arrays, the next most important noise source is σ_{VH}. Table 12-2 lists the FLIR92 recommended[3] values for PtSi staring arrays operating in the 3 to 5 μm range. Other systems using different detectors may perform differently. These values were obtained empirically.

For scanning systems, temporal row noise (σ_{TV}) and fixed row noise (σ_V) are the important contributors. Table 12-3 lists the FLIR92 default values appropriate for scanning systems with HgCdTe detectors operating in the 8 to 12 μm region. These values apply to all linear arrays (vertically oriented) with horizontal scanning.

Table 12-2
DEFAULT VALUES for STARING SYSTEMS

RELATIVE NOISE	DEFAULT VALUE
σ_{VH}/σ_{TVH}	0.40
σ_{TV}/σ_{TVH}	0
σ_V/σ_{TVH}	0
σ_{TH}/σ_{TVH}	0
σ_H/σ_{TVH}	0

Table 12-3
DEFAULT VALUES for SCANNING SYSTEMS

RELATIVE NOISE	DEFAULT VALUES
σ_{VH}/σ_{TVH}	0
σ_{TV}/σ_{TVH}	0.25
σ_V/σ_{TVH}	0.25
σ_{TH}/σ_{TVH}	0
σ_H/σ_{TVH}	0

12.3. THE MRC MODEL

The MRC is a measure of an observer's ability to detect a signal embedded in noise. The MRC model is a so-called *static* model in that the target is stationary and no search is required (or at least the observer knows where to look). The observer has an unlimited amount of time for target discrimination. As with most models, the system is assumed to be linear-shift-invariant.

This section applies the FLIR92 methodology to solid state camera performance. The fundamental difference between the MRC and MRT is the method by which the target is characterized. In the infrared spectral region, the target-to- background differential signal is specified by a temperature difference. In the visible spectral region, the differential signal is specified by the contrast. The MRT derivation can be found in the literature.[5,12-15]

Following FLIR92 methodology, the horizontal and vertical MRCs are

$$MRC_H(u_{ob}) = SNR_{TH} \frac{\pi^2}{8} \frac{1}{k\,MTF_H(u_{ob})} \sqrt{\frac{MRC + 2}{2n_{pe-B}}}\, K_H(u_{ob}) \quad (12\text{-}11)$$

and

$$MRC_V(v_{ob}) = SNR_{TH} \frac{\pi^2}{8} \frac{1}{k\,MTF_V(v_{ob})} \sqrt{\frac{MRC + 2}{2n_{pe-B}}}\, K_V(v_{ob}) \,. \quad (12\text{-}12)$$

The functions $K_H(u_{ob})$ and $K_V(v_{ob})$ are summary noise factors that combine the three-dimensional noise model components with the HVS's spatial and temporal integration capabilities. The variable $MTF_H(u_{ob})$ is the product of all the horizontal MTFs and $MTF_V(v_{ob})$ is the product of all the vertical MTFs (see Chapter 10, *System MTF*, page 267). In the original development of the MRC model, only the HVS spatial and temporal integration factors were considered. Later models included the HVS MTF.

The summary noise factors are

$$K_H(u_{ob}) = \left(E_T E_V E_H + \frac{\sigma^2_{VH}}{\sigma^2_{TVH}} E_V E_H + \frac{\sigma^2_{TH}}{\sigma^2_{TVH}} E_T E_H + \frac{\sigma^2_H}{\sigma^2_{TVH}} E_H \right)^{\frac{1}{2}} \quad (12\text{-}13)$$

and

$$K_V(v_{ob}) = \left(E_T E_V E_H + \frac{\sigma^2_{VH}}{\sigma^2_{TVH}} E_V E_H + \frac{\sigma^2_{TV}}{\sigma^2_{TVH}} E_T E_V + \frac{\sigma^2_V}{\sigma^2_{TVH}} E_V \right)^{\frac{1}{2}} . \quad (12\text{-}14)$$

For staring systems, the default (see Table 12-2) MRCs become

$$MRC_H(u_{ob}) = \frac{\pi^2}{8} \frac{SNR_{TH}}{k\,MTF_H(u_{ob})} \sqrt{\frac{MRC + 2}{2n_{pe\text{-}B}}} \sqrt{E_V E_H \left(E_T + \frac{4}{10} \right)} \quad (12\text{-}15)$$

and

$$MRC_V(v_{ob}) = \frac{\pi^2}{8} \frac{SNR_{TH}}{k\,MTF_V(v_{ob})} \sqrt{\frac{MRC + 2}{2n_{pe\text{-}B}}} \sqrt{E_V E_H \left(E_T + \frac{4}{10} \right)} . \quad (12\text{-}16)$$

For scanning systems (applies to linear arrays) the default (see Table 12-3) MRCs are

$$MRC_H(u_{ob}) = \frac{\pi^2}{8} \frac{SNR_{TH}}{k\,MTF_H(u_{ob})} \sqrt{\frac{MRC + 2}{2n_{pe\text{-}B}}} \sqrt{E_T E_V E_H} \quad (12\text{-}17)$$

and

$$MRC_V(v_{ob}) = \frac{\pi^2}{8} \frac{SNR_{TH}}{k\,MTF_V(v_{ob})} \sqrt{\frac{MRC + 2}{2n_{pe\text{-}B}}} \sqrt{E_V \left(E_T E_H + \frac{1}{4} E_T + \frac{1}{4} \right)} . \quad (12\text{-}18)$$

The values E_T, E_H, and E_V, respectively, represent the temporal, horizontal spatial and vertical spatial integration afforded by the HVS interpreter. The model assumes that the HVS temporally integrates perfectly and continuously over the HVS integration time. The noise is considered uncorrelated from frame-to-frame so that it adds in quadrature. When the information update rate, F_R, is high enough, and if there is signal correlation between adjacent frames:

$$E_T = \frac{1}{F_R t_e} . \qquad (12\text{-}19)$$

A 5:1 aspect ratio 3-bar target ($T_{ASPECT} = 5$) is used to measure the MRC, whereas a 7:1 aspect ratio 4-bar target is typically used for MRT measurements. The bars are oriented vertically for horizontal measurements and horizontally for vertical measurements. Thus, the equations used depend on the bar orientation. A vertical and horizontal MRC exists for each bar orientation. Table 12-4 lists the nomenclature. Before the two-dimensional MRC approach, only $E_{H\text{-}H}$ was used for horizontal MRC measurements. Similarly, only $E_{V\text{-}V}$ was used for vertical MRC measurements. The additional components, $E_{V\text{-}H}$ and $E_{H\text{-}V}$, are used in the two-dimensional MRC (discussed in Section 12.5., *Two-dimensional MRC*).

