Image Sequence Processing Using Spatiotemporal Segmentation

Gene K. Wu and Todd R. Reed, *Senior Member, IEEE*

***Abstract*— We investigate the improvements that can be obtained in several conventional video-processing algorithms through the incorporation of three-dimensional (3-D) (spatiotemporal) segmentation information. Four classes of image sequence processing techniques are considered: low-pass filtering, high-pass filtering, high-frequency emphasis, and 3-D Sobel filtering. It is demonstrated that segmentation information can improve the performance of these techniques substantially so that this approach may be promising for other applications (e.g., deinterlacing and resolution conversion) as well.**

***Index Terms*— Image edge analysis, image enhancement, image processing, image segmentation, image sequence analysis, Markov processes.**

I. Introduction

SEGMENTATION, one of the most commonly used techniques in image analysis, refers to the identification of uniform regions based on certain conditions. The conditions under which these regions are formed can be uniformity of gray value, color, texture, and so on. Segmentation has long been used in computer vision for a variety of applications, particularly object recognition. Recently, segmentation techniques have also been used in image compression and image processing. In image compression, segmentation-based compression methods are sometimes referred to as second-generation methods [1]. Techniques of this type are becoming increasingly important, as exemplified by the role of segmentation in the evolving MPEG-4 compression standard [2]. In image processing, as reported in [3] and [4], if two-dimensional (2-D) (spatial) segmentation information is used as an aid, the processed image's quality is much better than that processed without such an aid.

In this work, three-dimensional (3-D) (spatiotemporal) segmentation is applied to image sequences. The result of the segmentation process is a description of the image sequence consisting of 3-D volumetric regions. A concise approximation of the sequence can be formed based on this description by encoding the region boundary locations together with an expression for the intensities of the pixels within each region. The simplest such expression is the mean value of the pixel intensities for each region. Polynomials and various transforms can also be used to represent the interiors of regions. While more accurate, these expressions are also more complex.

How regions are related from one frame to another is very important for video processing. Such information is readily available in the results of 3-D segmentation. For example, if a uniform region spans several successive frames, the region boundary surfaces along the temporal dimension show how the locations of high contrast edges evolve with time while the interiors of regions indicate the correlation of pixels between frames. These facts can be exploited for improved filtering, compression, and representation of the image sequence [5]. The price for this improvement is the computational power and time required to obtain the segmentation. However, in some cases, the segmentation may already be available (e.g., if the sequence is encoded using segmentation-based compression).

In this paper, we demonstrate the use of 3-D segmentation to enhance a number of conventional image sequence processing techniques. In particular, we examine low-pass and high-pass filtering, high-frequency emphasis, and edge enhancement.

There are many segmentation approaches found in the literature [6], [7]. The image sequence processing techniques we will discuss in this paper do not rely on the details of the 3-D segmentation method. We require only that segmentation boundaries correspond closely to their associated object boundaries. Note that this does not require that all segmentation boundaries correspond to object boundaries (that there be no so-called "false contours"). While not especially desirable, we will see that for the most part, such contours do not have a strong impact on the algorithms that follow. The segmentation method used in this work is a region-growing method using a Gibbs–Markov random field (GMRF) model and contour relaxation. The segmentation model and algorithm are described in the following section.

II. 3-D Segmentation Using a Gibbs–Markov Model

The 3-D segmentation method we use employs a Gaussian model to represent the region interiors and a Gibbs–Markov random field model for the boundaries of regions in the sequence.

GMRF models are widely used in image processing. They are suitable for segmentation because they can capture the local properties of regions [8]. Three-dimensional segmentation using a GMRF model combined with contour relaxation has been found to produce accurate object boundaries when applied to volumetric data [9]. The model can also be used to produce accurate boundaries of uniform regions and smooth transitions from frame to frame in image sequences [5].

Manuscript received April 10, 1997; revised March 10, 1999. This paper was recommended by Associate Editor M.-T. Sun.

G. K. Wu was with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616 USA. He is now with Ricoh Silicon Valley, Menlo Park, CA 94025-7022 USA.

T. R. Reed is with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616 USA (e-mail: trreed@ucdavis.edu).

Publisher Item Identifier S 1051-8215(99)06189-3.

