UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO., LTD.; AND
SAMSUNG ELECTRONICS AMERICA, INC.
Petitioner

v.

IMAGE PROCESSING TECHNOLOGIES, LLC
Patent Owner

Patent No. 8,805,001

PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 8,805,001

TABLE OF CONTENTS

Contents

LIST OF EXHIBITS[1]

1001	U.S. Patent No. 8,805,001 ("the '001 Patent")
1002	Declaration of Dr. John C. Hart
1003	Curriculum Vitae for Dr. John C. Hart
1004	Prosecution File History of U.S. Patent No. 8,805,001
1005	Alton L. Gilbert et al., *A Real-Time Video Tracking System*, PAMI-2 No. 1 IEEE Transactions on Pattern Analysis and Machine Intelligence 47 (Jan. 1980) ("Gilbert")
1006	U.S. Patent 5,521,843 ("Hashima")
1007	U.S. Patent 5,150,432 ("Ueno")
1008	D. Trier, A. K. Jain and T. Taxt, "Feature Extraction Methods for Character Recognition-A Survey", Pattern Recognition, vol. 29, no. 4, 1996, pp. 641–662
1009	M. H. Glauberman, "Character recognition for business machines," Electronics, vol. 29, pp. 132-136, Feb. 1956
1010	Declaration of Gerard P. Grenier (authenticating Ex. 1005)

[1] Citations to non-patent publications are to the original page numbers of the publication, and citations to U.S. patents are to column:line number of the patents.

I. INTRODUCTION

Samsung Electronics Co., Ltd. and Samsung Electronics America, Inc. (collectively, "Petitioner") request *inter partes* review ("IPR") of Claims 1-4 of U.S. Patent No. 8,805,001 ("the '001 Patent") (Ex. 1001), which, on its face, is assigned to Image Processing Technologies, LLC ("Patent Owner"). This Petition presents several non-cumulative grounds of invalidity that the U.S. Patent and Trademark Office ("PTO") did not consider during prosecution. These grounds are each likely to prevail, and this Petition, accordingly, should be granted on all grounds and the challenged claims should be cancelled.

II. MANDATORY NOTICES UNDER 37 C.F.R. § 42.8

Real Parties-in-Interest: Petitioner identifies the following real parties-in-interest: Samsung Electronics Co., Ltd.; Samsung Electronics America, Inc.

Related Matters: Patent Owner has asserted the '001 Patent against Petitioner in *Image Processing Technologies LLC v. Samsung Elecs. Co.*, No. 2:16-cv-00505-JRG (E.D. Tex.). Patent Owner has also asserted U.S. Patent Nos. 6,959,293; 7,650,015; 8,983,134; and 8,989,445 in the related action. Petitioner is concurrently filing IPR petitions for all of these asserted patents.

Lead and Back-Up Counsel:

- Lead Counsel: John Kappos (Reg. No. 37,861), O'Melveny & Myers LLP, 610 Newport Center Drive, 17th Floor, Newport Beach,

California 92660. (Telephone: 949-823-6900; Fax: 949-823-6994; Email: jkappos@omm.com.)

- Backup Counsel: Nick Whilt (Reg. No. 72,081), Brian M. Cook (Reg. No. 59,356), O'Melveny & Myers LLP, 400 S. Hope Street, Los Angeles, CA 90071. (Telephone: 213-430-6000; Fax: 213-430-6407; Email: nwhilt@omm.com, bcook@omm.com.)

Service Information: Samsung consents to electronic service by email to IPTSAMSUNGOMM@OMM.COM. Please address all postal and hand-delivery correspondence to lead counsel at O'Melveny & Myers LLP, 610 Newport Center Drive, 17th Floor, Newport Beach, California 92660, with courtesy copies to the email address identified above.

III. PAYMENT OF FEES UNDER 37 C.F.R. § 42.15(a)

The Office is authorized to charge an amount in the sum of $23,000 to Deposit Account No. 50-0639 for the fee set forth in 37 CFR § 42.15(a), and any additional fees that might be due in connection with this Petition.

IV. GROUNDS FOR STANDING

Petitioner certifies that the '001 Patent is available for IPR and Petitioner is not barred or estopped from requesting IPR on the grounds identified herein.

V. PRECISE RELIEF REQUESTED

Petitioner respectfully requests review of Claims 1-4 of the '001 Patent, and

cancellation of these claims, based on the grounds listed below:

- **Ground 1**: Claims 1-4 are obvious under 35 U.S.C. § 103(a) over Gilbert in view of Hashima;
- **Ground 2**: Claims 1-4 are obvious under 35 U.S.C. § 103(a) over Hashima in view of Ueno; and
- **Ground 3:** Claims 1-4 are obvious under 35 U.S.C. § 103(a) over Ueno in view of Gilbert.

VI. LEGAL STANDARDS

A. Claim Construction

For expired claims, the Federal Circuit has held that the claims should be construed according to the *Phillips v. AWH Corp.* standard applicable in district court. *See In re Rambus Inc.* 753 F.3d 1253, 1256 (Fed. Cir. 2014). Under *Philips*, terms are given "the meaning that [a] term would have to a person of ordinary skill in the art in question at the time of the invention." *Phillips v. AWH Corp.*, 415 F.3d 1303, 1316 (Fed. Cir. 2005). Under 37 C.F.R. 42.100(b), the PTAB may also apply a district court-type claim construction if the patent is to expire within 18 months of the entry of the Notice of Filing Date.

The '001 Patent will expire on December 2, 2017—within 18 months of the Notice of Filing Date. Thus, for purposes of this proceeding, Petitioner has applied the *Phillips* standard to all claim construction issues. *See also* Ex. 1002, ¶49.

B. Level Of Ordinary Skill In The Art

One of ordinary skill in the art the time of the alleged invention of the '001 Patent would have had either (1) a Master's Degree in Electrical Engineering or Computer Science or the equivalent plus at least a year of experience in the field of image processing, image recognition, machine vision, or a related field or (2) a Bachelor's Degree in Electrical Engineering or Computer Science or the equivalent plus at least three years of experience in the field of image processing, image recognition, machine vision, or a related field. Additional education could substitute for work experience and vice versa. Ex. 1002, ¶¶44-48.

VII. OVERVIEW OF THE RELEVANT TECHNOLOGY AND '001 PATENT

The purported invention of the '001 Patent relates to identifying and tracking a target in an input signal using one or more histograms derived from an image frame in the video signal. *See e.g.*, Ex. 1001, Claims 1-4; Ex. 1002, ¶¶31-33. Video image processing and the use of histograms to identify and track targets, and to derive other information from a video signal were well known at the time the asserted patents were filed. Ex. 1002, ¶¶23-30, 60, 70, 77. An input signal used in the purported invention has "a succession of frames, each frame having a succession of pixels." Ex. 1001, 3:28-31. The input signal may be a video signal or any other signal that "generates an output in the form of an array of information corresponding to information observed by the imaging device," such as

"ultrasound, IR, Radar, tactile array, etc." Ex. 1001, 9:26-29; Ex. 1002, ¶32. The '001 Patent then constructs a histogram showing the frequency of pixels meeting a certain characteristic. The characteristics used to form histograms are referred to as "domains" in the '001 Patent. Ex. 1002, ¶34. The '001 Patent teaches that "the domains are preferably selected from the group consisting of i) luminance, ii) speed (V), iii) oriented direction (DI), iv) time constant (CO), v) hue, vi) saturation, and vii) first axis (x(m)), and viii) second axis (y(m))." Ex. 1001, 4:2-6; Ex. 1002, ¶¶34, 36. Figure 11 shows histogram processors that can create histograms in various domains:

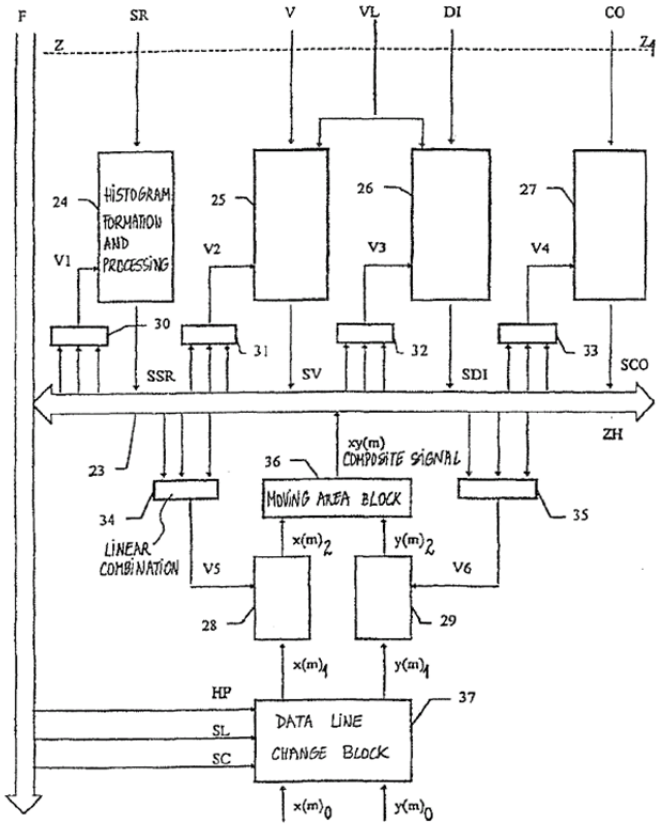
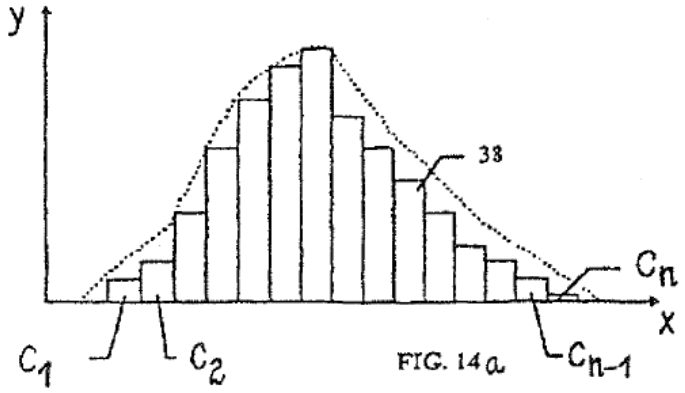


FIG. 11

The histograms include a plurality of "classes" within a given domain. Ex. 1002, ¶35. Figure 14a (and its accompanying description) illustrates an example of "classes" within a domain:



> FIG. 14*a* shows an example of the successive classes C_1 $C_{2...}C_{n-1}$ C_n, each representing a particular velocity, for a hypothetical velocity histogram, with their being categorization for up to 16 velocities (15 are shown) in this example. Also shown is envelope **38**, which is a smoothed representation of the histogram.

Ex. 1001, 20:54-59.

The '001 Patent then uses the histograms to identify a target in the input signal. For example, one embodiment of the '001 Patent performs "automatic framing of a person…during a video conference." *Id.* at 22:10-12, Figure 15:

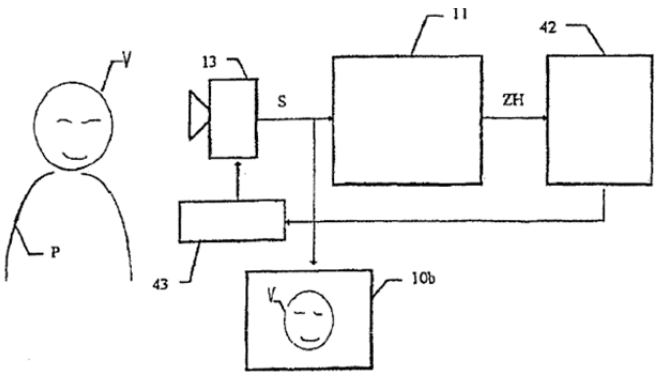


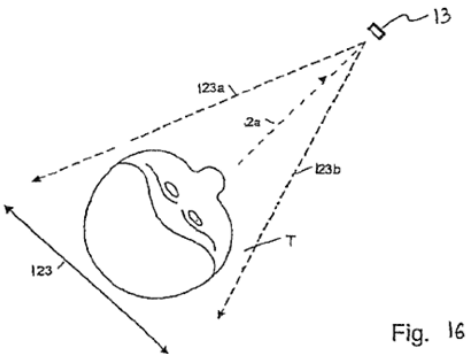
FIG.15

The system constructs histograms in the X- and Y-domains counting the number of pixels where the differences in luminance between successive frames are above certain threshold values:

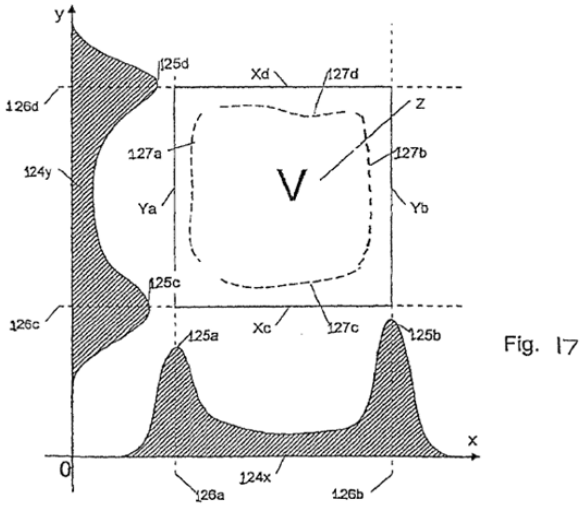
> The pixels with greatest movement within the image will normally occur at the peripheral edges of the head of the subject, where even due to slight movements, the pixels will vary between the luminance of the head of the subject and the luminance of the background. Thus, if the system of the invention is set to identify only pixels with DP=1, and to form a histogram of these pixels, the histogram will detect movement peaks along the edges of the face where variations in brightness, and therefore in pixel value, are the greatest, both in the horizontal projection along Ox and in the vertical projection along

Oy.

Id. at 22:49-59, 10:30-58 (explaining that DP is set to "1" when pixel value of the pixel under consideration has "undergone significant variation as compared to…the same pixel in the prior frame"); Ex. 1002, ¶¶36-37. Figures 16 and 17 show camera setup and the histogram constructed using this method:



Ex. 1001, Fig. 16



Ex. 1001, Fig. 17

In addition, the system may also be used to automatically track a target by "a spotlight or a camera. Using a spotlight the invention might be used on a helicopter to track a moving target on the ground, or to track a performer on a stage during an exhibition. The invention would similarly be applicable to weapons targeting systems." Ex. 1001, 23:40-46; Ex. 1002, ¶¶38-39. In such applications, the system uses X- and Y-minima and maxima of the histograms in X- and Y-domains to determine the center of the target:

> In a preferred embodiment, the new center of the area is determined to be $(X_{MIN}+X_{MAX})/2$, $(Y_{MIN}+Y_{MAX})/2$, where X_{MIN} and X_{MAX} are the positions of the minima and maxima of the x projection histogram, and Y_{MIN} and Y_{MAX} are the positions of the minima and maxima of the y projection histogram.