Table 12-4
MRC EQUATION NOMENCLATURE

MEASUREMENT	BAR ORIENTATION	MODEL NOMENCLATURE
Horizontal MRC	Vertical	Horizontal MRC: $E_{H\text{-}H}$ Vertical MRC: $E_{V\text{-}H}$
Vertical MRC	Horizontal	Horizontal MRC: $E_{H\text{-}V}$ Vertical MRC: $E_{V\text{-}V}$

The noise filters and noise power spectrum were described in Section 6.2.1., *Camera SNR* (page 186). The noise spectrum is modified by the electronic filters that occur after the insertion of noise. Because noise is assumed to originate in the detector, MTF_{OPTICS}, MTF_{MOTION}, and $MTF_{DETECTOR}$ are not part of the noise filters. As appropriate, the remaining MTFs provided in Chapter 10, *System MTF* (page 267), are components of the noise filters. For equation brevity, the horizontal and vertical noise filters are listed as $H_{NF\text{-}H}(u_{ob})$ and $H_{NF\text{-}V}(v_{ob})$, respectively. As appropriate, f_e, f_v, and u_d are expressed in equivalent object-space units (u_{ob}). Because electrical signals are typically read out in a serial data

stream (consistent with video standard timing), the noise appears in the horizontal direction only. The variable v_{ob} is the transposed variable v_d and also the vertical digital filter response if present.

For horizontal MRC, the bars are oriented vertically and the bar aspect ratio, T_{ASPECT}, appears in the vertical summary noise factors. For a bar target with fundamental spatial frequency u_o, the summary noise factors are

$$E_{H\text{-}H}(u_o) = \frac{d_{CCH}}{fl}\int_0^{\infty} S(u_{ob})\,|H_{NF\text{-}H}(u_{ob})|^2\, sinc^2\left(\frac{u_{ob}}{2u_o}\right) du_{ob} \quad (12\text{-}20)$$

and

$$E_{V\text{-}H}(u_o) = \frac{d_{CCV}}{fl}\int_0^{\infty} |H_{NF\text{-}V}(v_{ob})|^2\, sinc^2\left(\frac{T_{ASPECT}v_{ob}}{2u_o}\right) dv_{ob}\,. \quad (12\text{-}21)$$

For the vertical MRC, the bar aspect ratio is in the horizontal summary noise factor:

$$E_{H\text{-}V}(u_o) = \frac{d_{CCH}}{fl}\int_0^{\infty} S(u_{ob})\,|H_{NF\text{-}H}(u_{ob})|^2\, sinc^2\left(\frac{T_{ASPECT}u_{ob}}{2u_o}\right) du_{ob} \quad (12\text{-}22)$$

and

$$E_{V\text{-}V}(u_o) = \frac{d_{CCV}}{fl}\int_0^{\infty} |H_{NF\text{-}V}(v_{ob})|^2\, sinc^2\left(\frac{v_{ob}}{2u_o}\right) dv_{ob}\,. \quad (12\text{-}23)$$

In the early literature, only the horizontal MRC was calculated and only random temporal noise was modeled. With no other noise sources,

$$MRC_H(u_{ob}) = \frac{\pi^2}{8}\frac{SNR_{TH}}{k\,MTF_H(u_{ob})}\sqrt{\frac{MRC+2}{2n_{pe\text{-}B}}}\sqrt{E_V E_H E_T}, \quad (12\text{-}24)$$

where E_T, E_H, and E_V are given by Equations 12-19, 12-20, and 12-21, respectively.

If MTFs are near unity and the noise is white [$S(u_{ob}) \approx 1$] over the spatial frequencies of interest, then

$$E_H = E_{H\text{-}H}(u_o) \approx \frac{d_{CCH}}{fl} \frac{u_o}{T_{ASPECT}} \tag{12-25}$$

and

$$E_V = E_{V\text{-}H}(u_o) \approx \frac{d_{CCV}}{fl} u_o \, . \tag{12-26}$$

Then

$$MRC(u_o) = \frac{\pi^2}{8} \frac{SNR_{TH}}{k\,MTF(u_o)} \sqrt{\frac{MRC+2}{2n_{pe\text{-}B}}} \sqrt{\frac{1}{F_R t_e}} \sqrt{\frac{d_{CCH} d_{CCV}}{fl^2 T_{ASPECT}} u_o^2} \, . \tag{12-27}$$

The original SNR formulation[16] contained an $(a/A)^{1/2}$ factor where "a" is the target area at the image plane and "A" is the effective area of the photosensitive device. The MRC is the inverse of the SNR and therefore should contain $(A/a)^{1/2}$. This is identical to the last square-root term in Equation 12-27. "A" is the photosensitive area that contributes to noise. With vidicons, the "A" represented the entire image surface. With solid state arrays, it is the effective pixel size, $d_{CCH}d_{CCV}/fl^2$. The target size is $(1/u_o)(T_{ASPECT}/u_o)$. Equation 12-27 reduces to the basic MRC equation used by most researchers[17,18] prior to the introduction of the three-dimensional noise model.

12.4. SNR_{TH} and t_e

The FLIR92 documentation[19] states: *SNR_{TH} and the eye integration time, t_e, are most often the model parameters used to "tune" an MRTD prediction to a set of measurements.* These values are treated as constants that are adjusted so that the predicted values match laboratory data.