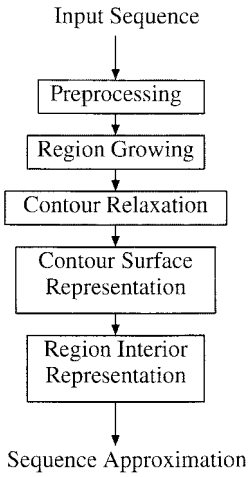


Fig. 1. The 3-D segmentation algorithm.

The steps in the segmentation algorithm are preprocessing, region growing, and contour relaxation, as shown in Fig. 1.

The preprocessing step models Weber's law, taking the response of the visual system with varying intensity into account. The following transformation was used:

$$C = \kappa * f^{1/3}. \tag{1}$$

C is the perceived brightness, κ is a proportionality constant, and f is the stimulus intensity. The value of 48.8 was used for κ in this work. Applying this transformation to the input sequence allows us to find a more natural partitioning of images, closer to the segmentation perceived by the human visual system. The value 1/3 was used for coding purposes, as suggested in [11].

After this step, region growing is performed. This portion of the algorithm consists of three different phases, using a more sophisticated model in each phase. First, only the mean of the luminance of the regions are used as a merging criterion. Second, the variance of each region is included in the model. A merging cost based on the mean and variance of each region is computed for each pair of neighboring regions. All neighboring regions having a merging cost smaller than a specified threshold are merged in the order of increasing cost.

A third model is introduced to obtain a more accurate partitioning of the sequence without increasing the number of regions. In this model, region contour and region interior information are used jointly to perform contour relaxation. A detailed description of this final phase of the algorithm is presented in the following section.

A. Contour Relaxation

Contour relaxation can be done in many ways. The method used here, first introduced for 3-D data in [9], consists of changing a boundary pixel label (which indicates region membership) if doing so results in a local maximum of a joint likelihood function. The main intent of the contour relaxation is to smooth rough boundaries and to obtain a segmentation that corresponds closely to the objects in the original image sequence.

Given the initial segmentation, boundaries are refined by examining each boundary pixel. The region label assigned to a boundary pixel is changed to that of one of its neighbors if the modification will locally maximize the joint likelihood of the segmentation and the image sequence data. The neighborhood is 26-connected (analogous to the 8-connected case for images). Because of the nature of the model used in the formulation of the likelihood function (discussed in detail in the following section), after the application of the contour relaxation algorithm, regions have more regularized boundaries that correspond to a more accurate partitioning of the underlying image data.

The contour relaxation algorithm is applied iteratively to all boundary pixels of the partitioned sequence. The algorithm is stopped when the number of boundary labels altered is less than a threshold.

B. Gibbs–Markov Model

An image segmentation Q can be modeled as a sample from a 3-D Gibbs–Markov random field. The probability of a segmentation Q is given by

$$P(Q) = \frac{1}{Z} e^{-U} \tag{2}$$

where

$$U = \sum_{c_i \subset Q} V(c_i) \tag{3}$$

is a potential function and $Z = \sum e^{-U}$ is a normalizing constant. For a GMRF of order two, the c_i (sometimes referred to as cliques) consist of each possible pair of adjacent pixels and $V(c_i)$ is defined as

$$V(c_i) = \begin{cases} \alpha & \text{all inhomo. nearest neighbor pairs parallel to the } x \text{ or } y \text{ axis} \\ \beta & \text{all inhomo. diagonal neighbor pairs in the same } x\text{–}y \text{ plane} \\ \gamma & \text{all inhomo. neighbor pairs parallel to the } t \text{ axis} \\ \delta & \text{all diagonal inhomo. neighbor pairs each pixel in different } x\text{–}y \text{ planes} \\ 0 & \text{otherwise.} \end{cases}$$

α is the penalty value for having inhomogeneous neighbor pairs parallel to the x- or y-axis. β is the penalty value for having inhomogeneous diagonal neighbor pairs on the same frame. γ is the penalty value for having inhomogeneous neighbor pairs parallel to the t axis. δ is the penalty value for having inhomogeneous diagonal neighbor pairs on the same frame.