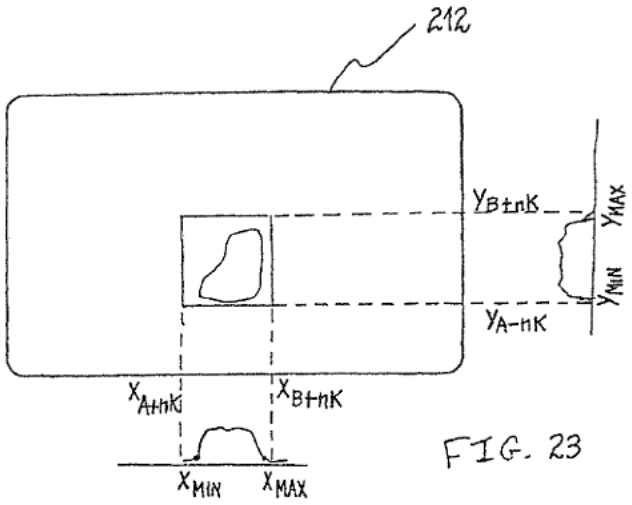
Ex. 1001, 24:50-55. The patent defines "the positions of the minima" of a projection histogram to be the smallest X- (and Y) coordinate of any pixel in the image region whose validation signal is "1." Ex. 1002, ¶¶40-41. Similarly the maximum is the largest X- (and Y) coordinate of any pixel in the image region whose validation signal is "1." *Id*. The system may recalculate the histograms at regular intervals and use those points to again find the center coordinates as the new frames are received. Ex. 1002, ¶42.

Once the center of the target is determined, the center is used to adjust the camera or spotlight to be directed to the moving target:

> Having acquired the target, controller **206** controls servomotors **208** to maintain the center of the target in the center of the image….
>
> It will be appreciated that as the target moves, the targeting box will move with the target, constantly

> adjusting the center of the targeting box based upon the movement of the target, and enlarging and reducing the size of the targeting box. The targeting box may be displayed on monitor **212**, or on another monitor as desired to visually track the target.

Ex. 1001, 25:12-25; Ex. 1002, ¶43. Figure 23 shows an example of the targeting box in a frame:



Ex. 1001, Fig. 23

VIII. DETAILED EXPLANATION OF GROUNDS

A. Overview Of The Prior Art References

1. Alton L. Gilbert et al., *A Real-Time Video Tracking System*, PAMI-2 No. 1 IEEE Transactions on Pattern Analysis and Machine Intelligence 47 (Jan. 1980) ("Gilbert") (Ex. 1005)

The purported invention of the '001 Patent relates to a process of identifying a target in digitized visual input by using histograms of pixel characteristics and

tracking the target. However, researchers at U.S. Army White Sands Missile Range, New Mexico, in collaboration with New Mexico State University, Las Cruces, had already developed a system that utilizes histograms to identify and track targets, and they published their findings in January 1980, more than 17 years before the earliest effective filing date of the '001 Patent. Ex. 1002, ¶50; Ex. 1010, Grenier Decl.

The article, entitled "A Real-Time Video Tracking System," published in IEEE Transactions on Pattern Analysis and Machine Intelligence in January 1980 ("Gilbert"), qualifies as prior art under pre-AIA § 102(b). Gilbert describes "a system for missile and aircraft identification and tracking…applied in real time to identify and track objects." Ex. 1002, ¶51; Ex. 1005, 47. Gilbert was not of record and was not considered during prosecution of the '001 Patent. The Gilbert system includes an image processing system comprising a video processor, a projection processor, a tracker processor, and a control processor as shown in Figure 1, reproduced below. Ex. 1002, ¶51; Ex. 1005, 48.

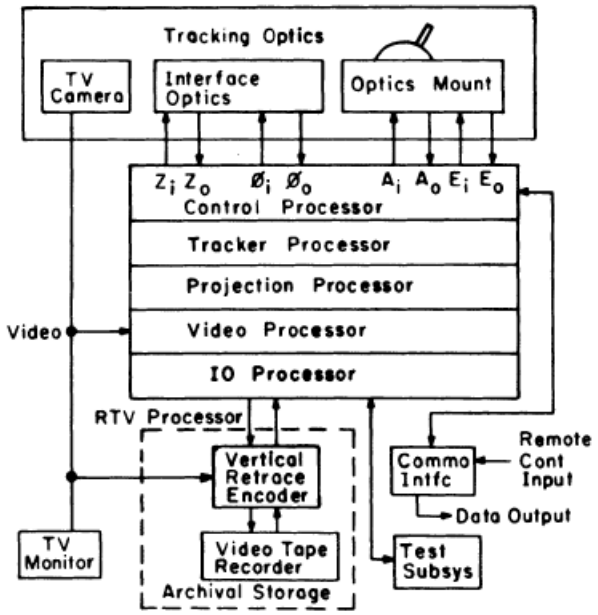


Fig. 1. RTV tracking system.

The video processor receives a digitized video signal comprising 60 fields/s, i.e., 30 frames/s, as input. Ex. 1002, ¶52; Ex. 1005, 48. Each field—*i.e.*, frame—consists of a succession of *n* X *m* pixels:

> As the TV camera scans the scene, the video signal is digitized at *m* equally spaced points across each horizontal scan. During each video field, there are *n* horizontal scans which generate an *n* X *m* discrete matrix representation at 60 fields/s.

Ex. 1005, 48. Although Gilbert uses the word "field" instead of "frame," a POSA would have understood that a "frame" consists of two "fields" in the context in which they are used in Gilbert. (At the time, video signals were interlaced such

that a frame of a non-interlaced video consisted of two fields. The first field would be the odd numbered scanlines and the second field would be the even numbered scanlines recorded 1/60th of a second later than the first field.) Ex. 1002, ¶52. Gilbert then constructs histograms of pixels in the 256 gray-level classes in the intensity domain:

> Every 96 ns, a pixel intensity is digitized and quantized into eight bits (256 gray levels), counted into one of six 256-level histogram memories, and then converted by a decision memory to a 2-bit code indicating its classification (target, plume, or background.).

Ex. 1005, 48, 49 ("Then the entire region has been scanned, h contains the distribution of pixels over intensity and is referred to as the feature histogram of the region R."). In other words, the Video Processor of Gilbert creates histograms using the intensity domain over classes of all intensity values. Ex. 1002, ¶52. Although Gilbert uses histograms in the intensity domain as examples, it also notes that other "features that can be functionally derived from relationship between pixels, e.g., texture, edge, and linearity measure" may be used. Ex. 1005, 48; Ex. 1002, ¶53.

Using the histograms, the video processor "separates the target from the background," *i.e.*, identifies the target. Ex. 1005, 48. Gilbert uses probability

estimates based on a 256 level grayscale histogram to determine whether a particular pixel belongs to the target, plume, or background region. Ex. 1002, ¶52. This identification is repeated for each frame. *See* Ex. 1005, 47 ("The camera output is statistically decomposed into background, foreground, target, and plume region by the video processor, with this operation carried on at video rate for up to the full frame.").

Once the video processor identifies the pixels belonging to the target, the video processor creates "a binary picture, where target presence is represented by a '1' and target absence by a '0.'" Ex. 1005, 50 ("In the projection processor, these matrices are analyzed field-by-field at 60 field/s using projection based classification algorithm to extract the structural and activity parameters needed to identify and track the target."); Ex. 1002, ¶54. This binary picture indicates to the projection processor whether or not a pixel should be considered further in constructing X- and Y-projections. Ex. 1002, ¶54. A projection processor creates projections using only the pixels identified for inclusion. Ex. 1002, ¶55. Although these projections are not explicitly referred to by Gilbert as projection histograms, reference to Figure 4 of Gilbert (annotated below) clearly shows four different projection histograms formed using the target pixels:

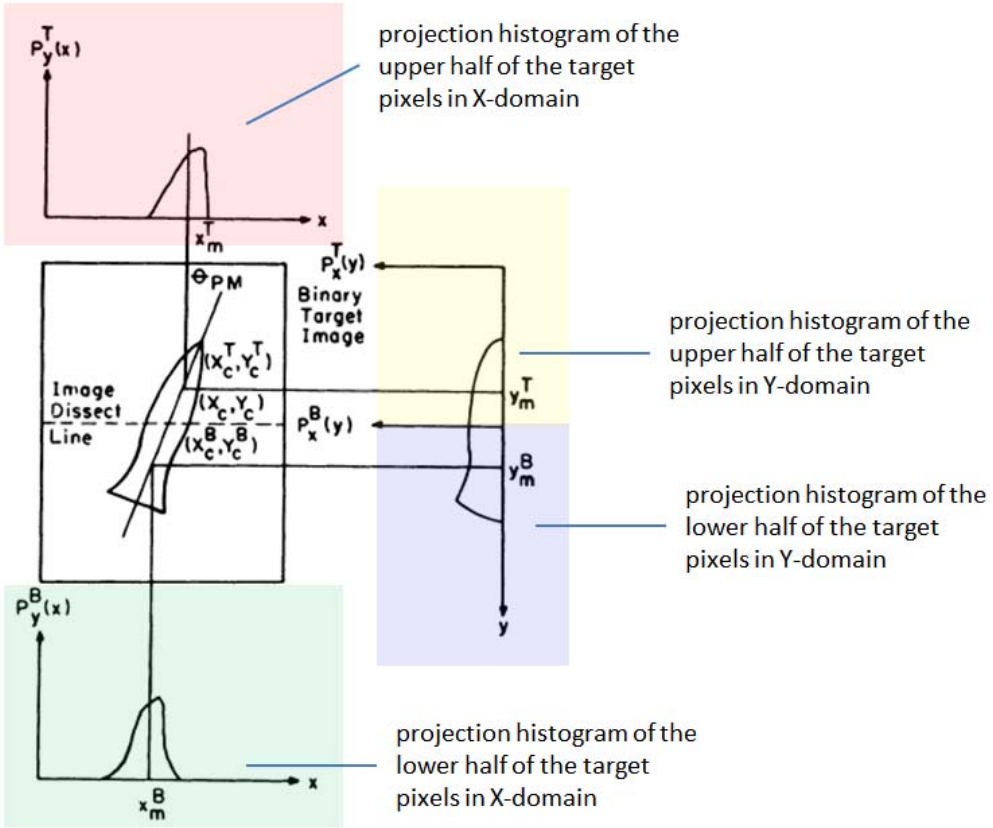


Fig. 4. Projection location technique.

In addition, Gilbert explains that a projection "gives the number of object points along parallel lines; hence it is a distribution of the target points for a given view angle." Ex. 1005, 50. Thus, these Figure 4 projections will be referred to as projection histograms throughout this petition.

The projection processor then identifies the target location, orientation, and structure using the projection histograms:

> The target location, orientation, and structure are characterized by the pattern of 1 entries in the binary picture matrix, and the target activity is characterized by

> a sequence of picture matrices. In the projection processor, these matrices are analyzed field-by-field at 60 fields/s [i.e., 30 frames/s]….

Ex. 1005, 50. The projection processor computes a center of area point for the target in order to "precisely determine the target position and orientation." *Id.*; Ex. 1002, ¶56. This calculation is done by first finding a center of area for each of top and bottom portions using the projection histograms in X- and Y-domains for the top and bottom portions of the target. Ex. 1002, ¶¶56-57. The projection processor then uses these center-of-area points to determine a target center-of-area point. This is shown in Figure 4, reproduced below (with annotations):

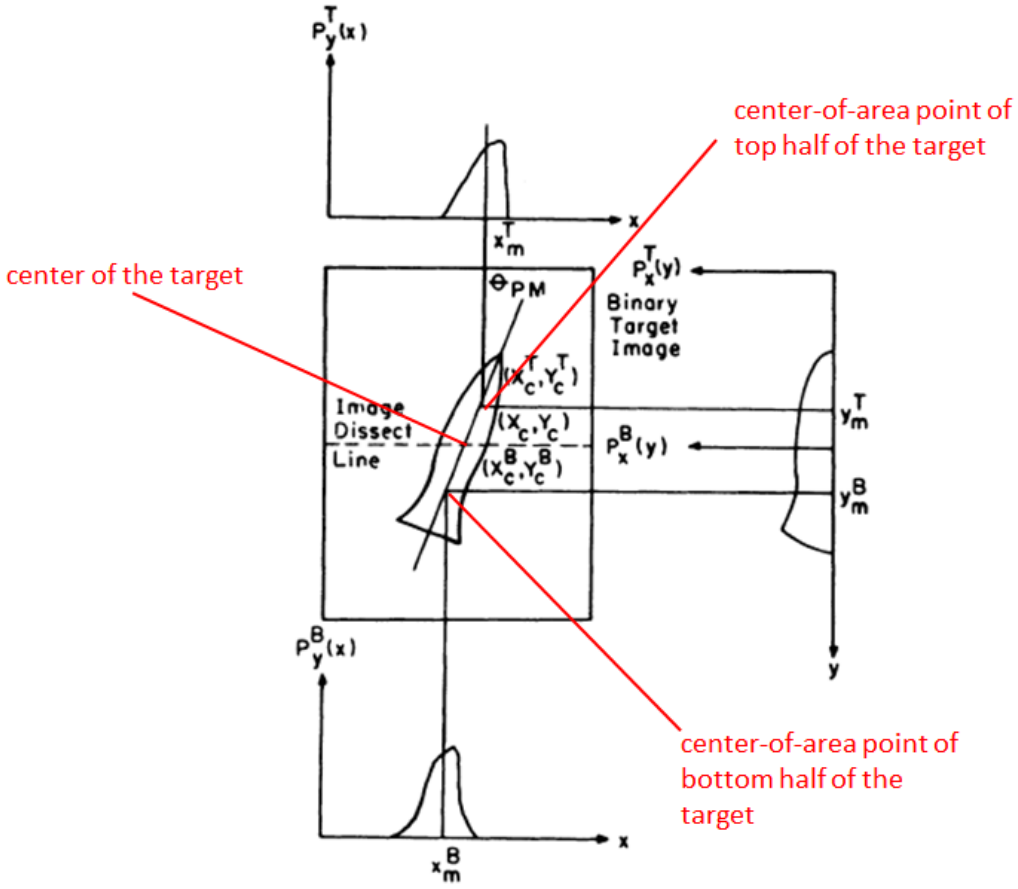


Fig. 4. Projection location technique.