The 1975 NVL model documentation[20] states: ... *unfortunately, universal values for these constants do not exist. The values recommended at this time are $SNR_{TH} = 2.25$ and $t_e = 0.2$.*

FLIR92 documentation continues[19] *The luminances associated with about a 0.1 second eye integration time agrees well with the display experiments*

conducted in 1988 ... (darkened room and optimal viewing) ... In higher ambient light levels conditions, such as may be encountered with fielded systems, greater display luminances can be expected and thus a faster eye integration time may be appropriate... NVESD recommends setting SNR_{TH} equal to 2.5 ... Though psychophysical data show that SNR_{TH} is a function of target frequency, a value of 2.5 represents a reasonable average

The value t_e is somewhat nebulous and its value depends on the specific task at hand and ambient light level. FLIR92 recommends changing t_e according to the available light level (Figure 12-9).

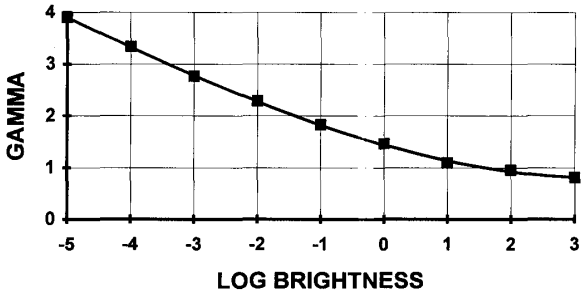


Figure 12-9. HVS integration values as a function of light level. FLIR92 documentation recommends using t_e = 0.1 s for laboratory MRT predictions. The same technique can be applied to MRC.

While many arguments exist for selecting the "correct" values, SNR_{TH} and t_e are simply constants. Both SNR_{TH} and t_e have a range of acceptable values. No matter what value is selected for t_e, consistency throughout the model must be maintained. The value t_e is also used in the linear motion MTF (see Section 10.5.1., *Linear Motion*, page 292).

A word of clarification is necessary here. SNR_{TH} is approximately the signal-to-noise value that the HVS perceives the image to be at threshold. It includes, in part, both temporal and spatial integration effects. It is *not* the signal-to-noise ratio of the video signal nor the SNR of the luminance on the display. Depending on the noise characteristics and the target characteristics, the HVS can detect[21] signals with video SNR as low as 0.05. Figure 12-10 provides an example of the HVS spatial integration capability.

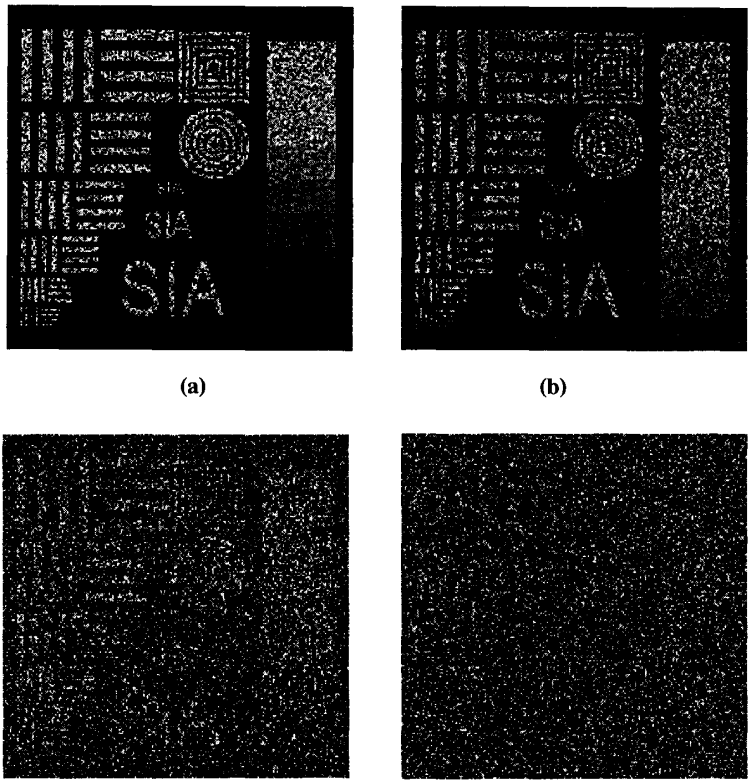


Figure 12-10. Variations in the signal-to-noise ratio, (a) SNR = 2, (b) SNR = 1, (c) SNR = 0.75, and (d) SNR = 0.5. The imagery was computer generated with the SNRs listed. The tonal transfer characteristics of the laser printer (to create the hard copy) and halftone (to print the image) may have affected the final SNR. The actual SNRs in this printed copy have not been measured. These images should be considered as representative of the ability to perceive low SNR imagery.

The ability to perceive targets embedded in noise depends on the target spatial extent with respect to the noise spatial frequencies. Large bars are easier to perceive compared to small targets that "disappear" into the noise. Figure 12-10 only illustrates the HVS's spatial integration capabilities. With live video, the random noise is changing from frame to frame and the HVS's temporal integration provides an additional enhancement of $F_R t_e$. For EIA 170 timing, this is about a factor of three to six.

12.5. TWO-DIMENSIONAL MRC

Although FLIR92 is called two-*dimensional*, it is a two-*directional* model. That is, the threshold is predicted along two orthogonal axes, taken as the vertical and horizontal directions. With unequal MTFs, the predicted MRC will be different in the two directions. The two directional values are brought together with a two-dimensional spatial frequency that approximates the response of an equivalent radially symmetric system.

As with the MRTs, the two-dimensional MRC is created by taking the geometric average of the horizontal and vertical MRC frequencies at each contrast value (Figure 12-11):

$$f_{2D} = \sqrt{u_1 v_1} \ . \tag{12-28}$$

In this manner, the two-dimensional MRC is forced to approach asymptotically the mean values of the vertical and horizontal cutoff frequencies. The two-dimensional MRC is a mathematical construct that is used for range performance predictions. It cannot be measured because f_{2D} does not exist.