Inhomogeneous pixel pairs are those with different labels (belonging to different regions). The coordinate x refers to the positions of pixels on the horizontal axis in an image, y refers to the positions of pixels on the vertical axis, and t refers to the positions of frames in the temporal direction. The joint

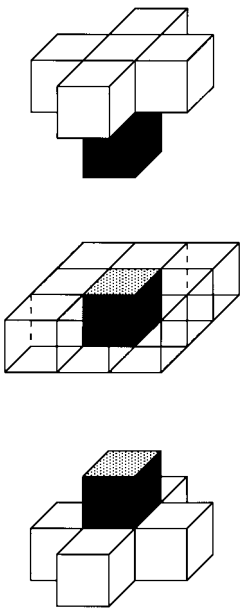


Fig. 2. The 3-D cliques.

likelihood that must be maximized is

$$P(Y, Q) = P(Q)P(Y|Q) \tag{4}$$

where Y is the original sequence data. Using a Gaussian approximation in the formulation of $P(Y|Q)$

$$P(Y|Q) = \prod_{R_j} \left(\sqrt{2\pi\tilde{\sigma}^2(R_j)} \right)^{-N_j} e^{-N_j/2} \tag{5}$$

the joint likelihood can be expressed as

$$P(Y, Q) = \frac{1}{Z} e^{-(n_a\alpha + n_b\beta + n_c\gamma + n_d\delta)} \prod_{R_j} \left(\sqrt{2\pi\tilde{\sigma}^2(R_j)} \right)^{-N_j} \cdot e^{-N_j/2}. \tag{6}$$

R_j is the jth region of the partition Q, N_j is the size of the jth region, and $\tilde{\sigma}^2$ is the estimated variance of the jth region [9]. The product comprises all regions of the partition Q. n_a is the number of inhomogeneous nearest neighbor cliques in R_j. They are the vertical and horizontal cliques on the middle frame in Fig. 2. n_b is the number of inhomogeneous diagonal neighbor cliques on the same frame in R_j, the diagonal cliques on the middle frame in Fig. 2. n_c is the number of inhomogeneous nearest neighbor cliques parallel to the t axis in R_j, the vertical cliques on the top and bottom frames in Fig. 2. n_d is the number of inhomogeneous diagonal neighbor cliques with each pixel lying on different frames in R_j, the diagonal cliques on the top and bottom frames in Fig. 2.

The values for α, β, γ, and δ can be determined based on the local surrounding pixels. In this paper, the values used in the model for α, β, γ, and δ are 1, 0.5, 1, and 0.5, respectively. If these values are increased, the joint likelihood is reduced given all cliques in the region remain the same. The choice of these parameters is not particularly critical. Fig. 2 shows all the possible 3-D cliques considered in the GMRF model.

Fig. 3. Selected frames from the *Miss America* sequence.

Fig. 4. Selected frames from the *Secretary* sequence.

C. Examples

Some selected frames from the *Miss America* and *Secretary* sequences are shown in Figs. 3 and 4. Each sequence is composed of 24 frames of image size 128 × 128 pixels. These sequences will be used as test data throughout the paper. Due to space constraints, only individual frames from a processed sequence will be presented.

The values used for obtaining the results are: the minimum number of pixels allowed per region is 30, the intensity difference between neighboring pixels is one, and the merging cost threshold is 20 in increments up to 500. The number of

(a) (b) (c)

Fig. 5. (a) Frame 20 of the original *Miss America* sequence, (b) contour map, and (c) approximation image.

(a) (b) (c)

Fig. 6. (a) Frame 3 of the original *Secretary* sequence, (b) contour map, and (c) approximation image.

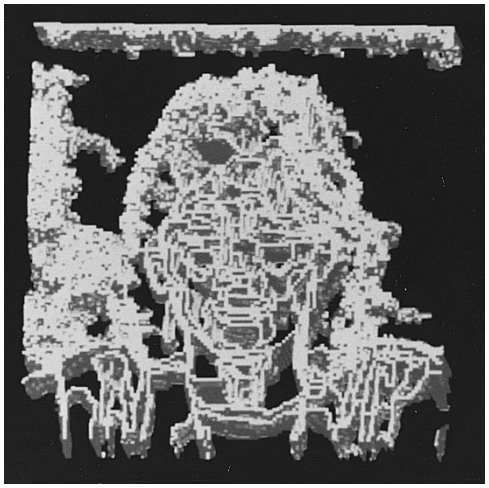


Fig. 7. Contour frames of the *Miss America* approximation sequence viewed in a 3-D perspective.