The tracker processor analyzes information, such as the target's size, location, and orientation, from the projection processor and outputs information such as "1) tracking window size, 2) tracking window shape, and 3) tracking window position." Ex. 1005, 52; Ex. 1002, ¶59. The video processor then uses this information to draw a tracking window, i.e., a tracking box, around the target. *See id.* The display of the tracking box in the frame is shown in Figure 2.

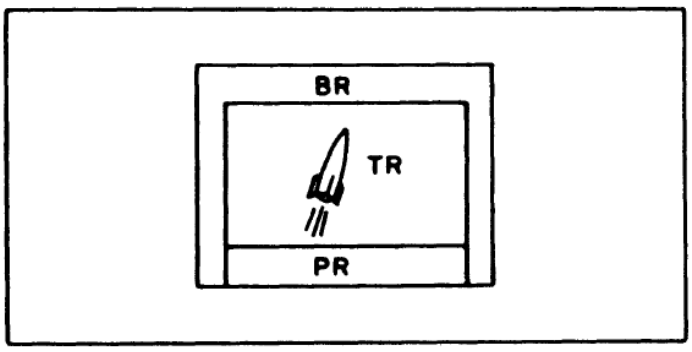


Fig. 2. Tracking window.

The tracker processor also outputs the target's movements to the control processor, which controls the direction and zoom scale of the lens to follow the target:

> The outputs to the control processor are used to control the target location and size for the next frame. The bore-sight correction signals are used to control the azimuth and elevation pointing angles of the telescope. The desired zoom is used to control the zoom lens, keeping the target visible within the FOV. The desired image rotation controls the image rotation element to keep the target image vertical.

Ex. 1005, 52.

The histogram formation and target identification process of Gilbert is performed on a frame-by-frame basis. Ex. 1005, 52; Ex. 1002, ¶58.

2. U.S. Patent No. 5,521,843 ("Hashima") (Ex. 1006)

Hashima also discloses a "system for and method of recognizing and tracking a target mark…by processing an image of the target mark produced by a video camera." Ex. 1006, 1:6-12. Hashima qualifies as prior art under pre-AIA 35 U.S.C. §§ 102(a), 102(b), 102(e), and § 119 ("but no patent shall be granted…for an invention which had been…described in a printed publication in any country more than one year before the date of the actual filing of the application in this country."). Although Hashima was of record, it was not applied during prosecution of the '001 Patent. Hashima uses an image processing system that takes input from a video camera and identifies the target shape based on histograms constructed from pixel characteristics. Ex. 1002, ¶¶61-62. Once the target shape is detected, Hashima uses the histogram information to determine the location of the target, and move a robot arm based on the information to grip the target object. One embodiment of this system is shown in Figure 1, reproduced below.

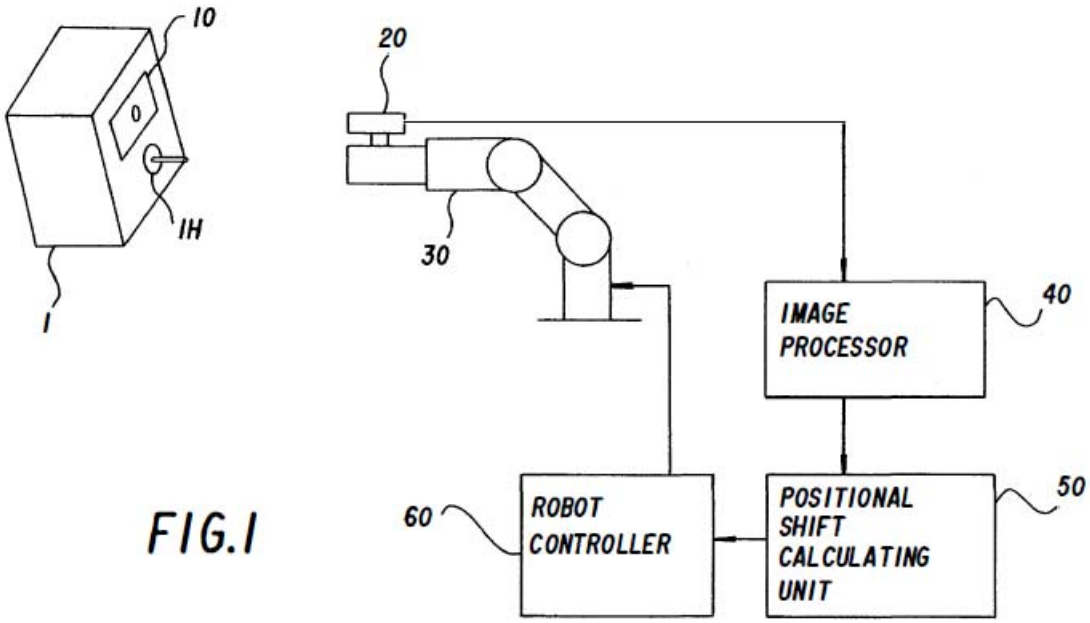
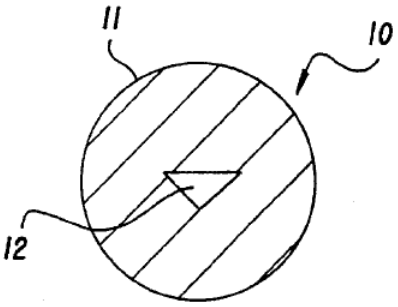
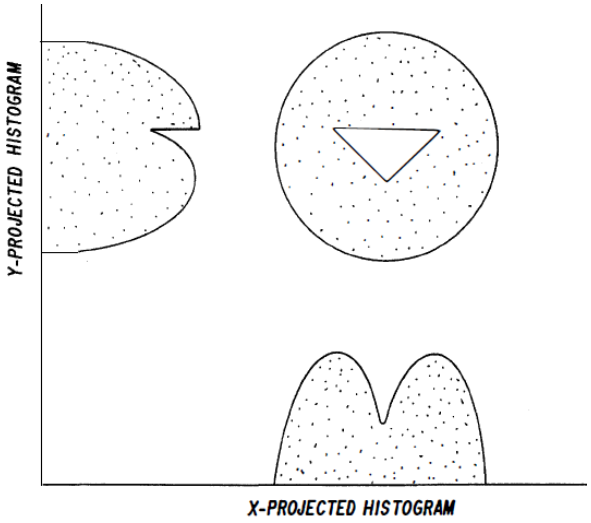


FIG.1

Hashima uses a pre-determined mark, a black circle with white triangle inside, as shown in Figure 3 below, and Hashima notes that "target marks of various configurations…can be detected by the [disclosed] process." Ex. 1006, 10:20-23; Ex. 1002, ¶62. The histograms of the exemplary mark in X- and Y-domains counting the number of black pixels are shown in Figure 6.



Hashima Fig. 3



Hashima Fig. 6

The image processor reads input from the video camera, converts the image into a binary image, and constructs a histogram of the binary images. Ex. 1006, 8:22-30 ("X- and Y-projected histograms of the target mark image are determined …. The X-projected histogram represents the sum of pixels having the same X-coordinates, and the Y-projected histogram represents the sum of pixels having the same Y-coordinates."); Ex. 1002, ¶62. *See generally*, Ex. 1006, 8:18-9:7. Then the image processor determines whether the image represents the target by counting the number of peaks and valleys in the projected histogram. Ex. 1006, 9:8-9:13; Ex. 1002, ¶63. If the number matches the number of peaks and valleys of the X- and Y-projected histograms as shown in Figure 6 (reproduced above), the system identifies the image as that of the target mark. *Id.* at 9:13-9:23; Ex. 1002, ¶63. Figure 5 shows the flow chart describing the detection process:

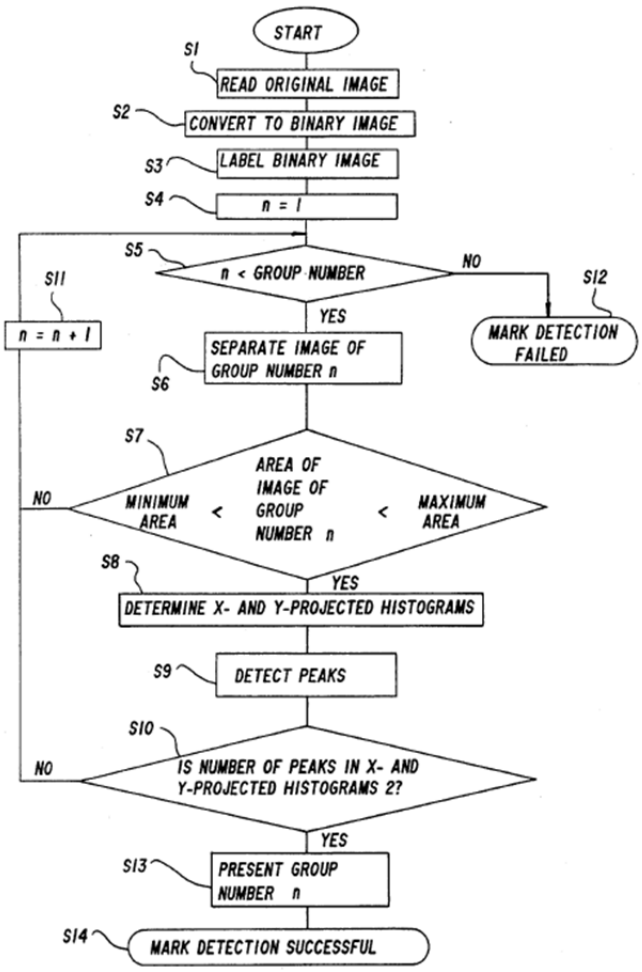


FIG.5

Once the target mark is detected, Hashima determines the center of the detected mark from the X- and Y-maxima and minima of the X- and Y-histograms:

> The image processor 40…generat[es] a histogram 15 projected onto the X-axis and a histogram 16 projected onto the Y-axis, and then determines the X and Y coordinates of the central position Pm (mx, my) from the projected histograms 15, 16. Specifically, the X

coordinate mx of the central position Pm (mx, my) can be determined using opposite end positions Xb1, Xb2 obtained from the X-projected histogram 15 according to the following equation (3):

$$mx=(Xb1+Xb2)/2 \quad (3).$$

The Y coordinate my of the central position Pm (mx, my) can be determined using opposite end positions Yb1, Yb2 obtained from the Y-projected histogram 16 according to the following equation (4):

$$my=(Yb1+Yb2)/2 \quad (4).$$

Id. at 11:6-25; Ex. 1002, ¶¶65-66. Figure 15 illustrates the process for finding the center position of the detected target:

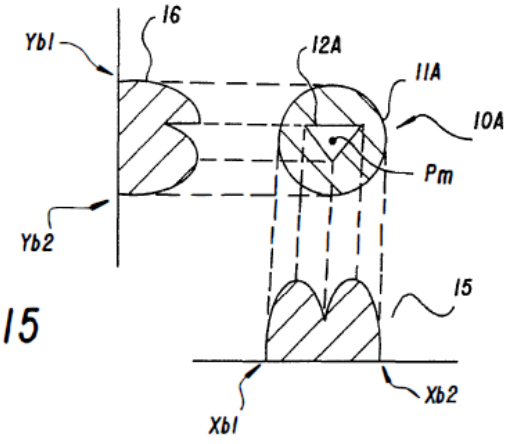


FIG.15

The center of the target is then compared to the center of the image memory from the previous frame to find the shift of the target in the X- and Y-directions. The shift amount is used to move the robot arm (and the camera mounted thereon)

toward the target, and is recalculated based on new images as the robot arm and the camera move toward the target, until the object is gripped. Ex. 1002, ¶¶64, 67; Ex. 1006, 14:61-15:37. Figure 27 shows a flow chart describing the process:

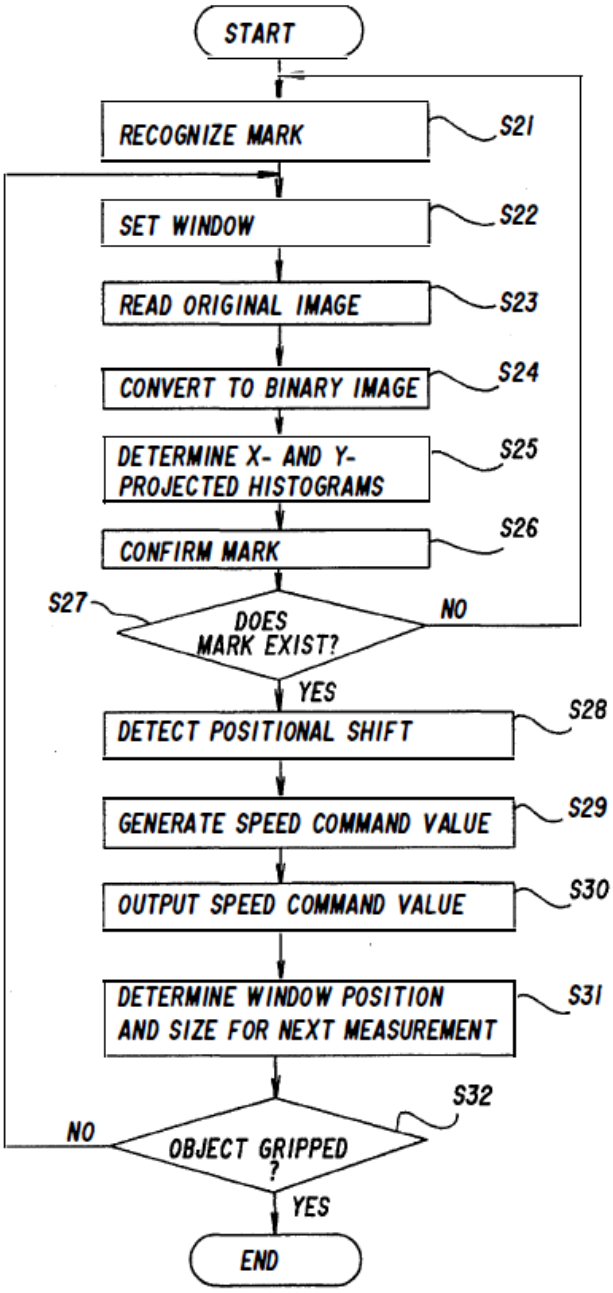
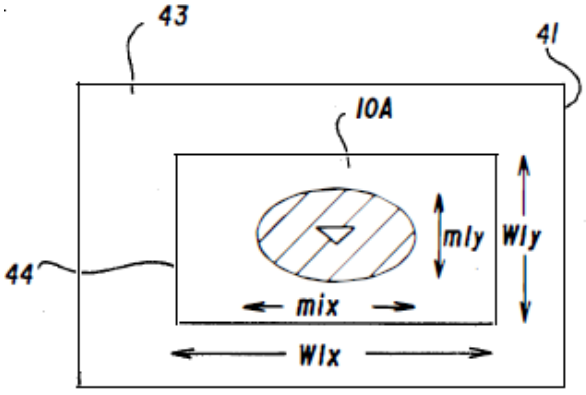


FIG.27

As the target is tracked, Hashima displays a rectangular window around the target—i.e., a tracking box. Ex. 1006, 14:29-34, Figure 23. As the system receives new frames from the video camera, the target and window locations are recalculated using the new histograms created from each new frame:

> When the target mark **10** starts to be tracked, the window is established using the projected histogram information obtained when the target mark image is recognized. When the target mark **10** is subsequently tracked, the window [**44**] is established using new projected histogram information obtained upon each measurement made by the camera **20**.