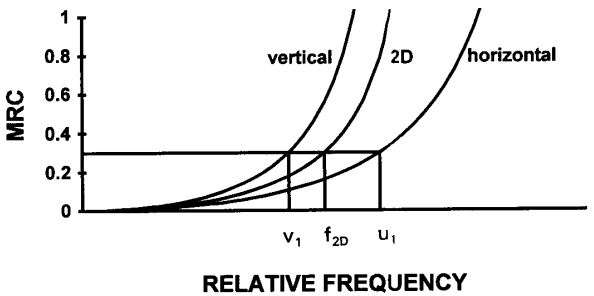


Figure 12-11. The two-dimensional MRC is mathematically created from the horizontal and vertical MRCs. It is really a two-*directional* MRC.

12.6. RANGE PREDICTIONS

Performance modeling assumes that the target is a single spatial frequency. This is an approximation because real targets consist of a complex superposition of spatial frequencies. Targets are actually a continuum of spatial frequencies from zero on up. The system MTF limits the target spatial frequencies that can be displayed. This section provides an overview of the range prediction methodology. A detailed discussion of system range performance is available.[22]

The Johnson criterion provides the link between the target angular subtense and the spatial frequency scale on the MRC graph. The target's inherent contrast, C_O, is modified by the atmospheric conditions (attenuation and path radiance). When the target contrast is reduced to the MRC curve, the performance range is obtained. Depending on the Johnson criterion selected, this becomes the detection, recognition, or identification range.

12.6.1. CONTRAST TRANSMITTANCE

Intuitively, it seems that any phenomena that prevent us from visually perceiving an object will affect all imagery the same way. During a hazy day, light scattered into the HVS (path radiance) reduces the visual contrast. Here, objects that are far away appear as a neutral white and we are unable to distinguish any features. While the HVS acts in this manner, imaging systems do not. Imaging systems simply respond to radiance differences.

However, real systems have automatic gain circuitry that may be activated by the path radiance signal. If so, the effect is to reduce system gain. Path radiance will partially fill the charge wells.

For most applications, the target and background are illuminated by the same source (sun, moon, etc.). The apparent contrast, C_R, is the inherent contrast modified by the atmosphere. Dropping the wavelength notation for equation brevity,

$$C_O = \frac{\Delta L}{L} = \frac{e^{-\sigma_{ATM}R}\left[\rho_T - \rho_B\right]L_{SCENE}}{e^{-\sigma_{ATM}R}\rho_B L_{SCENE} + L_{ATM}} = C_o \frac{1}{1 + \dfrac{L_{ATM}}{\rho_b L_{SCENE}} e^{\sigma_{ATM}R}}, \quad (12\text{-}29)$$

where L_{SCENE} is the ambient illumination that irradiates the target and background.

It is the sky, sun, or moon modified by the intervening atmospheric transmittance. The value σ_{ATM} is the atmospheric scattering cross section and $\exp(-\sigma_{ATM}R)$ is the atmospheric transmittance. The path radiance, L_{ATM}, is due to light scattered into the line-of-sight and the total amount received is an integral over the path length.[23] When only scattering is present (a reasonable approximation for the visible spectral band), $L_{ATM} \approx [1 - \exp(-\sigma_{ATM}R)]L_{SKY}$. The value L_{SKY} depends on the viewing direction and the location of the sun. The value $\rho_B L_{SCENE}$ is the background luminance. For convenience, the $L_{SKY}/\rho_B L_{SCENE}$ is called the sky-to-ground ratio (SGR):

$$C_R = C_O \frac{1}{1 + SGR\left(e^{\sigma_{ATM}R} - 1\right)} . \tag{12-30}$$

The SGR is approximately $0.2/\rho_B$ for a clear day ($L_{SKY}/L_{SCENE} \approx 0.2$) and $1/\rho_B$ for an overcast day ($L_{SKY}/L_{SCENE} \approx 1$). When the SGR is one, the received contrast is simply the inherent contrast reduced by the atmospheric transmittance:

$$C_R \approx C_O e^{-\sigma_{ATM}R} . \tag{12-31}$$

Meteorological range is defined quantitatively by the Koschmieder formula[24]:

$$R_{VIS} = \frac{1}{\sigma_{ATM}} \ln\left(\frac{1}{C_{TH}}\right) , \tag{12-32}$$

where C_{TH} is the modulation threshold at which 50% of the observers would see the target. The use of the scattering cross section in this definition rather than the extinction coefficient implies that absorption of particles at visual wavelengths is small enough to ignore. This view is probably justified except in cases of polluted air. Koschmieder set C_{TH} to 0.02 and evaluated σ_{ATM} at $\lambda = 0.555$ μm.Then

$$Transmittance = T_{ATM} = e^{-\sigma_{ATM}R} = e^{-\frac{3.912}{R_{VIS}}R} . \tag{12-33}$$

Table 12-5 lists σ_{ATM} as a function of standard weather conditions. However, σ_{ATM} decreases with increasing wavelength. For visibilities less than 6 km, σ_{ATM} can be approximated by

$$\sigma_{ATM}(\lambda) \approx \left(\frac{3.912}{R_{VIS}}\right)\left(\frac{0.55}{\lambda}\right)^{0.585(R_{VIS})^{1/3}} . \qquad (12\text{-}34)$$

For much greater visibilities,

$$\sigma_{ATM}(\lambda) \approx \left(\frac{3.912}{R_{VIS}}\right)\left(\frac{0.55}{\lambda}\right)^{0.5(R_{VIS})^{1/3}} . \qquad (12\text{-}35)$$

Table 12-5
INTERNATIONAL VISIBILITY CODE

DESIGNATION	VISIBILITY	SCATTERING COEFFICIENT, σ_{ATM}
Dense fog	0 - 50 m	$> 78.2\ km^{-1}$
Thick fog	50 - 200 m	$19.6 - 78.2\ km^{-1}$
Moderate fog	200 - 500 m	$7.82 - 19.6\ km^{-1}$
Light fog	500 - 1 km	$3.92 - 7.82\ km^{-1}$
Thin fog	1 - 2 km	$1.96 - 3.92\ km^{-1}$
Haze	2 - 4 km	$0.978 - 1.96\ km^{-1}$
Light haze	4 - 10 km	$0.391 - 0.978\ km^{-1}$
Clear	10 - 20 km	$0.196 - 0.391\ km^{-1}$
Very clear	20 - 50 km	$0.0782 - 0.196\ km^{-1}$
Exceptionally clear	> 50 km	$< 0.0782\ km^{-1}$