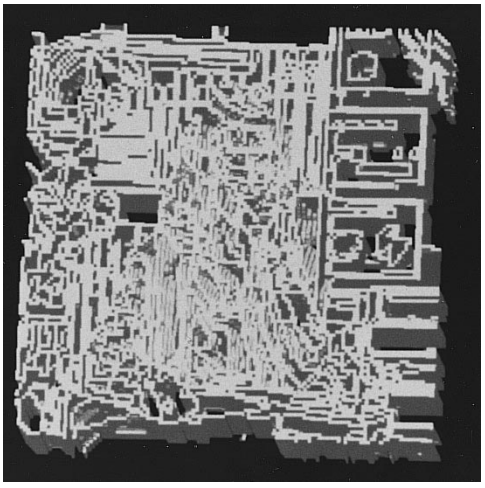


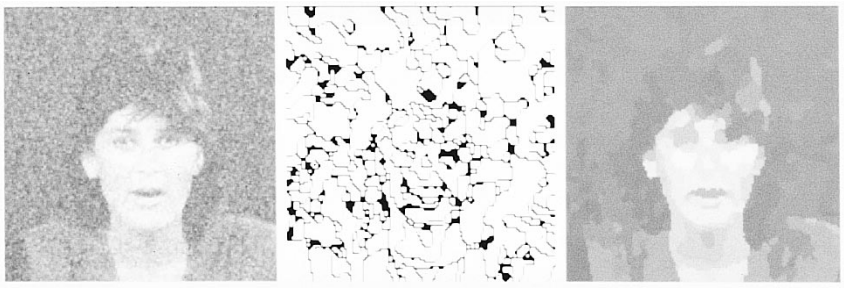
Fig. 8. Contour frames of the *Secretary* approximation sequence viewed in a 3-D perspective.

regions present are 85 and 210 in the *Miss America* sequence and the *Secretary* sequence, respectively. It takes about 15 h on an HP 715/75 workstation to process a sequence of eight frames with a frame size of 256 $\times$ 256 pixels.

The results of 3-D segmentation on a typical head-and-shoulders sequence *(Miss America)* are shown in Fig. 5. The image in Fig. 5(a) is a selected frame from the original sequence, the image in (b) is the contour map for this frame, and the image in (c) is the associated frame from the image sequence approximation formed by filling each region with its mean value. For the *Secretary* sequence, sample frames from each sequence are shown in Fig. 6. The images in Fig. 6(c) may look blocky due to the low resolution of the image frame size (128 $\times$ 128 pixels) being displayed on a 2.15 $\times$ 2.15 in^2 square area. The segmentation method does not introduce the blocking effect.

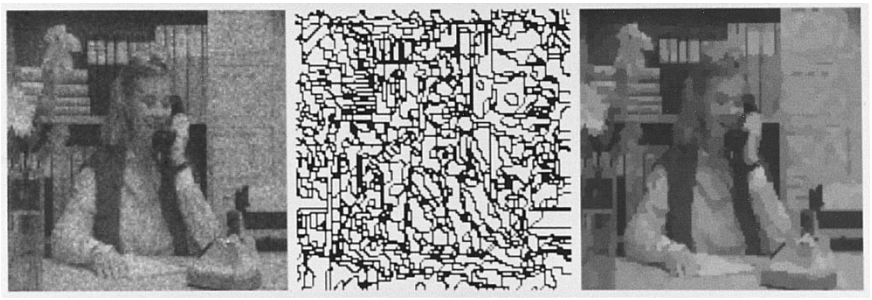
Figs. 7 and 8 show the contour frames derived from the *Miss America* sequence and the *Secretary* sequence, respectively, assembled into 3-D (spatiotemporal) volumes. These 3-D contour volumes are viewed in perspective. The holes in the volume indicate locations where consecutive frames share common regions.

To prove the viability of this method under nonideal circumstances, and to provide two test sequences for some of



(a) (b) (c)

Fig. 9. (a) Frame 20 from the noisy *Miss America* sequence, (b) contour map, and (c) approximation image.



(a) (b) (c)

Fig. 10. (a) Frame 3 from the noisy *Secretary* sequence, (b) contour map, and (c) approximation image.

the methods that follow, two noisy image sequences were produced by normalizing and adding Gaussian noise with a variance of 150 to the original sequences. The results of 3-D segmentation of the noisy *Miss America* sequence are shown in Fig. 9. The image in Fig. 9(a) is a selected frame from the noisy sequence, the image in (b) is the contour map for this frame, and the image in (c) is the approximation image.