Id. at 14:29-34; Ex. 1002, ¶68. Figure 23 shows the target window **44** around the target mark **10A**.



Hashima Fig. 23

A block diagram of the system in Hashima shows that after determining an appropriate window, the image is displayed on a monitor, 315. *See* also, Ex. 1007, Hashima at 25:34–38, Figure 52; Ex. 1002, ¶69.

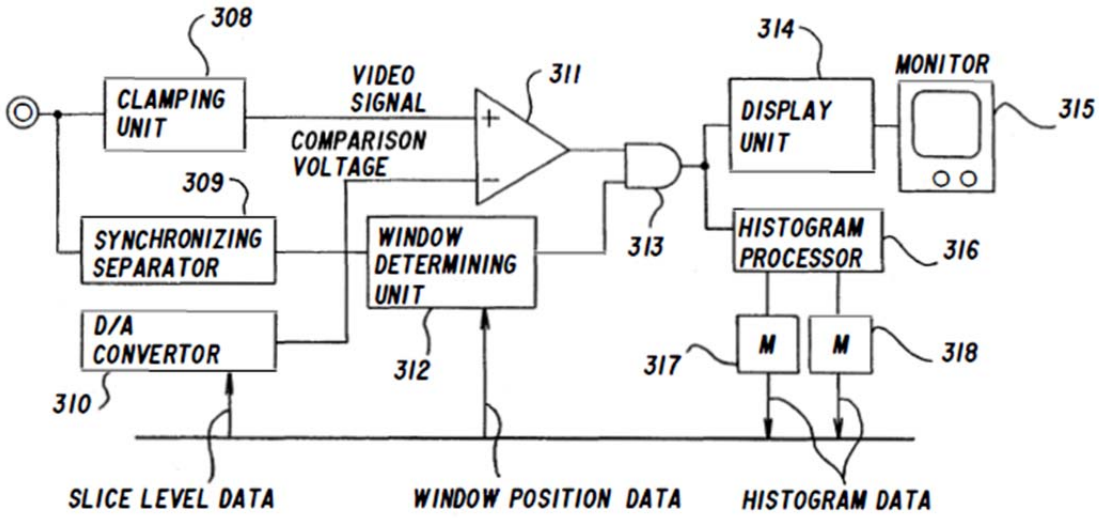


FIG.53

3. U.S. Patent No. 5,150,432 ("Ueno") (Ex. 1007)

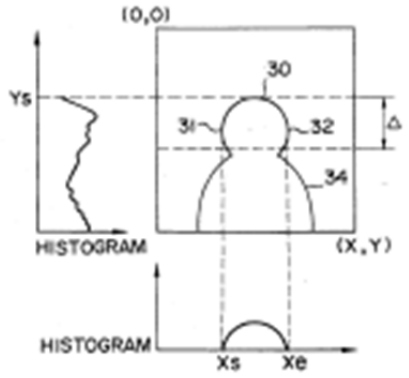
A similar process and apparatus is described in Ueno. Ueno uses histograms in X- and Y-domains to identify the facial region in the frames of video signals, and tracks the facial region with a rectangle constructed from X- and Y-minima and maxima derived from the histograms. Ex. 1007, 1:39-68; Ex. 1002, ¶71. Ueno qualifies as prior art under pre-AIA § 102(b). Ueno was not of record and was not considered during prosecution of the '001 Patent.

Ueno describes a video encoding system that includes "frame memory **101** for storing image signals for one frame" connected to "a facial region detecting

circuit **102** for detecting human facial region of the frame data…." Ex. 1007, 3:5-7; Ex. 1002, ¶72. In Ueno, "video signals are sequentially input to the image input terminal **100** in frames." Ex. 1007, 4:25-26. "The image signal is written in the frame memory **101** as frame data one frame by one frame. The frame data written in the frame memory **101**…is supplied to the facial region detecting circuit **102**…." Ex. 1007, 4:33-37. The facial region detecting circuit of Ueno uses differences between two consecutive frames and creates "interframe difference images." Ueno then constructs histograms in X- and Y-domains using the interframe difference images. Ex. 1002, ¶¶72-73.

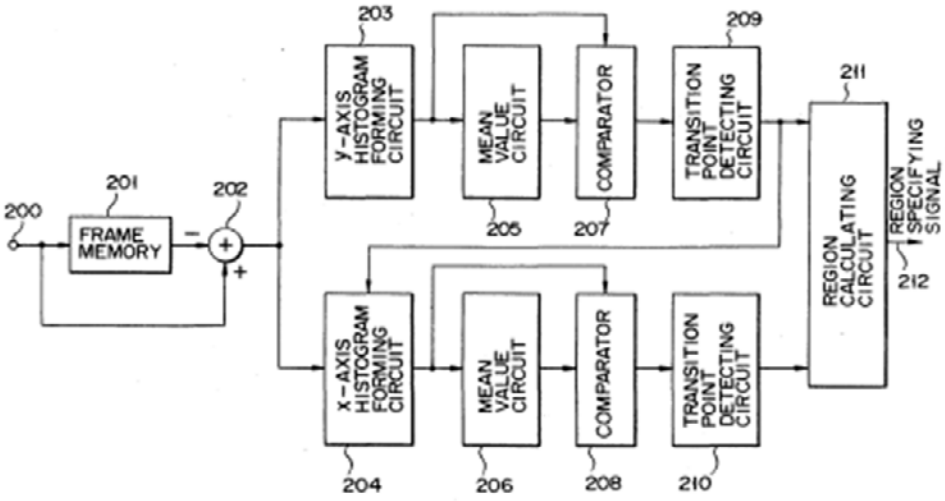
> FIG. 3 shows an interframe difference image input to the facial region detecting circuit 102. Information for the interframe difference image is converted into binary data of "0" or "1" by the preset first threshold level. Then the number of pixels having the binary data value of "1" or the value equal to or more than the first threshold value is counted in the vertical and horizontal directions of the screen and histograms of the pixels (x-and y-axis histograms) are formed. Facial detection is executed according to the histograms.

Ex. 1007, 7:7-16.



F I G. 3

Figure 6 shows block diagrams of the parallel histogram processors that create the X- and Y-histograms. Ex. 1007, 8:22-49; Ex. 1002, ¶74.



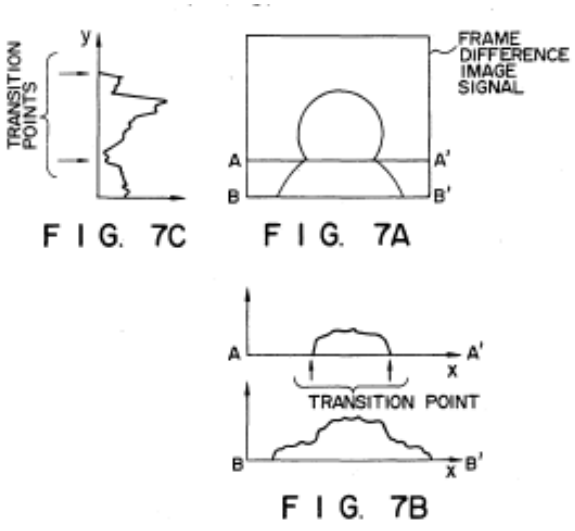
F I G. 6

Ueno uses these X- and Y-histograms to determine where the facial region of the image starts and ends:

> Face detection starts with the top of a head 30. The top of the head 30 is detected by searching the y-axis

> histogram from the top and selecting the point Ys which first exceeds the present second threshold value.
>
> After the top of the head 30 is detected, the left end 31 and right end 32 of the head 30 are detected by searching the x-axis histogram from the left and right and selecting the points Xs and Xe which first exceed the second threshold value….
>
> Then, the bottom end of the facial region is decided. To decide the bottom end of it, the head hight [sic] is obtained by multiplying the head width by a certain ratio r. The r values of 1.3 to 1.6 are proper. The bottom end of the facial region is obtained by adding the head length to the value for the top of the head previously detected. The facial region is specified by a rectangle 50 designated by coordinates Xs, Xe, Ys, and Ye as shown in FIG. 5.

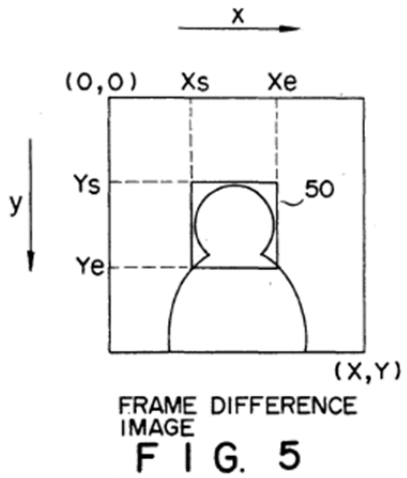
Ex. 1007, 7:17-45. The use of these histograms in identifying and tracking the target is further seen in Ueno Figures 3 (reproduced above) and 7 (reproduced below):



F I G. 7C F I G. 7A

F I G. 7B

Once a target is identified, Ueno constructs a tracking box around the target—"[b]y extracting the coordinates of the transition points of these histograms, the face region is specified by a rectangle as shown in Fig. 5." Ex. 1007, 8:9-11; Ex. 1002, ¶¶74-75.

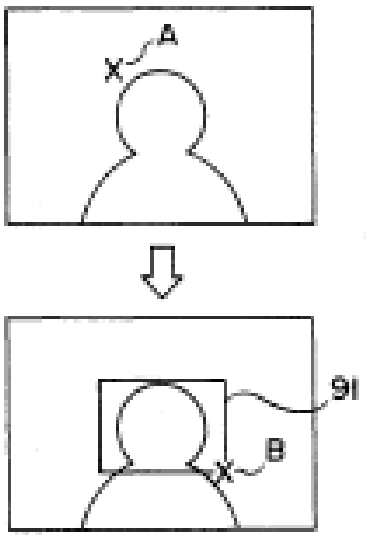


F I G. 5

This tracking box is then displayed on a monitor and continuously tracks the moving target:

> In the self-monitoring mode, low-pass-filtered images are displayed on the monitoring display 710. After a region is specified by the region specifying apparatus 120, the region specifying signal is continuously output. Thus, said frame line showing the specified region is superimposed on the image and displayed on the monitoring display 710.

Ex. 1007, 13:17-23; Ex. 1002, ¶76. Examples of the tracking boxes taught by Ueno are shown in Figures 5 (above) and 18, reproduced below.



F I G. 18

IX. Specific Explanation Of Grounds For Invalidity

A. Ground 1: Gilbert In View Of Hashima Renders Obvious Claims 1-4

1. Reasons To Combine Gilbert And Hashima

A POSA would have been motivated to combine Gilbert and Hashima because they are directed toward very similar systems that operate in a similar manner. As set forth above, both Gilbert and Hashima are directed to identifying and tracking a target using an image processing system with a video camera input, (Ex. 1005, 47-48, Figure 1; Ex. 1006, 1:6-15, 2:45-3:12, Figure 1; Ex. 1002, ¶79), where the video input comprises a succession of frames, each frame comprising a succession of pixels (Ex. 1005, 47-48, 52; Ex. 1006, 10:34-45, 24:13-25:33; Ex. 1002, ¶79), in which the target comprises pixels in one or more of a plurality of classes and one or more of a plurality of domains (Ex. 1005, 48; Ex. 1006, 8:18-30, 10:34-45; Ex. 1002, ¶79). The Gilbert and Hashima systems perform this process on a frame-by-frame basis (Ex. 1005, 47-48, 52; Ex. 1006, 24:13-25:23; Ex. 1002, ¶79).

Additionally, both Gilbert and Hashima identify and track a moving target using histograms (Ex. 1005, 48-49; Ex. 1006, 4:8-47; Ex. 1002, ¶80), employ a camera that keeps the target in the field of view (Ex. 1005, 52; Ex. 1006, 3:41-58; Ex. 1002, ¶80), and recognize the importance of finding the center of a target (Ex. 1005, 50-51, Figure 4; Ex. 1006, 11:1-30; Ex. 1002, ¶80). Thus, they are not only

in the same field of endeavor but also are directed to very similar systems and processes. Ex. 1002, ¶80. A POSA would have recognized that Hashima's invention could improve a similar device, such as the system in Gilbert, in the same way, and thus would have been motivated to combine Gilbert and Hashima. Ex. 1002, ¶81. *See KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398 (2007) ("if a technique has been used to improve one device, and a person of ordinary skill in the art would recognize that it would improve similar devices in the same way, using the technique is obvious unless its actual application is beyond his or her skill."). A POSA would have also expected the combination to yield a predictable result because it would have involved applying known techniques to similar systems. Ex. 1002, ¶82.

Moreover, the references themselves, taken as a whole, demonstrate reasons why a POSA would be motivated to combine Gilbert with Hashima. *See In re Beattie*, 974 F.2d 1309, 1312 (Fed. Cir. 1992) ("As long as some motivation or suggestion to combine the references is provided by the prior art taken as a whole, the law does not require the references be combined for the reasons contemplated by the invention."). For example, Gilbert recognizes "that there are many features that can be functionally derived from relationships between pixels" but notes that it has chosen to use the domain of intensity (i.e., gray scale) for its histograms to identify the target. Ex. 1005, 48; Ex. 1002, ¶83. Gilbert also discloses the

construction of projections to determine the position of the target. These projections are entirely analogous to the projection histograms in the X- and Y-domain in Hashima (and entirely analogous to the position histograms in the '001 Patent at 20:54-21:29).

Gilbert also uses X- and Y-projection histograms to identify the center of the target, and its orientation, after the pixels associated with the target have been separated from the plume and background. A POSA reading Gilbert's disclosure would have recognized that plotting histograms in other domains would result in certain advantages over the disclosure in Gilbert. Ex. 1002, ¶84. For example, by having the capability of plotting histograms in additional domains, the system would have a higher likelihood of successfully recognizing a target in an image frame. This is because each additional domain provides another opportunity for correlation between that domain and the unique target being tracked. Additionally, by computing histograms across multiple domains simultaneously, a POSA would have been able to utilize the existing design of the histogram component of the "video processor" to increase system effectiveness.