12.6.2. JOHNSON CRITERIA

The Johnson[25] methodology has become known as the equivalent bar pattern approach. Observers viewed military targets through image intensifiers and were asked to detect, decide the orientation, recognize, and identify the targets. Air Force tri-bar charts with bars the same contrast as the scaled models were also viewed and the maximum resolvable bar pattern frequency was determined. The number of bars per critical object dimension was increased until the bars could just be individually resolved. In this way, detectability was correlated with the sensor's threshold bar pattern resolution (Table 12-6). These results became the foundation for the discrimination methodology used today. *Critical* dimension refers to the target dimension that must be discerned for detection, recognition, or identification. In the early literature, the minimum dimension was called the critical dimension.

Table 12-6
JOHNSON'S RESULTS
(From Reference 25)

DISCRIMINATION LEVEL	MEANING	CYCLES ACROSS CRITICAL TARGET DIMENSION
Detection	An object is present (object versus noise).	1.0 ± 0.025
Orientation	The object is approximately symmetrical or unsymmetrical and its orientation may be discerned (side view versus front view).	1.4 ± 0.35
Recognition	The class to which the object belongs (e.g., tank, truck, man).	4.0 ± 0.80
Identification	The object is discerned with sufficient clarity to specify the type (e.g., T-52 tank, friendly jeep).	6.4 ± 1.50

Although Johnson used 6.4 cycles across the critical target dimension for identification, studies with thermal imaging systems at the Night Vision Laboratory suggested that eight cycles were more appropriate for identification. Orientation is a less popular discrimination level. Because today's standards are based on Johnson's work, they are labeled as the Johnson criteria though they are not the precise values recommended by him.

The original Johnson methodology was validated using common module thermal imaging systems. Those systems typically had a 2:1 difference in horizontal to vertical resolution. The vertical resolution was limited by the detector pitch Nyquist frequency. For validation purposes, military vehicles were used and the critical dimension tended to be in the vertical direction although the predictions were based on horizontal performance. Converting to a two-dimension model requires removing the directional bias imposed by the one-dimensional model. This is achieved by reducing all values by 25%, or equivalently, multiplying[26] all discrimination values by 0.75. This adjustment then provides the same range performance predictions whether predicted by the one-dimensional or two-dimensional model. Table 12-7 provides the commonly used discrimination levels. By convention, each 50% probability is labeled as N_{50} (e.g., in one-dimension, N_{50} is 1, 4, or 8 cycles for detection, recognition, or identification, respectively).

Table 12-7
COMMON DISCRIMINATION LEVELS

TASK	DESCRIPTION	One-dimension N_{50}	Two-dimension N_{50}
Detection	The blob has a reasonable probability of being an object being sought.	1.0	0.75
Classical recognition	Object discerned with sufficient clarity that its specific class can be differentiated.	4.0	3.0
Identification	Object discerned with sufficient clarity to specify the type within the class.	8.0	6.0

12.6.3. TARGET TRANSFER PROBABILITY FUNCTION

The results of threshold experiments, such as Johnson's, provide an approximate measure of the 50% probability of discrimination level. Results of several field tests[20] provided the cumulative probability of discrimination or target transfer probability function (TTPF) (Table 12-8). The TTPF can be used for all discrimination tasks by simply multiplying the 50% probability of performing the task by the TTPF multiplier. For example, from Tables 12-7 and 12-8, the probability of 95% recognition is $2N_{50} = 2(4) = 8$ cycles across the target's critical dimension.

Table 12-8
DISCRIMINATION CUMULATIVE PROBABILITY

PROBABILITY of DISCRIMINATION	MULTIPLIER
1.00	3.0
0.95	2.0
0.80	1.5
0.50	1.0
0.30	0.75
0.10	0.50
0.02	0.25
0	0

An empirical fit[27] to the data provides

$$P(N) = \frac{\left(\frac{N}{N_{50}}\right)^E}{1 + \left(\frac{N}{N_{50}}\right)^E}, \tag{12-36}$$

where

$$E = 2.7 + 0.7\left(\frac{N}{N_{50}}\right). \tag{12-37}$$

Figure 12-12 illustrates the full TTPF for the three levels of discrimination as identified by Johnson. Because of the variability in the population, some people can only detect the target, others will recognize it, and still a smaller portion will identify the target. This leads to the wide variations seen in field test results.

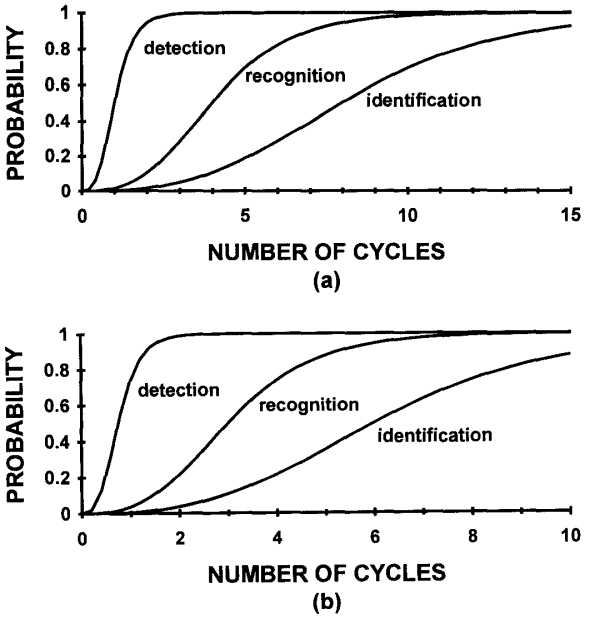


Figure 12-12. Target transfer probability function for detection, recognition, and identification. (a) one-dimensional and (b) two-dimensional. The curves overlap so that when only a small portion of the population can identify the target, a larger portion will recognize the target and an even larger portion will detect the target.