The results of 3-D segmentation of the noisy *Secretary* sequence are shown in Fig. 10. The image in Fig. 10(a) is a selected frame from the noisy sequence, the image in (b) is the contour map for this frame, and the image in (c) is the approximation image.

Comparing the results in Fig. 9 with those in Fig. 5, it is clear that the segmentation results obtained from the noisy sequence have more regions than those from the original image sequence. These extra regions can be formed in two ways. Either regions found in the noise-free sequence are broken up into smaller regions or new regions are added in locations where they did not previously exist. It is important to note, however, that the approximation resulting from the segmentation of the noisy sequence still captures the majority of the important structure of the sequence.

III. Segmentation-Based Video-Processing Methods

In this section, we consider the improvements that can be obtained by including segmentation information in algorithms to perform some common video-processing tasks. Four applications will be investigated: segmentation-aided low-pass filtering, segmentation-aided high-pass filtering, high-frequency emphasis, and edge enhancement.

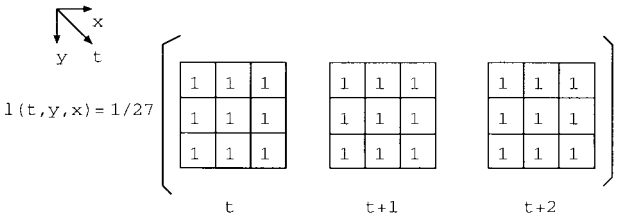


Fig. 11. The 3-D low-pass filtering operator.

A. Low-Pass Filtering

In conventional low-pass filtering for noise reduction, the price paid for the reduction in noise is the blurring of the image (especially noticeable around high contrast edges). A tradeoff must be made between noise reduction and sharp feature reduction.

This being the case, it would be very desirable if we could perform the low-pass filtering only in the interior of regions but not in areas where there are sharp features (such as at the boundaries of objects). Three-dimensional segmentation makes such selective filtering possible. A simple example of this approach is selective averaging. Whenever the kernel is completely within a uniform region, averaging is performed. However, at region boundaries, only those pixels that belong to the same region as the pixel being evaluated are considered in the average. The low-pass operator $l(t, y, x)$ used for low-pass filtering is shown in Fig. 11.

The results of both the conventional and segmentation-aided methods are shown in Fig. 12 (frame 3 from each processed sequence). The image in Fig. 12(a) is from the noisy sequence.

(a) (b) (c)

Fig. 12. (a) Frame 3 from the noisy *Miss America* sequence, (b) conventionally filtered sequence, and (c) guided low-pass filtered sequence.

(a) (b) (c)

Fig. 13. (a) Frame 3 from the noisy *Secretary* sequence, (b) conventionally filtered sequence, and (c) guided low-pass filtered sequence.

The image in (b) is the result of conventional low-pass filtering. Shown in (c) is the result of the segmentation-guided low-pass filtering.

The image in the middle is blurred due to the indiscriminate averaging of pixel values across all regions during the low-pass filtering process. As a result, noise in the sequence is reduced but with an overall reduction in sharpness.

With the aid of 3-D segmentation, low-pass filtering is performed in a controlled manner. Averaging takes place only within regions, not across region boundaries. Consequently, high-frequency components that are important in maintaining sharpness (those near object and feature boundaries) are retained. At the same time, the noise inside regions is reduced. Rather than appearing blurred, the filtered sequence often appears sharper than the noisy original. Results using the *Secretary* sequence are shown in Fig. 13.

Even though the signal-to-noise ratio (SNR) does not always reflect the quality of the sequence, it is still one of the most frequently used quality metrics (besides the widely used but relatively subjective visual evaluation). The SNR's of the noisy unfiltered sequence, the sequence filtered using the conventional low-pass filter, and the aided low-pass filtered sequence are 6.5, 9.8, and 14.8 dB, respectively, for the *Miss America* sequence. The SNR's of the noisy sequence, the conventional low-pass filtered sequence, and the aided low-pass filtered sequence are 10.42, 9.58, and 14.91 dB, respectively, for the *Secretary* sequence.