A POSA reading Gilbert would look to improve upon Gilbert by constructing histograms in other domains and, in so doing, would look to references that teach using histograms in other domains. Ex. 1002, ¶85. Hashima is one such reference—it creates histograms based on the X- and Y- domains to

identify the target. Ex. 1006, 4:8-47. Thus, to identify the target, and improve on processing speed, a POSA would have been motivated to combine Gilbert's image processing method with Hashima to replace Gilbert's intensity histograms with the X- and Y-domain histograms disclosed in Hashima. Ex. 1002, ¶¶84-85. Moreover, there would have been a reasonable expectation of success that in doing so a more accurate target identification system would have resulted.

Gilbert teaches a center calculation that involves splitting the target into two halves, calculating the center for each of the top and the bottom halves, and then finding the mid-point between the top center point and the bottom center point. Ex. 1005, 50-51; Ex. 1002, ¶86. Gilbert uses this center calculation to keep the target centered within the frame, making automatic adjustments to the direction of the camera to achieve this end. Ex. 1005, 52; Ex. 1002, ¶86. But Gilbert teaches a complex method of determining a center and orientation that involves many computational steps. Gilbert teaches the center calculation using center of area in part to allow for a determination of the rocket's orientation in order to predict its trajectory. Ex. 1002, ¶86; Ex. 1005, 50-51. Gilbert also uses the center calculated using the two points to direct the lens toward the moving target such that the target remains in the camera's field of view. Ex. 1005, 52; Ex. 1002, ¶86. However, a POSA at the time of the '001 Patent would have recognized that a simpler center calculation for the entire target image would have been able to achieve all of the

benefits of Gilbert, while imposing a significantly smaller computational burden. Ex. 1002, ¶86. For example, by calculating the center point of the target over at least two frames, the velocity and orientation can be calculated more efficiently by simply using the time elapsed between any two frames and calculating the distance travelled by the target from the change in location of their respective center points from the first frame to the second. Ex. 1002, ¶86. Moreover, a POSA would have recognized that this improvement would also allow the Gilbert system to track other similar objects with different shapes because the center of area calculation disclosed by Gilbert becomes unreliable for determining orientation when tracking objects with differing shapes. Ex. 1002, ¶87.

Thus, it would have been obvious for a POSA to combine Gilbert with other references teaching simpler center calculations based on the end-points of a target in a histogram (i.e., X-minima and maxima and Y-minima and maxima), such as the calculation disclosed by Hashima. Ex. 1002, ¶88. A POSA would be motivated to make such a combination because it would reduce the number of calculations needed to determine the center, resulting in faster, more efficient processing to improve tracking of a target. Ex. 1002, ¶89. These improvements could also reduce hardware expenses and result in more efficient use of circuit die space. *Id.* Moreover, there would have been a reasonable expectation of success that in doing so a more efficient computation system would have resulted. *Id.*

A detailed comparison will now be made of Gilbert in view of Hashima to the asserted claims of the '001 Patent.

2. Claim 1

a. Gilbert and Hashima each disclose: "a process of tracking a target, in an input signal implemented using a system comprising an image processing system, the input signal comprising a succession of frames, each frame comprising a succession of pixels, the target comprising pixels in one or more of a plurality of classes in one or more of a plurality of domains, the process performed by said system comprising, on a frame-by-frame basis" (1[pre])

Gilbert discloses a process of tracking a target (a missile). Ex. 1005, 47; Ex. 1002, ¶¶90-91. Gilbert's tracking system uses a video signal (input signal) comprising digitized fields with frame rate of 60 fields/s (i.e., 30 frames/s in a succession of frames), each image frame comprising an $n \times m$ matrix of digitized points (a succession of pixels). Ex. 1005, 48. Gilbert tracks the image of the target (which comprises pixels in the video signal), by categorizing the pixels into 256 gray-scale levels (a plurality of classes) according to their pixel intensity (domain). *See* Ex. 1005, 48. Although Gilbert uses pixel intensity, Gilbert discloses that a plurality of other domains, such as "texture, edge, and linearity measures" can also be used. *Id.* Gilbert performs the target tracking process "during each field," i.e., on a frame-by-frame basis. *Id.*; Ex. 1002, ¶93.

Hashima similarly discloses "[a] system for and a method of recognizing and tracking a target mark with a video camera." Ex. 1006, 1:7-9. Hashima's input signal is a digitized video signal, which comprises a sequence of frames. Ex. 1006, 2:45-3:13; Ex. 1002, ¶92. Hashima also teaches that "one frame of scanned image data from a camera is stored in a frame memory," (Ex. 1006, 24:14-17) and that "[a]n image…is read into an image memory…[which] is composed of 512 pixels in each of the X- and Y-directions." Ex. 1006, 10:43-45. The pixels in the input image in Hashima are categorized into classes within the X- and Y-domain. Ex. 1006, 8:27-31; Ex. 1002, ¶93. Hashima's process is performed on a frame-by-frame basis. Ex. 1006, 24:24-25 ("it is necessary to write and rea[d] one frame of image data each time it is supplied from the camera.").

b. Gilbert and Hashima each disclose the step of "forming at least one histogram of the pixels in the one or more of a plurality of classes in the one or more of a plurality of domains, said at least one histogram referring to classes defining said target" (1[a])

In both Gilbert and Hashima, at least one histogram is formed based on pixel characteristics (or domains) such as intensity or coordinate axis. Ex. 1002, ¶95. Further, target pixels are categorized according to classes within the chosen domain—for example, one of 256 gray levels for intensity histograms (Gilbert) or position along a relevant axis for X- and Y-projection histograms (Hashima). Ex. 1005, 48; Ex. 1006, 24:14-21; Ex. 1002, ¶¶95-98. These X- and Y- projection

histograms in Hashima are analogous to the X- and Y-position histograms discussed in the '001 Patent at 20:60-21:29. Ex. 1002, ¶98.

Hashima also explicitly discloses a plurality of domains by showing histograms in the X- and Y-domains. Ex. 1002, ¶98. Moreover, to the extent the Patent Owner argues that Gilbert's intensity histogram does not include a plurality of "classes" because the histogram contains every pixel (sorted into bins based on intensity level) rather than a subset of pixels that meet a threshold intensity characteristic, this argument should be rejected. Gilbert's intensity histogram is entirely analogous to the velocity histogram in Figure 14a of the '001 Patent, as described at 20:54-59. Ex. 1002, ¶96.

Moreover, Gilbert also discloses a projection processor that forms X- and Y-projection histograms using binary pictures generated by the video processor. Ex. 1005, Figure 4; Ex. 1002, ¶99. These projection histograms are analogous to the X- and Y-position histograms discussed in the '001 Patent at 20:60-21:29. Ex. 1002, ¶99. Thus, Gilbert discloses creating histograms in a plurality of domains, including the intensity, the X-, and the Y-domains. Ex. 1002, ¶99.

In these histograms, each "class" within a domain counts the number of pixels meeting certain threshold criteria. For example, Gilbert's intensity histogram has 256 grayscale levels (i.e., 256 classes). Ex. 1005, 48. The histogram is formed as each pixel in the frame is sorted into whichever of the 256

classes that pixel's intensity level fits. Ex. 1002, ¶101. In Hashima, the number of pixels having the value "1" (which indicates that a pixel meets a certain intensity threshold) within a given column defined by an X- coordinate is counted into the class that corresponds to that X-coordinate value. Ex. 1006, 8:22-30; Ex. 1002, ¶101. Similarly, in the X- and Y-projection histograms of Gilbert Figure 4, the number of pixels having the value "1" in the binary picture (which indicates that a pixel belongs to the target) within a given column defined by an X- or Y-coordinate is counted into the class that corresponds to that X- or Y-coordinate value. Ex. 1005, 50-51; Ex. 1002, ¶101. These binary selection processes in Gilbert and Hashima are entirely analogous to the histogram formation process discussed in the '001 Patent at 18:16-19:24. Ex. 1002, ¶98-99.

c. Gilbert and Hashima each disclose the step of "identifying the target in said at least one histogram itself" (1[b])

In both Gilbert and Hashima, histograms are used to identify the target. For example, Gilbert uses probability estimates based on the 256 level grayscale histograms to determine whether a particular pixel belongs to the target, plume, or background region. Ex. 1005, 48-50; Ex. 1002, ¶102. Similarly, in Hashima, peaks and valleys in the X- and Y-axis histograms are used to determine the location of the target. Ex. 1006, 8:18-10:24; Ex. 1002, ¶103. The '001 Patent

describes an identical process for identifying a face from the histograms in Figure 17 and the accompanying description at 22:60-23:14. Ex. 1002, ¶103.

d. Gilbert and Hashima each disclose the step of "wherein identifying the target in said at least one histogram further comprises determining a center of the target to be between X and Y minima and maxima of the target" (1[c])

Gilbert determines a center by computing the center of areas of the target's upper and lower halves. Ex. 1005, 50-51, Figure 4. However, Gilbert also teaches that "target nose and tail points" could be used. Ex. 1005, 50; Ex. 1002, ¶104. Hashima also explicitly discloses determining X-maxima and minima and Y-maxima and minima of the boundaries of the target in the histogram to calculate the center point of the target using the equations $(X_{MIN} + X_{MAX})/2$ and $(Y_{MIN} + Y_{MAX})/2$. Ex. 1006, 11:13-24, Figure 15; Ex. 1002, ¶¶105-107. The equations used to find the center in Hashima are identical to those described for the same purpose in the '001 Patent at 24:50-55. Ex. 1002, ¶105.

e. Gilbert in view of Hashima renders Claim 1 obvious

To the extent Patent Owner argues Gilbert does not disclose the claimed plurality of domains or method of finding the center of the target, it would have been obvious to replace these features of Gilbert with those of Hashima, as discussed in Section **IX.A.2.b**, *supra*; Ex. 1002, ¶101.

Patent Owner may argue that Hashima does not expressly disclose "an input signal comprising a succession of frames, each frame comprising a succession of pixels" (see claim limitation 1[pre]) because Hashima only explicitly discusses the input signal and the resulting output at a high level. But Hashima inherently discloses this input signal when it discloses an input from a video camera and the generation of histograms based on individual pixel location, (Ex. 1006, 1:7-9; 8:22-30; 10:37-44; 24:15-16; 24:24-25) because the input signal must comprise a succession of frames, each frame comprising a succession of pixels in order to form such histograms on a frame-by-frame basis. Ex. 1002, ¶94 (1[pre]). Furthermore, Gilbert explicitly describes an input signal comprising a succession of frames, each frame comprising a succession of pixels (Ex. 1005, 47-48, 52). Thus, the combination of Gilbert and Hashima renders Claim 1 obvious. Ex. 1002, ¶¶78-107, 115.

3. Claim 2: "The process according to claim 1, wherein identifying the target in said at least one histogram further comprises determining a center of the target to be between X and Y minima and maxima of the target"

Claim 2 contains a limitation identical to the limitation in 1[c] already discussed. For the same reasons, Gilbert and Hashima both disclose this limitation and render Claim 2 obvious. *See* Ex. 1002, ¶¶108, 115.

4. Claim 3: "The process according to claim 1, wherein identifying the target in said at least one histogram further comprises determining the center of the target at regular intervals"

Gilbert discloses forming X- and Y-projection histograms on a frame-by-frame basis at a rate of 60 fields/second (i.e., 30 frames/second). *See, e.g.*, Ex. 1005, 50 ("In the projection processor, these matrices are analyzed field-by-field at 60 fields/s [i.e., 30 frames/s] using the projection based algorithm"); Ex. 1002, ¶109; Ex. 1005, 55 ("The result after six processing frames…"). Because part of this analysis involves generating a histogram to identify the target, and further finding the center of the target using the generated histogram, this process is necessarily performed at regular intervals, namely once per frame or once every 1/30 second. Ex. 1002, ¶109.

Hashima similarly discloses processing images on a frame-by-frame basis as they are supplied from the camera. *See, e.g.,* Ex. 1006, 24:13-35 ("When the above general-purpose image processor is employed, it is necessary to write and rea[d] one frame of 25 image data each time it is supplied from the camera, a window is established, the data is converted into binary image data, and projected histograms are generated."). Hashima discloses embodiments where this process requires "at least 200 ms," *id.*, and other embodiments where the "processing time corresponds to only one frame (33 ms)." Ex. 1006, 25:16-33. Thus, because determining the center of the target is part of the tracking process in both

embodiments, it is done at the regular interval of 200 ms or on a frame-by-frame basis every 33 ms. Ex. 1006, 25:16-33. Ex. 1002, ¶110. The frame-by-frame determination of the center of the target in Gilbert and Hashima is identical to the description in the '001 Patent of determining the center position at regular intervals at 25:3-25. *See* Ex. 1001, 25:3-25; Ex. 1002, ¶111.

To the extent Patent Owner argues Gilbert or Hashima does not explicitly disclose that the center of the target is determined at regular intervals, it would have been obvious to implement this limitation in order to achieve smooth, predictable processing. Ex. 1002, ¶111. Thus, Gilbert and Hashima both disclose this limitation and render it obvious. Ex. 1002, ¶¶109-111, 115.

5. Claim 4: "The process according to claim 1 further comprising drawing a tracking box around the target"

Gilbert discloses the step of drawing a tracking box around the target. *See, e.g.*, Ex. 1005, 48 ("A tracking window is placed about the target image,…The tracking window frame is partitioned into a background region (BR) and a plume region (PR). The region inside the frame is called the target region (TR) as shown in Fig. 2."). This tracking window is shown in Ex. 1005, Figure 2 as annotated below:

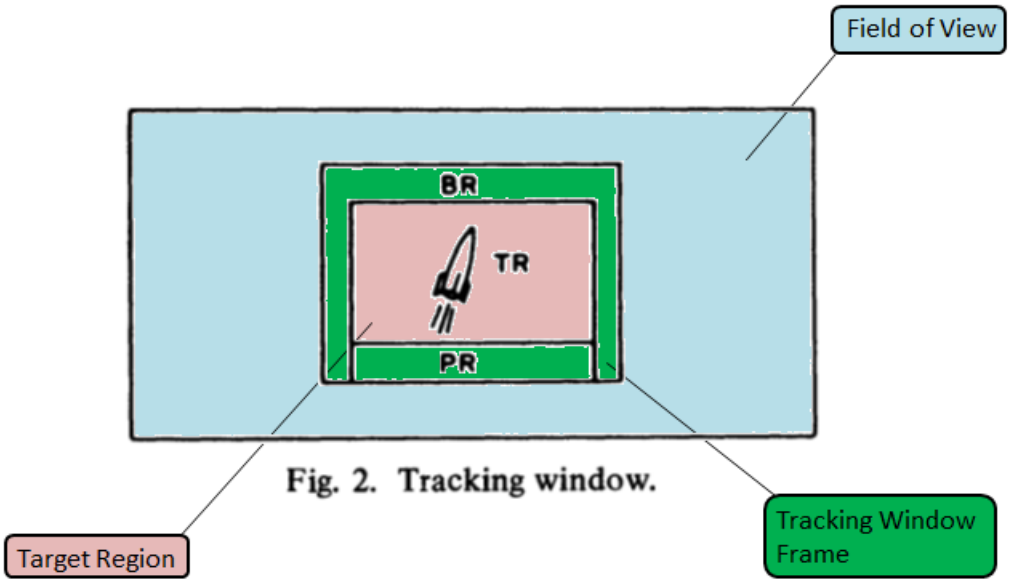


Fig. 2. Tracking window.