12.6.4. RANGE PREDICTION METHODOLOGY

The MRC abscissa is converted into a range scale using a target discrimination value. With the two-dimensional model, the critical dimension, h_c, is the square root of the target area, A_{TARGET}. For rectangular targets, $h_c = (HW)^{1/2}$ where H and W are the target's height and width, respectively. The number of equivalent cycles across the target is

$$N_T = \frac{\sqrt{A_{TARGET}}}{R} f_{2D} = \frac{h_c}{R} f_{2D} \,. \qquad (12\text{-}38)$$

Consider a target that is 2 × 8 m and six cycles are required across the critical dimension. Then the conversion from spatial frequency to range is

$$R = \frac{\sqrt{2 \cdot 8}}{6} f_{2D} = \frac{4}{6} f_{2D}, \tag{12-39}$$

where six cycles represent 95% probability of recognition for the two-dimensional model. The intersection of the MRC curve and the target apparent contrast, C_R (using Equation 12-31, page 360), versus range curve is the range at which the target can be discerned according to the discrimination level selected (Figure 12-13).

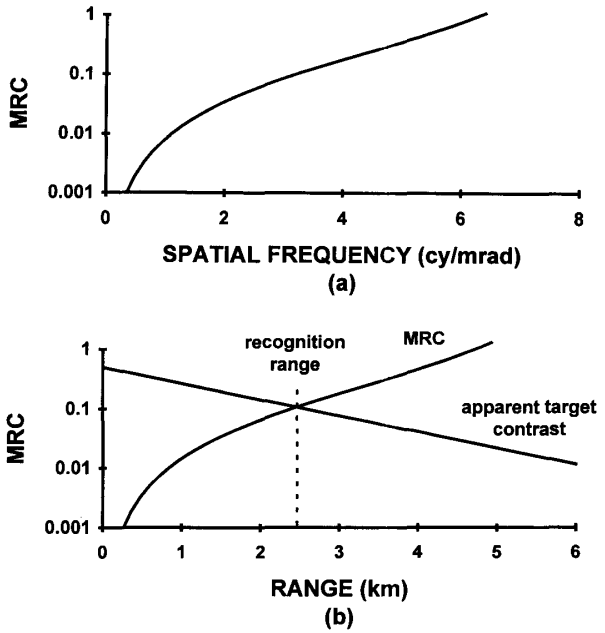


Figure 12-13. Graphical method to obtain the recognition range. (a) Two-dimensional MRC and (b) conversion of spatial frequency into range. The values $N_{50} = 3$, $h_c = 4$ m, and $C_O = 0.5$. The value C_R is a straight line when plotted in semi-logarithmic coordinates. The visibility is 6.2 km. The two-dimensional 50% probability recognition range is approximately 2.5 km.

The target transfer probability function is used to predict the probability of range performance. Here, a range is selected and C_R is calculated. This value intersects the MRC curve at what is called the critical frequency. Multiplying the angular subtense by the critical frequency provides the number of cycles across the target. Referring to the TTPF (see Figure 12-12), the probability is determined for that particular range. Then a new range is selected and the process is repeated until the entire probability function is determined (Figure 12-14).

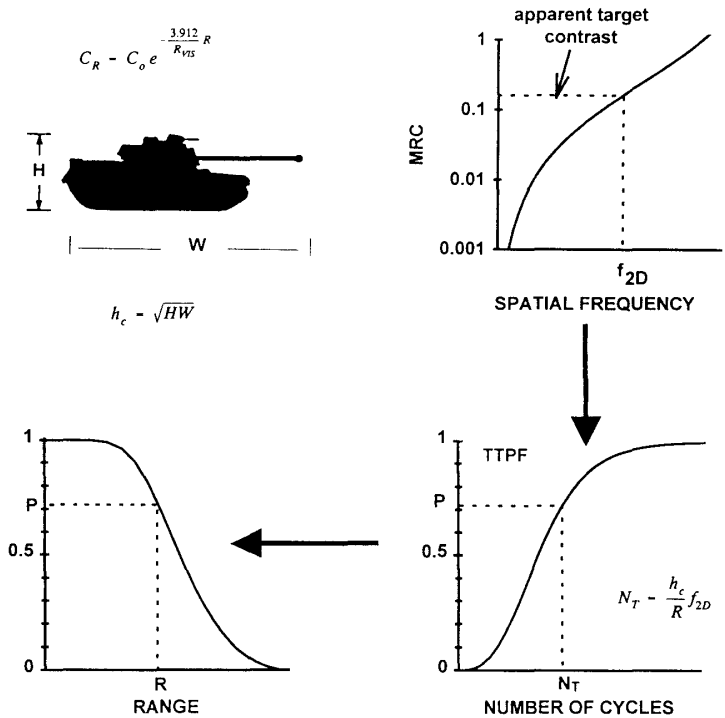


Figure 12-14. Methodology to determine range performance probability.

12.6.5. SAMPLING EFFECTS

In Chapter 8, *Sampling Theory*, page 231, it was stated that the highest frequency that could be faithfully reproduced is the Nyquist frequency. Solid state cameras are undersampled systems and are limited to a spatial frequency dictated by the detector center-to-center spacing. In the development of the MRC equation, sampling effects were ignored.

Because no frequency can exist above the Nyquist frequency, many researchers represent the MTF as zero above the Nyquist frequency. This forces the MRC to approach one at the Nyquist frequency (Figure 12-15). When the probability of detection is calculated using the target transfer probability function, the Nyquist limit reduces the range (Figure 12-16).

This range reduction is an artifact of the model. It is real when trying to discriminate sinusoids or bar targets. Real targets, however, are composed of many spatial frequencies ranging from zero and on up. Target features that have spatial frequencies above Nyquist will be distorted. No data exist on the detection or recognition of distorted imagery and it is, therefore, unwise at this juncture to extend the MRC model past the Nyquist frequency. The analyst must recognize that the Nyquist frequency limit may have an overly severe impact on target discrimination range. The specific manner in which the Nyquist frequency is included in the model has a dramatic effect on range performance. This is a result of the assumptions used and is not necessarily representative of actual system performance.