B. High-Pass Filtering

Following an argument dual to that used for low-pass filtering, it is desirable to be able to confine high-pass filtering

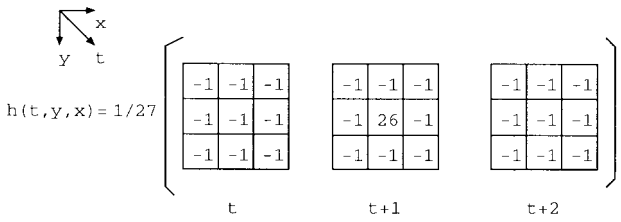


Fig. 14. The 3-D high-pass filtering operator.

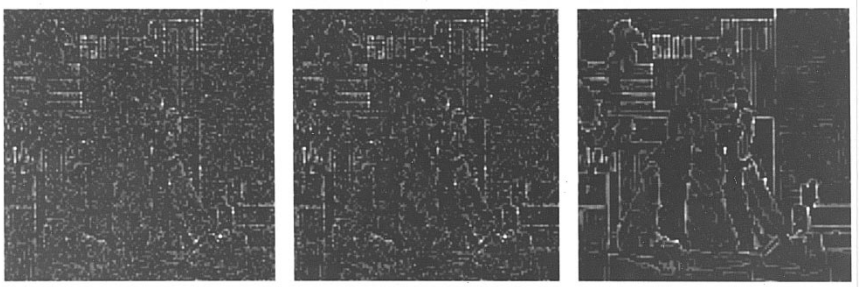
to only areas where edges are present. Region boundaries (due to the fact that they often coincide with object boundaries) are useful in finding these areas. One approach to selective high-pass filtering, then, is to calculate the output of the filter whenever the high-pass operator covers pixels from at least two different regions, where there is a possibility that it is in the vicinity of an edge. Otherwise, the filter output is set to zero.

Three high-pass filtering procedures are compared in this section: conventional high-pass filtering, boundary-constrained high-pass filtering, and the combination of segmentation-guided low-pass filtering and boundary-constrained high-pass filtering.

In conventional high-pass filtering, both edges and noise are enhanced in the image. This is particularly objectionable in smooth areas, where noise is very visible. In boundary-constrained high-pass filtering, only high-frequency components near boundaries are enhanced. This is an improvement over conventional high-pass filtering since the noise inside regions (which corresponds to smooth areas of the image) is

(a) (b) (c)

Fig. 15. (a) Frame 6 from the *Miss America* sequence after conventional high-pass filtering, (b) boundary-constrained filtering, and (c) low-pass/high-pass filtering.



(a) (b) (c)

Fig. 16. (a) Frame 3 from the *Secretary* sequence after conventional high-pass filtering, (b) boundary-constrained filtering, and (c) low-pass/high-pass filtering.

not enhanced. If the segmentation used to guide the filtering process includes false contours, however, the noise around these contours is also enhanced. In very noisy sequences (which tend to result in segmentations with numerous false contours), this can be a major shortcoming.

To reduce the noise around the boundaries, the sequence can be segmentation-guided low-pass filtered first, followed by boundary-constrained high-pass filtering. The combination of these two types of filtering reduces noise much more effectively. The aided low-pass filtering reduces noise near boundaries and in the region interiors, while the boundary-constrained high-pass filtering enhances the edges. Due to the nonlinear nature of the 3-D segmentation-aided filtering, one process does not undo the effects of the other. As a result, the benefits of both processes can be retained. The high-pass operator $h(t, y, x)$ used for high-pass filtering is shown in Fig. 14.

The results obtained by applying the three techniques described above to the noisy *Miss America* sequence are shown in Fig. 15. The image on the left is the result of conventional high-pass filtering. Noise is amplified so much that it overwhelms the image. The image in the middle is the result of the boundary-constrained high-pass filtering. Noise is somewhat reduced. However, because the segmentation was obtained from a noisy image sequence, noise still remains, particularly in the neighborhood of the false segmentation contours. As discussed above, a remedy for this problem is to perform the segmentation-aided low-pass filtering first and then perform the segmentation-aided high-pass filtering. The image on the right is the result of this combination. The improvement over the other two methods is very clear. For the *Secretary* sequence, similar results can be seen in Fig. 16.