In simulation outputs (using pre-recorded video of a rocket in flight as an input signal),, the system similarly tracks the target by drawing a tracking box. *See, e.g.*, Ex. 1005, 55 ("Fig. 7 contains two representative simulation outputs selected from the first 100 fields of simulated digitized video."). This is shown in Ex. 1005, Figure 7, where a rocket is being tracked as annotated below:

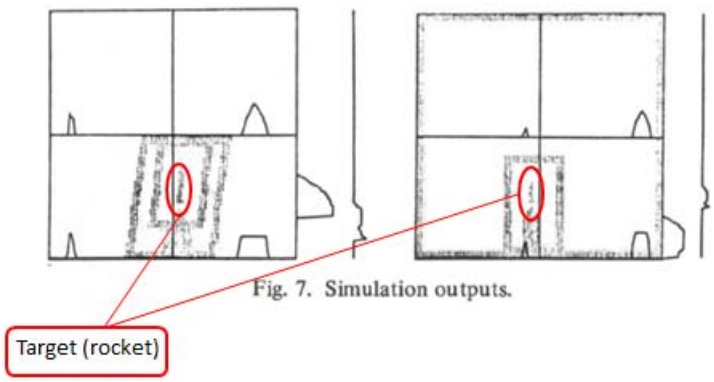


Fig. 7. Simulation outputs.

Ex. 1002, ¶112.

Thus, Gilbert discloses this limitation. Ex. 1002, ¶113. Hashima also discloses a tracking box around the target: "each of the images grouped by the labeling process is separated by a window process." Hashima at 8:55-57. This window is "established" by Hashima using the technique described at 13:61-14:34. Furthermore, the system disclosed in Figure 53 described by Hashima at 24:44-25:23 includes a "window determining unit 312" that uses the "window position data" to restrict the video signal to only include the window contents through the "AND gate 313" and out to the "display unit 314" to the "monitor 315." In this configuration, only the contents of the window would be displayed on the monitor and the remainder of the screen would be black, thereby drawing a (black) tracking box around the target currently under consideration. Ex. 1002, ¶114. Thus, it would have been obvious to implement this limitation in order to aid a user at a display to visually locate and follow a target. Ex. 1002, ¶¶112-115.

6. Gilbert And Hashima Are Not Cumulative

While Gilbert and Hashima both disclose image processing systems and methods very similar to those claimed in the '001 Patent, they are not cumulative of one another. Each compliments the other such that strong disclosures of one make up for weaknesses in the other. For example, Gilbert determines a center of the target by finding a center of area point in both the top and bottom portions of the target and using these calculations to determine the center of the target. Ex.

1005, 50-51, Figure 4; Ex. 1002, ¶104. Gilbert also recognizes the possibility of using the X- and Y-minima and maxima to find the center, but Gilbert elects to use an alternative to make this calculation. Ex. 1002, ¶104; Ex. 1005, 50. To the extent Patent Owner argues that Gilbert does not disclose the step of "determining a center of the target to be between X- and Y-minima and maxima of the target," Hashima discloses this limitation. Ex. 1006, 11:13-24, Figure 15. Hashima describes precisely the same center calculation used in the '001 Patent—namely a center calculation that uses the X- and Y-minima and maxima determined from the histogram to calculate the center of the target. Ex. 1002, ¶¶105-106.

Similarly, Gilbert generally teaches the step of forming a histogram in one or more of a plurality of domains, such as the intensity (luminance) domain used to construct the intensity histogram, or the X- and Y-domains used to construct the projection histograms. Ex. 1005, 48-51, Ex. 1002, ¶99. In addition, Hashima explicitly teaches identifying the target using both X- and Y-histograms. Ex. 1006, 24:14-21; Ex. 1002, ¶¶99-101.

To the extent the Patent Owner argues that the only histograms disclosed in Gilbert are the histograms formed in the intensity domain (i.e., that the X- and Y-projections in Gilbert are not "histograms"), Hashima makes up for this purported weakness because it discloses creating histograms in the X-axis and Y-axis domains (i.e., a plurality of domains). Ex. 1006, 24:14-21; Ex. 1002, ¶¶98-100.

Nevertheless, Gilbert also teaches that "[t]here are many feature that can be functionally derived from relationships between pixels, e.g., texture, edge, and linearity measures." Ex. 1005, 48; Ex. 1002, ¶97. A POSA would have understood that Gilbert thus explains that other histograms based on domains of texture, edge, or linearity measures may be used to achieve similar result. Ex. 1002, ¶85.

As another example, Hashima discloses imaging a target mark using a video camera, reading it into an image memory, creating a pixel map in the image memory, creating histograms based on the X- and Y-coordinates of individual pixels, and performing this process on the image data one frame at a time, but it does not explicitly disclose an input signal comprising a succession of frames, each frame comprising a succession of pixels. Ex. 1005, 48; Ex. 1002, ¶92. Of course a video camera input that is processed one frame at a time to create a histogram based on pixel location necessarily receives an input comprising a succession of frames, each frame comprising a succession of pixels, so this limitation is inherently disclosed by Hashima. Ex. 1002, ¶92. On the other hand, this limitation is explicitly disclosed by Gilbert, and the combination of Gilbert and Hashima, therefore, strengthens the overall disclosure. Ex. 1002, ¶92.

Thus, while each of Gilbert and Hashima renders the challenged claims of the '001 Patent obvious in combination with the knowledge of a POSA, the

combination of Gilbert and Hashima provides a more complete disclosure, and Petitioner urges the Board to adopt this ground as presented because Gilbert and Hashima are not cumulative. *See* Ex. 1002, ¶115.

B. Ground 2: Hashima In View Of Ueno Renders Obvious Claims 1-4

1. Reasons To Combine Hashima And Ueno

A POSA would have been motivated to combine Hashima and Ueno because they are directed toward very similar systems that operate in a similar manner. As set forth above, both Hashima and Ueno are directed to identifying and tracking a target using an image processing system with a video camera input, (Ex. 1006, 1:6-15, 2:45-3:12, Figure 1; Ex. 1007, 3:1-8, 4:25-41), where the video input comprises a succession of frames, each frame comprising a succession of pixels (Ex. 1006, 10:34-45, 24:13-25:33; Ex. 1007, 1:44-49, 4:25-41), in which the target comprises pixels in one or more of a plurality of classes and one or more of a plurality of domains (Ex. 1006, 8:18-30, 10:34-45; Ex. 1007, 7:8-16), and performing the process on a frame-by-frame basis (Ex. 1006, 24:13-25:23; Ex. 1007, 4:25-41, 5:31-40). *See also* Ex. 1002, ¶117.

Additionally, both Hashima and Ueno identify and track a moving target using histograms. Ex. 1006, 4:8-47; Ex. 1007, 7:7-8:22; Ex. 1002, ¶118. Thus, they are not only in the same field of endeavor but also are directed to very similar systems and processes. Ex. 1002, ¶¶118-119. A POSA would have recognized

that Ueno's invention could improve a similar device, such as the system in Hashima, in the same way, and thus would have been motivated to combine Hashima and Ueno *See KSR Int'l Co. v. Teleflex Inc*., 550 U.S. 398 (2007) ("if a technique has been used to improve one device, and a person of ordinary skill in the art would recognize that it would improve similar devices in the same way, using the technique is obvious unless its actual application is beyond his or her skill."). A POSA would have also expected the combination to yield a predictable result because it would have involved applying known techniques to similar systems. Ex. 1002, ¶120.

Moreover, the references themselves, taken as a whole, demonstrate reasons why a POSA would be motivated to combine Hashima with Ueno. *See In re Beattie*, 974 F.2d 1309, 1312 (Fed. Cir. 1992) ("As long as some motivation or suggestion to combine the references is provided by the prior art taken as a whole, the law does not require the references be combined for the reasons contemplated by the invention."). In particular, although Hashima describes a system that identifies and tracks a target and projects the tracked object on a display for viewing by a user, it does not teach a system that highlights the target for easier visualization by a user. However, being able to visualize a target more easily on a screen provides numerous advantages. Ex. 1002, ¶121. For example, a tracking box enables the user to visually confirm that the correct object has been identified

and is being tracked by the image processing system. Ex. 1002, ¶121. Additionally, the user may use a mouse or keyboard input to specify a target mark to be tracked. Ex. 1002, ¶121. A tracking box allows the user to more easily pick out the target object in a situation where noise or other objects in the field of view otherwise make it more difficult to visually identify the target. Ex. 1002, ¶121. For these reasons, a POSA would be motivated to combine Hashima with Ueno, which teaches a tracking box and allows a user to select a target using an input to the image processing system. Ex. 1002, ¶121. Moreover, there would have been a reasonable expectation of success that in doing so a more user friendly system would have resulted. Ex. 1002, ¶121.

A detailed comparison will now be made of Hashima in view of Ueno to the asserted claims of the '001 Patent.

2. Claim 1

a. Hashima and Ueno each disclose: "a process of tracking a target…on a frame-by-frame basis" (1[pre])

Hashima discloses a process of tracking a target mark from a video signal (an input signal comprising a succession of frames, each frame comprising a succession of pixels). Ex. 1006, 1:7-15; Ex. 1002, ¶123-124. The target mark in the video signal (which comprises pixels in the video signal) is identified using the number of peaks and valleys in the histograms of the target mark in the X- and Y-domains. Ex. 1006, 8:18-9:40; Ex. 1002, ¶¶124, 126. Hashima repeats this

process "each time [a new frame of image data] is supplied from the camera," i.e., on a frame-by-frame basis. Ex. 1006, 24:24-25; Ex. 1002, ¶128. Ueno also discloses a process of tracking a human face (target) from a video signal (an input signal comprising a succession of frames, each frame comprising a succession of pixels). Ex. 1002, ¶125. The image of human face (which, like Hashima, comprises pixels in the video signal) is identified in the histograms of inter-frame difference image in the X- and Y-domains. Ex. 1007, 7:7-16; Ex. 1002, ¶127. Ueno also constructs the histogram on a field-by-field basis. Ex. 1007, 1-44-48; Ex. 1002, ¶128.

b. Hashima and Ueno each disclose the step of "forming at least one histogram…referring to classes defining said target" (1[a])

In both Hashima and Ueno, at least one histogram is formed based on pixel characteristics (domains) such as intensity or coordinate axis. Ex. 1002, ¶129. For example, in Hashima, an image processor converts an image from a camera into binary image data and generates X- and Y-projected histograms from the binary data. Ex. 1006, 8:27-30; Ex. 1002, ¶129. In Ueno, a similar process occurs, but the histograms are generated on interframe images and specifically identify pixels that are different in the two frames. Ex. 1007, 7:7-16; Ex. 1002, ¶130. These X- and Y- projection histograms in Hashima and Ueno are analogous to the X- and Y-position histograms discussed in the '001 Patent at 20:60-21:29. Ex. 1002, ¶131.

In these histograms, each "class" within a domain counts the number of pixels meeting certain threshold criteria. Ex. 1002, ¶129. For example, in both Hashima and Ueno, the number of pixels having a value of "1" within a given column defined by an X-coordinate is counted into the class that corresponds to that X-coordinate value. Ex. 1006, 8:27-30; Ex. 1007, 7:7-16; Ex. 1002, ¶130. These binary selection processes in Hashima and Ueno are entirely analogous to the histogram formation process discussed in the '001 Patent at 18:16-19:24. In each case, at least one histogram is formed based on pixel characteristics (or domains) such as coordinate axis. Ex. 1002, ¶¶129-131.

c. Hashima and Ueno each disclose the step of "identifying the target in said at least one histogram itself" (1[b])

In both Hashima and Ueno, histograms are used to identify the target. For example, in Hashima, peaks and valleys in the X- and Y-axis histograms are used to determine the location of the target. Ex. 1006, 10:20-24; Ex. 1002, ¶132. Similarly, in Ueno, interframe differences are formed into X- and Y-histograms and used to identify a human face. Ex. 1007, 7:7-16, Figure 3, Ex. 1002, ¶133. The '001 Patent describes an identical process for identifying a face from the histograms in Figure 17 and the accompanying description at 22:60-23:14.

d. Hashima discloses the step of "wherein identifying the target…comprises determining a center of the target to be between X and Y minima and maxima of the target" (1[c])

Hashima uses histograms to find the X- and Y-minima and maxima of the target and then calculates the center based on those points using the equations $(X_{MIN} + X_{MAX})/2$ and $(Y_{MIN} + Y_{MAX})/2$. Ex. 1006, 11:13-24, Figure 15; Ex. 1002, ¶135. The equations used to find the center in Hashima are identical to those described for the same purpose in the '001 Patent at 24:50-55.

e. Hashima in view of Ueno renders Claim 1 obvious

To the extent Patent Owner argues Hashima does not disclose the claimed input signal comprised of a plurality of frames, it would have been obvious to replace this feature of Hashima with that of Ueno, as discussed in Section **IX.B.2.b** *supra*; Ex. 1006, 24:24-25; Ex. 1002, ¶128.

Patent Owner may argue that Ueno does not expressly disclose the step of "determining a center of the target to be between X- and Y-minima and maxima of the target" (see claim limitation 1[c]), but Hashima explicitly discloses finding the center of the target as a function of the X- and Y-minima and maxima of the target (Ex. 1006, 11:1-25, Figure 15). Thus, the combination of Hashima and Ueno renders Claim 1 obvious. Ex. 1002, ¶¶116-135, 143.