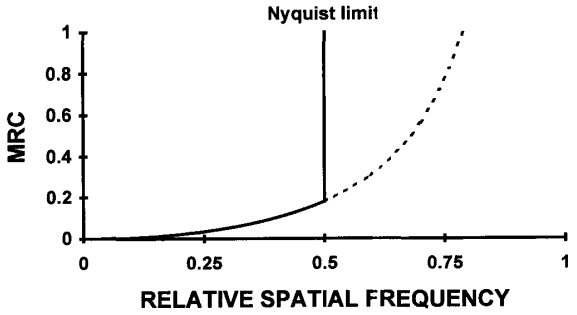


Figure 12-15. Representative MRC for a staring array sensor. At Nyquist frequency, the MRC is assumed to be one. The abrupt increase at Nyquist frequency is used by some researchers.

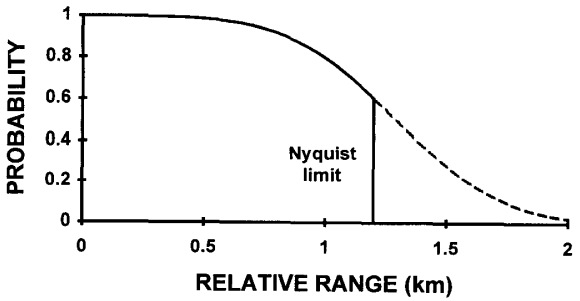


Figure 12-16. Probability of detection versus range for a staring array. The shapes of the curves depend on the MRC, contrast, atmospheric transmittance, and N_{50}. Because the MTF is forced to zero at the Nyquist frequency, the range hits a brick wall at Nyquist frequency. This is an artifact created by the model assumptions. Targets can be seen at longer ranges but aliasing will distort edges.

12.7. REFERENCES

1. O. H. Shade, Sr., "Image Gradation, Graininess, and Sharpness in Television and Motion Picture Systems," published in four parts in *SMPTE Journal*: "Part I: Image Structure and Transfer Characteristics," Vol. 56(2), pp. 137-171 (1951); "Part II: The Grain Structure of Motion Pictures - An Analysis of Deviations and Fluctuations of the Sample Number," Vol. 58(2), pp. 181-222 (1952); "Part III: The Grain Structure of Television Images," Vol. 61(2), pp. 97-164 (1953); "Part IV: Image Analysis in Photographic and Television Systems," Vol. 64(11), pp. 593-617 (1952).
2. F. A. Rosell and R. H. Willson, "Performance Synthesis of Electro-Optical Sensors," Air Force Avionics Laboratory Report AFAL-TR-72-229, Wright Patterson AFB, OH (1972).
3. *FLIR92 Thermal Imaging Systems Performance Model, User's Guide*, NVESD document UG5008993, Fort Belvoir, VA (1993).
4. J. D'Agostino and C. Webb, "3-D Analysis Framework and Measurement Methodology for Imaging System Noise," in *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing II*, G. C. Holst, ed., SPIE Proceedings Vol. 1488, pp. 110-121 (1991).
5. W. R. Lawson and J. A. Ratches, "The Night Vision Laboratory Static Performance Model Based on the Matched Filter Concept," in *The Fundamentals of Thermal Imaging Systems*, F. Rosell and G. Harvey, eds., NRL Report 8311, pp. 159-179, Naval Research Laboratory, Washington, D.C. (1979).
6. S. Daly, "Application of a Noise Adaptive Contrast Sensitivity Function in Image Data Compression," *Optical Engineering*, Vol. 29(8), pp. 977-987 (1990).
7. H. Pollehn and H. Roehrig, "Effect of Noise on the Modulation Transfer Function of the Visual Channel," *Journal of the Optical Society of America*, Vol. 60, pp. 842-848 (1970).
8. A. Van Meeteren and J. M. Valeton, "Effects of Pictorial Noise Interfering With Visual Detection," *Journal of the Optical Society of America A*, Vol. 5(3), pp. 438-444 (1988).