3. Claim 2: "The process according to claim 1, wherein identifying the target in said at least one histogram further comprises determining a center of the target to be between X and Y minima and maxima of the target"

Claim 2 contains a limitation identical to the limitation in 1[c] already discussed. For the same reasons, Hashima and Ueno both disclose this limitation and render Claim 2 obvious. *See* Ex. 1002, ¶¶136, 143.

4. Claim 3: "The process according to claim 1, wherein identifying the target in said at least one histogram further comprises determining the center of the target at regular intervals"

Hashima discloses processing images on a frame-by-frame basis as they are supplied from the camera. *See, e.g.*, Ex. 1006, 24:13-35 ("When the above general-purpose image processor is employed, it is necessary to write and rea[d] one frame of 25 image data each time it is supplied from the camera, a window is established, the data is converted into binary image data, and projected histograms are generated."). Hashima discloses embodiments where this process requires "at least 200 ms," *id.*, and other embodiments where the "processing time corresponds to only one frame (33 ms)." *Id.* at 25:16-33. Thus, because determining the center of the target is part of the tracking process in both embodiments, it is done at the regular interval of 200 ms or on a frame-by-frame basis at every 33 ms. *Id.* at 25:16-33; Ex. 1002, ¶137. This frame-by-frame determination of the center of the target in Hashima is identical to the description in the '001 Patent at 25:3-25 of

determining the center position at regular intervals. *See* Ex. 1001, 25:3-25; Ex. 1002, ¶137.

Ueno similarly discloses processing the camera input one frame at a time. *See, e.g.*, Ex. 1007, 1:45-46 ("[A] specific region in image signals sequentially input for every frame is detected which corresponds to a movable image…."); 4:25-41 ("Video signals are sequentially input to the image input terminal 100 in frames…. The image signal is written in the frame memory 101 as frame data one frame by one frame…. The data can also be read out every frame."). Because determining the X- and Y-minima and maxima is part of target location and identification process, it necessarily takes place on the regular interval of each frame. Ex. 1002, ¶138.

To the extent Ueno or Hashima does not explicitly disclose that the center of the target is determined at regular intervals, it would have been obvious to implement this limitation in order to achieve smooth, predictable processing. Ex. 1002, ¶139. Thus, Hashima and Ueno both disclose this limitation and render it obvious. Ex. 1002, ¶¶137-139, 143.

5. Claim 4: "The process according to claim 1 further comprising drawing a tracking box around the target"

Ueno discloses the step of drawing a tracking box around the target. *See, e.g.*, Ex. 1007, 7:43-45 ("The facial region is specified by a rectangle 50 designated by coordinates Xs, Xe, Ys, and Ye as shown in FIG. 5."); Ex. 1002,

¶140. This tracking box is shown in Ex. 1007, Figure 5, annotated below (the solid black line around the box is in the original figure):

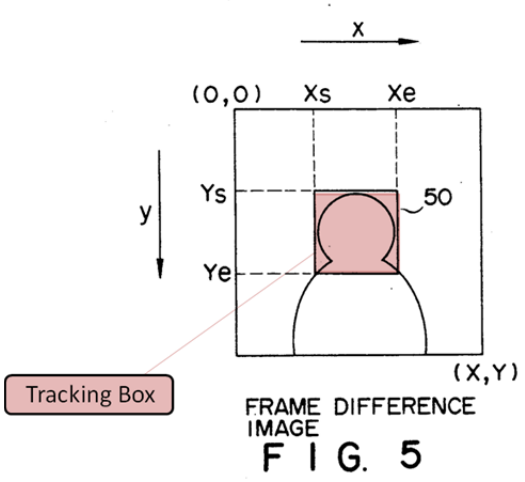


FIG. 5

Ueno also clarifies that the tracking box is displayed on a monitor or display. Ex. 1002, ¶141. For example, Ueno teaches that the region of interest is displayed with a frame that is superimposed on the original image and displayed on the monitor:

> In the self-monitoring mode, low-pass-filtered images are displayed on the monitoring display 710. After a region is specified by the region specifying apparatus 120, the region specifying signal is continuously output. Thus, said frame line showing the specified region is superimposed on the image and displayed on the monitoring display 710.

Ex. 1007, 13:17-23; *id.* at 12:6-25, Figure 18; Ex. 1002, ¶141.

Thus, Ueno discloses this limitation. Ex. 1002, ¶142. Hashima also discloses a tracking box around the target: "each of the images grouped by the labeling process is separated by a window process." Hashima at 8:55-57. This window is "established" by Hashima using the technique described at 13:61-14:34. Furthermore, the system disclosed in Figure 53 described by Hashima at 24:44-25:23 includes a "window determining unit 312" that uses the "window position data" to restrict the video signal to only include the window contents through the "AND gate 313" and out to the "display unit 314" to the "monitor 315." In this configuration, only the contents of the window would be displayed on the monitor and the remainder of the screen would be black, thereby drawing a (black) tracking box around the target currently under consideration. Ex. 1002, ¶¶114, 140. Thus, it would have been obvious to implement this limitation in order to aid a user at a display to visually locate and follow a target. Ex. 1002, ¶¶140-143.

6. Hashima And Ueno Are Not Cumulative

While Hashima and Ueno both disclose image processing systems and methods very similar to those claimed in the '001 Patent, they are not cumulative of one another. Each compliments the other such that strong disclosures of one make up for weaknesses in the other. For example, Hashima discloses imaging a target mark using a camera and reading it into an image memory that stores pixel data. Ex. 1006, 8:27-31. Hashima also discloses reading and writing image data

one frame at a time and creating histograms based on the number of pixels meeting certain criteria the pixels being sorted into classes and counted according to their X- and Y-coordinates (i.e., the domain) of individual pixels. Ex. 1006, 24:24-25. But Hashima does not explicitly discuss an input signal comprising a succession of frames, each frame comprising a succession of pixels. Ex. 1002, ¶124. Of course a video camera input that is processed one frame at a time to create a histogram based on pixels meeting certain characteristics at a given location necessarily includes an input comprising a succession of frames, each frame comprising a succession of pixels. Thus, an input signal comprising a succession of frames is inherently disclosed by Hashima. Ex. 1002, ¶124. On the other hand, this limitation is explicitly disclosed by Ueno, which teaches that "a specific region in the image signals sequentially input for every frame" and that a block size in the conversion unit is 8 x 8 pixels large. Ex. 1007, 4:25-53; Ex. 1002, ¶125. Thus, the combination of Hashima and Ueno strengthens the overall disclosure. Ex. 1002, ¶¶122-128.

Similarly, although Ueno specifically discloses using the X-axis histogram to locate the left and right end of the target (i.e., X-minima and maxima), and using the Y-axis histogram to locate the top and bottom of the target (i.e., Y-minima and maxima), it does not explicitly disclose finding the center point of the target from these maxima and minima. Of course calculating a center point is simply a

mathematical identity based on determining the X- and Y-minima and maxima, so this limitation is inherently disclosed by Ueno or at a minimum would have been obvious in view of the knowledge of a POSA. Ex. 1002, ¶136. However, to the extent Patent Owner argues that the claims of the '001 Patent require calculating the center point using the calculation described in the '001 specification, Hashima explicitly discloses calculating the center point as the point between the X-minima and maxima and between the Y-minima and maxima as described in the '001 Patent. Ex. 1006, 11:13-24, Figure 15; Ex. 1002, ¶135.

Thus, while each of Hashima and Ueno arguably renders the challenged claims of the '001 Patent obvious on its own in view of the knowledge of a POSA, the combination of Hashima and Ueno provides a more complete disclosure, and Petitioner urges the Board to adopt this ground as presented because Hashima and Ueno are not cumulative. *See* Ex. 1002, ¶143.

C. Ground 3: Ueno In View Of Gilbert Renders Obvious Claims 1-4

1. Reasons To Combine Ueno And Gilbert

A POSA would have been motivated to combine Ueno and Gilbert because they are directed toward very similar systems that operate in a similar manner. As set forth above, both Ueno and Gilbert are directed to identifying and tracking a target using an image processing system with a video camera input, (Ex. 1007, 3:1-8, 4:25-41; Ex. 1005, 47-48, Figure 1), where the video input comprises a

succession of frames, each frame comprising a succession of pixels (Ex. 1007, 1:44-49, 4:25-41; Ex. 1005, 47-48, 52), in which the target comprises pixels in one or more of a plurality of classes and one or more of a plurality of domains (Ex. 1007, 7:8-16; Ex. 1005, 48; Ex. 1002). *See also*, Ex. 1002, ¶145. The Ueno and Gilbert systems also perform this process on a frame-by-frame basis. Ex. 1007, 4:25-41, 5:31-40; Ex. 1005, 47-48, 52; Ex. 1002, ¶145.

Additionally, both Ueno and Gilbert identify and track a moving target using histograms. Ex. 1007, 7:7-8:22; Ex. 1005, 48-49. Thus, they are not only in the same field of endeavor but also are directed to very similar systems and processes. Ex. 1002, ¶146. A POSA would have recognized that Gilbert's invention could improve a similar device, such as the system in Ueno, in the same way, and thus would have been motivated to combine Ueno and Gilbert. Ex. 1002, ¶146; *see KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398 (2007) ("if a technique has been used to improve one device, and a person of ordinary skill in the art would recognize that it would improve similar devices in the same way, using the technique is obvious unless its actual application is beyond his or her skill."). A POSA would have also expected the combination to yield a predictable result because it would have involved applying known techniques to similar systems. Ex. 1002, ¶147.

Moreover, the references themselves, taken as a whole, demonstrate reasons why a POSA would be motivated to combine Ueno with Gilbert. *See In re Beattie*,

974 F.2d 1309, 1312 (Fed. Cir. 1992) ("As long as some motivation or suggestion to combine the references is provided by the prior art taken as a whole, the law does not require the references be combined for the reasons contemplated by the invention."). In particular, although Ueno describes a system that identifies and tracks a target and draws a tracking box around the target, to the extent the Patent Owner argues Ueno does not disclose a means of finding a center of the target, a POSA would have been motivated to improve Ueno's system to add a step of calculating the target's center, because this center calculation enables the addition of other functionality. Ex. 1002, ¶148. For example a POSA would then have been able to add functionality to allow the camera to adjust its field of view to follow the target, a capability not present in Ueno's camera, which is stationary. *Id.* Adding the center calculation would have allowed the Ueno system more flexibility to track targets physically moving around a room during a video call and would have allowed the Ueno system to be useful in other related applications. *Id.* Thus, a POSA would have been motivated to look for prior art that teaches a tracking system where the system calculates the target center and where the camera automatically adjusts to keep the target centered in the field of view. *Id.* As described above, Gilbert is an example of such a system and teaches that "[t]he boresight correction signals are used to control the azimuth and elevation pointing angles of the telescope…keeping the target visible within the FOV." Ex. 1005, 52.

Moreover, there would have been a reasonable expectation of success that in doing so, a more accurate target identification system would have resulted.

Although Gilbert recognizes that the center of the target can be found using "the target nose and tail points," i.e., the X- and Y-maxima and minima, Gilbert considers this method to be prone to noise perturbations and thus not suitable for missile tracking systems. Gilbert thus employs a center calculation using the center of area of two halves of the target. A POSA would have recognized that the limitations recognized by Gilbert would not have been applicable to a system like that described in Ueno. Ex. 1002, ¶149. For example, Gilbert teaches that the difficulty in finding the exact nose and tail points of a rocket and a significant risk that noise in the tracking system would make reliance on the nose and tail points unreliable. Ex. 1005, 50. However, in Ueno, the targets are human faces moving at slow speeds in contexts where noise would be at a minimum and not prohibit finding the extreme values. Ex. 1005, 50. Thus, in order to benefit from faster, more efficient processing, a POSA would have been motivated to apply the simpler center calculation using the nose and tail points suggested in Gilbert, rather than the more complex center of area calculation explicitly described. Ex. 1002, ¶149. This would have been especially obvious to a POSA given that Ueno already calculates the X- and Y-minima and maxima, making the application of using the

extreme values to find the center a simple calculation involving only one additional step. Ex. 1002, ¶149.

Thus, a POSA would have been motivated to combine Ueno's image processing system with that described by Gilbert to enhance the video camera to add center calculation to automatically track a target by physically adjusting the optical axis of the camera as disclosed in Gilbert, and would have expected such a combination to have a reasonable expectation of success. Ex. 1002, ¶¶147-150.

A detailed comparison will now be made of Ueno in view of Gilbert to the asserted claims of the '001 Patent.

2. Claim 1

a. Ueno and Gilbert each disclose: "a process of tracking a target…on a frame-by-frame basis" (1[pre])

Gilbert describes 1[pre] as explained in Section **IX.A.2.a** above. Ex. 1005, 47-48. Ueno also describe 1[pre] as explained in Section **IX.B.2.a** above. Ex. 1007, 3:1-8, 4:25-30, 4:39-41, 7:8-16. Ueno also discloses a plurality of domains by disclosing histograms in the X- and Y- domains. Ex. 1007, 7:8-16; Ex. 1002, ¶¶141-159.

b. Ueno and Gilbert each disclose the step of "forming at least one histogram…referring to classes defining said target" (1[a])

In both Ueno and Gilbert, at least one histogram is formed based on pixel characteristics (domains) such as intensity or coordinate axis. Ex. 1002, ¶160. For

example, Ueno forms histograms in X- and Y-domains. Ex. 1007, 7:7-16, Figure 3; Ex. 1002, ¶160. These histograms in Ueno are analogous to the X- and Y-position histograms discussed in the '001 Patent at 20:60-21:29. Ex. 1002, ¶160. Similarly, Gilbert forms a histogram of the intensity domain, and teaches other "features that can be functionally derived from relationships between pixels, e.g., texture, edge, and linearity measures" may also be used. Ex. 1005, 48; Ex. 1002, ¶161.

Ueno also explicitly discloses a plurality of domains by showing histograms in the X- and Y-domains. *See* Ex. 1007, 7:11-15. Moreover, to extent the Patent Owner argues that Gilbert's intensity histogram does not include a plurality of "classes," because the histogram contains every pixel (sorted into bins based on intensity level) rather than a subset of pixels that meet a threshold intensity characteristic, this argument should be rejected. Gilbert's intensity histogram is entirely analogous to the velocity histogram in Figure 14a of the '001 Patent, as described at 20:54-59. Ex. 1002, ¶161.

Moreover, Gilbert also discloses a projection processor that forms X- and Y-projection histograms using binary pictures generated by the video processor. Ex. 1005, Figure 4. These projection histograms are analogous to the X- and Y-position histograms discussed in the '001 Patent at 20:60-21:29. Thus, Gilbert

discloses creating histograms in a plurality of domains, including the intensity, the X-, and the Y-domains. Ex. 1002, ¶162.

In these histograms, each "class" within a domain counts the number of pixels meeting certain threshold criteria. For example, Gilbert's intensity histogram has 256 grayscale levels (i.e., 256 classes). Ex. 1005, 48. The histogram is formed as each pixel in the frame is sorted into one of the 256 classes for which that pixel's intensity level fits. Ex. 1005, 48; Ex. 1002, ¶163. In Ueno, the number of pixels having the value "1" (which indicates that a pixel meets a certain intensity threshold) within a given column defined by an X- coordinate is counted into the class that corresponds to that X-coordinate value. Ex. 1007, 7:7-45; Ex. 1002, ¶163. Similarly, in the X- and Y-projection histograms of Gilbert Figure 4, the number of pixels having the value "1" in the binary picture (which indicates that a pixel belongs to the target) within a given column defined by an X- or Y-coordinate is counted into the class that corresponds to that X- or Y-coordinate value. Ex. 1005, 50-51; Ex. 1002, ¶¶163-164. These binary selection processes in Ueno and Gilbert are entirely analogous to the histogram formation process discussed in the '001 Patent at 18:16-19:24. Ex. 1002, ¶163.

c. Ueno and Gilbert each disclose the step of "identifying the target in said at least one histogram itself" (1[b])

In both Ueno and Gilbert, histograms are used to identify the target. Ex. 1002, ¶165. For example, in Ueno, histograms are generated on interframe images

and specifically to identify pixels that are different in the two frames, creating histograms that are used to determine the location of the target. Ex. 1007, 7:7-16, Figure 3; Ex. 1002, ¶165. Similarly, in Gilbert, an intensity histogram is used to separate the target from the background and the plume. Ex. 1005, 48-50; Ex. 1002, ¶166.

d. Gilbert discloses the step of "wherein identifying the target…comprises determining a center of the target to be between X and Y minima and maxima of the target" (1[c])

Gilbert discloses this limitation as explained in Section **IX.A.2.d.** In particular, Gilbert suggests that the center point may either be calculated by simply using the extreme values from the histogram or by a more complicated center of area calculation that involves first bisecting the image. Ex. 1005, 50; Ex. 1002, ¶¶167-168. Gilbert also discloses the technique of using the "target nose and tail points," i.e., the X- and Y-minima and maxima, to calculate the target center. Ex. 1005, 50; Ex. 1002, ¶167. In addition, the standard centering calculation disclosed in the '001 Patent would have been well known to a POSA. A POSA thus would have desired to implement this standard centering calculation as the most cost effective and simple way to calculate a center point. Ex. 1002, ¶¶168-169.

Thus, the combination of Ueno and Gilbert renders Claim 1 obvious. Ex. 1002, ¶¶144-169, 178.

3. Claim 2: "The process according to claim 1, wherein identifying the target in said at least one histogram further comprises determining a center of the target to be between X and Y minima and maxima of the target"

Claim 2 contains a limitation identical to the limitation in 1[c] already discussed. For the same reasons, Ueno and Gilbert both disclose this limitation and render Claim 2 obvious. *See* Ex. 1002, ¶¶170, 178.

4. Claim 3: "The process according to claim 1, wherein identifying the target in said at least one histogram further comprises determining the center of the target at regular intervals"

Ueno discloses the step of processing the camera input one frame at a time. *See, e.g.,* Ex. 1007, 1:45-46 ("[A] specific region in image signals sequentially input for every frame is detected which corresponds to a movable image…."); 4:25-41 ("Video signals are sequentially input to the image input terminal 100 in frames…. The image signal is written in the frame memory 101 as frame data one frame by one frame…. The data can also be read out every frame."). *See* Ex. 1002, ¶171. Because determining the X- and Y-minima and maxima is part of target location and identification process, it necessarily takes place on the regular interval of each frame. *Id.*

Gilbert similarly discloses forming projection histograms on a frame-by-frame basis at a rate of 30 frames/second. *See, e.g.*, Ex. 1005, 50 ("In the projection processor, these matrices are analyzed field-by-field at 60 fields/s [i.e.,

30 frames/s] using the projection based algorithm"); Ex. 1005, 55 ("The result after six processing frames…."). Because part of this analysis involves generating a histogram to identify the target, and further finding the center of the target using the generated histogram, this process is necessarily performed at regular intervals, namely once per frame or once every 1/30 second. Ex. 1002, ¶172. This frame-by-frame determination of the center of the target in Gilbert is identical to the description in the '001 Patent at 25:3-25 of determining the center position at regular intervals. *See* Ex. 1002, ¶172.

To the extent Ueno or Gilbert do not explicitly disclose that the center of the target is determined at regular intervals, it would have been obvious to implement this limitation in order to achieve smooth, predictable processing. Ex. 1002, ¶173. Thus, Ueno and Gilbert both disclose this limitation and render it obvious. Ex. 1002, ¶¶171-173, 178.

5. Claim 4: "The process according to claim 1 further comprising drawing a tracking box around the target"

Ueno discloses the step of drawing a tracking box around the target. *See, e.g.*, Ex. 1007, 7:43-45 ("The facial region is specified by a rectangle 50 designated by coordinates Xs, Xe, Ys, and Ye as shown in FIG. 5."); Ex. 1002, ¶174. This tracking box is shown in Ex. 1007, Figure 5, annotated below (the solid black line around the box is in the original figure):

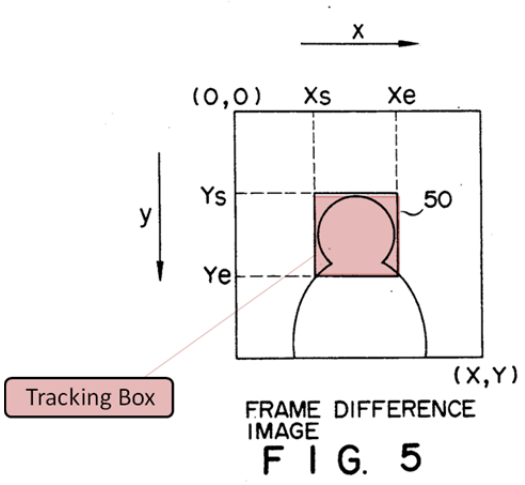


FIG. 5

Ueno also states that the tracking box is displayed on a monitor or display. For example, Ueno teaches that the region of interest is displayed with a frame that is superimposed on the original image and displayed on the monitor:

> In the self-monitoring mode, low-pass-filtered images are displayed on the monitoring display 710. After a region is specified by the region specifying apparatus 120, the region specifying signal is continuously output. Thus, said frame line showing the specified region is superimposed on the image and displayed on the monitoring display 710.

Ex. 1007, 13:17-23, 12:6-25, Figure 18; Ex. 1002, ¶175.

Gilbert similarly discloses drawing a tracking box around the target. *See, e.g.*, Ex. 1005, 48 ("A tracking window is placed about the target image,…The tracking window frame is partitioned into a background region (BR) and a plume

region (PR). The region inside the frame is called the target region (TR) as shown in Fig. 2."); Ex. 1002, ¶176. This tracking window is shown in Ex. 1005, Figure 2 as annotated below:

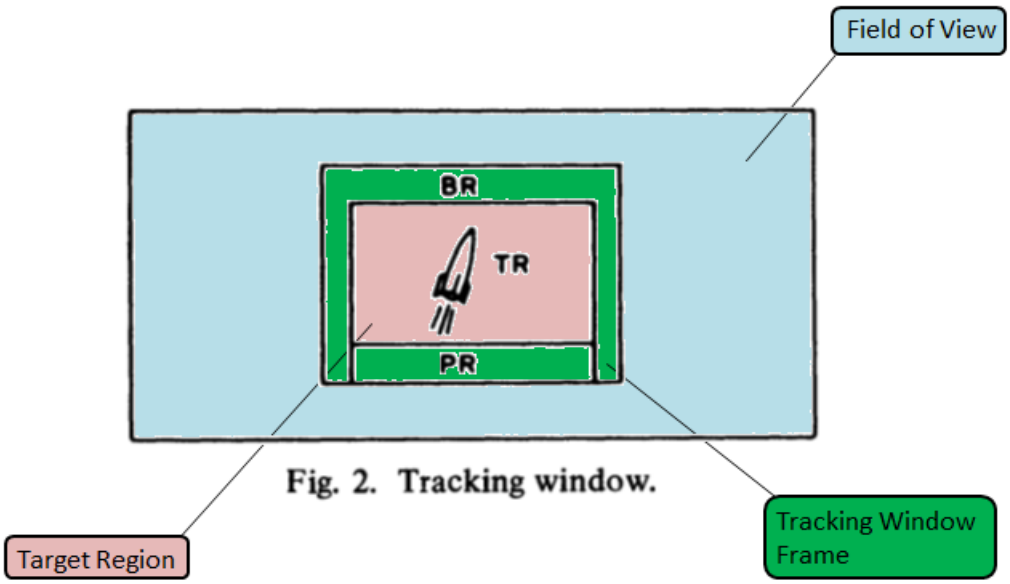


Fig. 2. Tracking window.

In simulation outputs (using pre-recorded video of a rocket in flight as an input signal), the system similarly tracks the target by drawing a tracking box. *See, e.g.*, Ex. 1005, 55 ("Fig. 7 contains two representative simulation outputs selected from the first 100 fields of simulated digitized video."). Ex. 1002, ¶176. These simulation outputs are shown in Ex. 1005, Figure 7, where a rocket is being tracked, as annotated below:

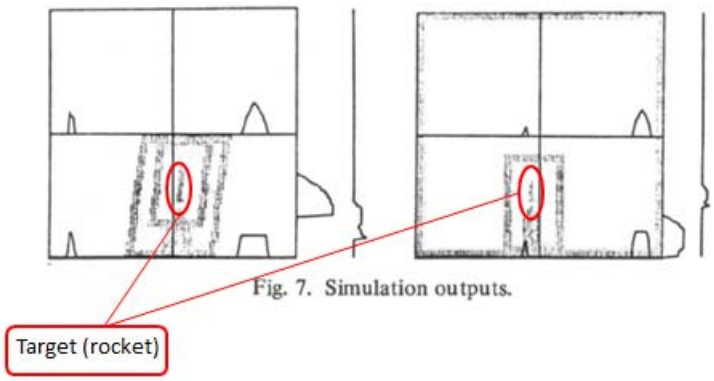


Fig. 7. Simulation outputs.

Ex. 1002, ¶176.

Thus, Ueno and Gilbert both disclose this limitation. Ex. 1002, ¶177. To the extent Patent Owner argues that Gilbert does not explicitly disclose drawing a tracking box around the target, it would have been obvious to implement this limitation in order to aid a user at a display to visually locate and follow a target. Ex. 1002, ¶¶174-178.

6. Ueno and Gilbert Are Not Cumulative

While Ueno and Gilbert both disclose image processing systems and methods very similar to those claimed in the '001 Patent, they are not cumulative of one another. Each compliments the other such that strong disclosures of one make up for weaknesses in the other. For example, although Ueno teaches finding the X- and Y-minima and maxima and drawing a tracking box around the target, it does not explicitly teach the step of determining the center of the target. Gilbert, on the other hand, expressly teaches one method of finding the center of a target and recognizes that a similar calculation may be done using the X- and Y-minima

and maxima of the target. Ex. 1002, ¶167; Ex. 1005, 50. To the extent Patent Owner argues that finding the X- and Y-minima and maxima is not disclosed in either Gilbert or Ueno, it would have been obvious to a POSA, especially in view of the fact that Ueno calculates the X- and Y-minima and maxima, and Gilbert teaches using these values as one way of finding the center of the target. Ex. 1002, ¶¶168-169.

Similarly, Gilbert generally teaches the step of forming a histogram in one or more of a plurality of domains, such as the intensity (luminance) domain used to construct the intensity histogram, or the X- and Y-domains used to construct the projection histograms. Ex. 1005, 48-51, Figure 4. Ueno provides more explicit teaching of identifying the target using both X- and Y-histograms. Ex. 1007, 7:7-16, Figure 3; Ex. 1002, ¶¶160-163.

To the extent the Patent Owner argues that the only histogram disclosed in Gilbert are the histograms formed in the intensity domain (i.e., that the X- and Y-projections in Gilbert are not "histograms"), Ueno makes up for this purported weakness because it discloses creating histograms in the X-axis and Y-axis domains (i.e., a plurality of domains). Ex. 1002, ¶162. Moreover, Gilbert teaches that "[t]here are many feature that can be functionally derived from relationships between pixels, e.g., texture, edge, and linearity measures." Ex. 1005, 48. A POSA would have understood that Gilbert thus teaches other histograms formed

based on domains of texture, edge, or linearity measures may be used to achieve similar results. Ex. 1002, ¶146.

Thus, while each of Ueno and Gilbert renders the challenged claims of the '001 Patent obvious in combination with the knowledge of a POSA, the combination of Ueno and Gilbert provides a more complete disclosure, and Petitioner urges the Board to adopt this ground as presented because Ueno and Gilbert are not cumulative. Ex. 1002, ¶¶144, 178.

X. CONCLUSION

For the reasons given above, Petitioner requests institution of IPR for Claims 1-4 of the '001 Patent based on each of the grounds specified in this petition.

Dated: November 29, 2016

By: */s/ John Kappos*
John Kappos (Reg. No. 37,861)

CERTIFICATION OF WORD COUNT

Pursuant to 37 C.F.R. § 42.24(d), Petitioner certifies that this petition includes 13,754 words, as measured by Microsoft Word, exclusive of the table of contents, mandatory notices under § 42.8, certificates of service and word count, and exhibits.

CERTIFICATE OF SERVICE

The undersigned certifies pursuant to 37 C.F.R. § 42.6(e) and § 42.105 that on November 29, 2016, a true and correct copy of Petitioner **Petition for *Inter Partes* Review of U.S. Patent No. 8,805,001** was served via express mail on the Petitioner at the following correspondence address of record:

Gregory Nelson
FOX ROTHSCHILD LLP
997 Lennox Drive
Lawrenceville, NJ 08648

Courtesy copies were also sent to Patent Owner's litigation counsel at the following addresses:

S. Calvin Capshaw
CAPSHAW DERIEUX, LLP
114 E. Commerce Ave.
Gladewater, TX 75647

Christopher J. Coulson
KENYON & KENYON LLP
One Broadway
New York, NY 10004

Michael N. Zachary
KENYON & KENYON LLP
110 Peachtree Street, Suite 2800
Atlanta, GA 30309

Respectfully submitted,

/s/ John Kappos
John Kappos (Reg. No. 37,861)
O'Melveny & Myers LLP
610 Newport Center Drive, 17th Floor
Newport Beach, CA 92660
Telephone: (949) 760-9600

Attorney for Petitioner