9. L. Scott and J. D'Agostino, "NVEOD FLIR92 Thermal Imaging Systems Performance Model," in *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing III*, G. C. Holst, ed., SPIE Proceedings Vol. 1689, pp. 194-203 (1992).
10. System Image Analyzer software is available from JCD Publishing, 2932 Cove Trail, Winter Park, FL 32789.
11. T. S. Lomheim and L. S. Kalman, "Analytical Modeling and Digital Simulation of Scanning Charge-Coupled Device Imaging Systems," in *Electro-Optical Displays*, M. A. Karim, ed., pp. 551-560, Marcel Dekker, New York (1992).
12. J. A. Ratches, W. R. Lawson, L. P. Obert, R. J. Bergemann, T. W. Cassidy, and J. M. Swenson, "Night Vision Laboratory Static Performance Model for Thermal Viewing Systems," ECOM Report ECOM-7043, p. 34, Fort Monmouth, NJ (1975).
13. J. M. Lloyd, *Thermal Imaging Systems*, pp. 182-194, Plenum Press, NY (1975).
14. F. A. Rosell, "Laboratory Performance Model," in *The Fundamentals of Thermal Imaging Systems*, F. Rosell and G. Harvey, eds., NRL Report 8311, pp. 85-95, Naval Research Laboratory, Washington, D.C. (1979).
15. R. L. Sendall and F. A. Rosell, "Static Performance Model Based on the Perfect Synchronous Integrator Model," in *The Fundamentals of Thermal Imaging Systems*, F. Rosell and G. Harvey, eds., NRL Report 8311, pp. 181-230, Naval Research Laboratory, Washington, D.C. (1979).
16. F. A. Rosell, "Video, Display, and Perceived Signal-to-Noise Ratios," in *The Fundamentals of Thermal Imaging Systems*, F. Rosell and G. Harvey, eds., NRL Report # 8311, pp. 49-83, Naval Research Laboratory, Washington, D.C. (1979)
17. W. W. Frame, "Minimum Resolvable and Minimum Detectable Contrast Prediction for Vidicon Cameras," *SMPTE Journal*, Vol. 104(1), pp. 21-26 (1985).
18. W. W. Frame, "Minimum Resolvable and Minimum Detectable Contrast Prediction for Monochrome Solid-State Imagers," *SMPTE Journal*, Vol. 96(5), pp. 454-459 (1987).
19. *FLIR92 Thermal Imaging Systems Performance Model, Analyst's Reference Guide*, NVESD document RG5008993, pp. ARG-12 - 13, Fort Belvoir, VA (1993).
20. J. A. Ratches, W. R. Lawson, L. P. Obert, R. J. Bergemann, T. W. Cassidy, and J. M. Swenson, "Night Vision Laboratory Static Performance Model for Thermal Viewing Systems," ECOM Report ECOM-7043, p. 56, Fort Monmouth, NJ (1975).
21. F. A. Rosell, "Psychophysical Experimentation," in *The Fundamentals of Thermal Imaging Systems*, F. Rosell and G. Harvey, eds., NRL Report # 8311, p. 225, Naval Research Laboratory, Washington, D.C. (1979).
22. G. C. Holst, *Electro-Optical Imaging System Performance*, Chapters 19, 20, and 21, JCD Publishing, Winter Park, FL (1995).
23. L. Levi, *Applied Optics*, pp. 118-124, Wiley and Sons (1980).
24. W. E. K. Middleton, *Vision Through the Atmosphere*, University of Toronto Press (1958).
25. J. Johnson, "Analysis of Imaging Forming Systems," in *Proceedings of the Image Intensifier Symposium*, pp. 249-273, Warfare Electrical Engineering Dept., U.S. Army Engineering Research and Development Laboratories, Ft. Belvoir, VA (1958). This article is reprinted in *Selected Papers on Infrared Design*, R. B. Johnson and W. L. Wolfe, eds., SPIE Proceedings Vol. 513, pp. 761-781 (1985).
26. L. Scott and R. Tomkinson, "An Update on the C^{2}NVEO FLIR90 and ACQUIRE Sensor Performance Model," in *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing II*, G. C. Holst, ed., SPIE Proceedings Vol. 1488, pp. 99-109 (1991).
27. J. D. Howe, "Electro-Optical Imaging System Performance Prediction," in *Electro-Optical Systems Design, Analysis, and Testing*, M. C. Dudzik, ed., p. 92. This is Volume 4 of the *Infrared and Electro-Optical Systems Handbook*, J. S Accetta and D. L. Shumaker, eds., copublished by Environmental Research Institute of Michigan, Ann Arbor, MI, and SPIE Press, Bellingham, WA (1993).

APPENDIX

A.1. EFFECTIVE FOCAL LENGTH

Most optical systems are composed of many individual lenses and mirrors. Each element may have a different refractive index and shape to minimize the system aberrations. The optical design can be considered as a single thick element (Figure A-1). For modeling purposes, it is treated as a single thin element with an effective focal length (Figure 2-6, page 23). The clear aperture is not necessarily the diameter of an optical element. The clear aperture limits the amount of scene radiation reaching the detector and is determined by the optical design.

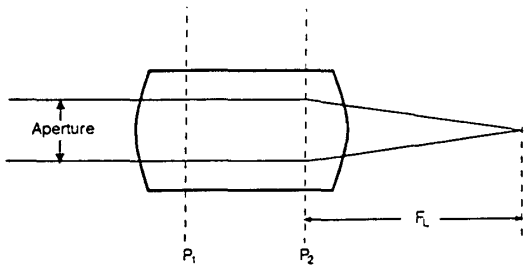


Figure A-1. The optical system can be considered as a single thick lens. P_1 and P_2 are the principal surfaces. The effective focal length is measured from the second principal surface. The clear aperture limits the amount of scene radiation reaching the detector. Although shown as planes, the principal surfaces are typically spherical.

A.2. f-NUMBER

The radiometric equations were derived from plane geometry and paraxial ray approximations. When using solid angles, the image incidance is proportional to $\sin^2(U')$ where U' is the maximum angle subtended by the lens. For paraxial rays, the principal surface is considered a plane. This representation is shown in most textbooks.

Lens design theory[1] assumes that the principal surface is spherical - every point on the surface is exactly a focal-length distance away (Figure A-2). This "sine corrected" lens has zero spherical aberration. Thus, $\sin^2(U')$ is equal to $1/(4F^2)$. The numerical aperture is another measure of the energy collected by the optical system. When the image is in air (refractive index of unity) the numerical aperture is

$$NA = \sin U' = \frac{1}{2F} . \qquad \text{(A-1)}$$

Since the largest angle is $\pi/2$, the smallest theoretical value for F is ½. This theoretical limit on F is not obvious from the radiometric equation.

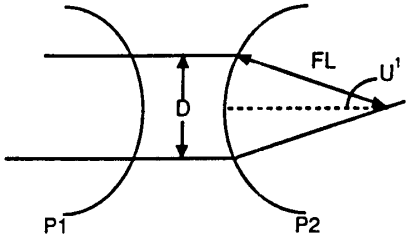


Figure A-2. Principal planes are typically spherical. Sin(U') = D/2fl = 1/2F (in air).

Not all optical systems have spherical principal planes. Often we assume the principal plane is flat. For hand-drawn optical layouts, the lenses are approximated by thin lenses and paraxial rays are used (large f-numbers). Here, $\sin(\theta) \approx \theta$ and Equation A-1 reduces to F = fl/D. Then $\sin^2(U')$ is $1/(4F^2+1)$. Whether $1/(4F^2+1)$ or $1/4F^2$ is used in the radiometric equations depends upon the f-number definition. If the analyst calculates the f-number from the effective lens diameter and focal length, then $1/(4F^2+1)$ should be used. If the optical designer provides the f-number, then the analyst must consult with him to ensure that the appropriate factor is used. For large f-numbers, the factors are approximately equal.

A.3. REFERENCE

1. W. J. Smith, *Modern Optical Engineering*, second edition, pp. 142-145, McGraw-Hill, New York (1990).

INDEX