C. High-Frequency Emphasis

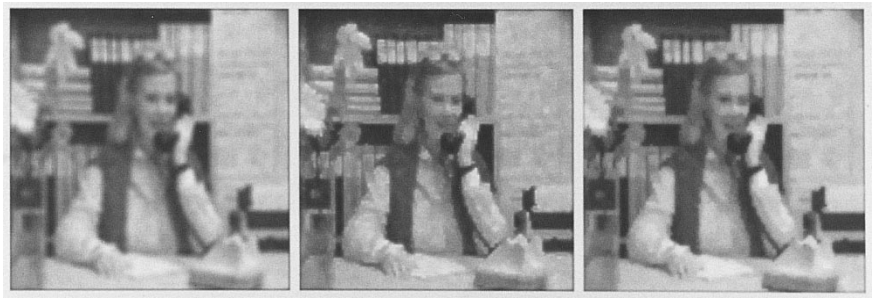
A straightforward approach to sharpening an image sequence is to add a scaled high-pass filtered version of the sequence to the original. Using a similar approach, sharpening can be performed by adding a scaled boundary-constrained high-pass filtered version of the sequence to the original. This second approach allows a more favorable tradeoff between sharpening and noise enhancement. Using conventional high-pass filtering, noise inside regions is amplified and added to the original. In the case of boundary-constrained high-pass filtering, only high-frequency components around region boundaries are enhanced. Consequently, the overall quality of a sequence sharpened with the later approach can be expected to be superior.

Unsharp masking is a technique that subtracts a blurred version of an image from its original. This technique is a basic tool in photographic image processing. In a method that could be considered dual to unsharp masking, which we will refer to as "sharp masking," the segmentation-based approximation described earlier is weighted and added to the original, and the result is rescaled. This process results in sharpening due to the sharp transition between intensity values at region boundaries in the approximation.

For the *Miss America* sequence, the results of the two approaches are shown in Fig. 17. The figure consists of frame 3 of the processed sequence. The image in Fig. 17(a) is

(a) (b) (c)

Fig. 17. (a) Frame 3 from the *Miss America* blurred sequence, (b) result of sharpening using segmentation-guided low-pass filtering followed by boundary-constrained high-pass filtering, and (c) sharp masking approach.



(a) (b) (c)

Fig. 18. (a) Frame 8 from the *Secretary* blurred sequence, (b) result of sharpening using segmentation-guided low-pass filtering followed by boundary-constrained high-pass filtering, and (c) sharp masking approach.

the blurred image, the image in (b) is the result obtained using boundary-constrained high-pass filtering, and the image in (c) is the result produced by sharp masking. The first approach allows a better tradeoff between sharpening and noise enhancement than would be expected from conventional sharpening. In this example, it can be seen that the facial features have been well enhanced over those in the blurred image. The eyes can be clearly seen, and the contrast in the area of the mouth has been increased. There has been only a modest degree of noise enhancement.

The sharp masking approach enhances the high-frequency components in the image such that the face can be seen clearly and at the same time the noise is *reduced.* This is due to the fact that at no point in this method is a difference calculated (the source of noise enhancement in both conventional and segmentation-guided high-pass filtering). Instead, a pointwise weighted average is calculated between the approximation and the sequence. As a result, this method performs significantly better for the sharpening of noisy sequences. Similar results using the *Secretary* sequence are presented in Fig. 18.

D. Sobel Filtering

Following the approach used for high-pass filtering in Section III-B but using a 3-D Sobel operator, edge enhancement can be performed. The Sobel operator is shown in Fig. 19 [10]. This operator is applied to the original sequence three times, oriented along the x-, y-, and t-axes. The resulting sequences contain gradient estimates along their associated dimensions. To suppress noise, outputs are evaluated only at

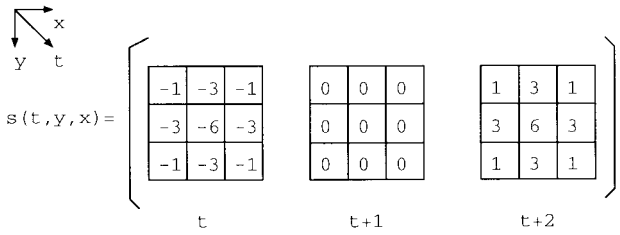


Fig. 19. The 3-D Sobel operator.

positions in the image where the operator spans pixels from at least two different regions. Otherwise, they are set to zero.

In general, the gradient magnitude is calculated as the square root of the sum of the square of the pixel values at each location from the three different sequences

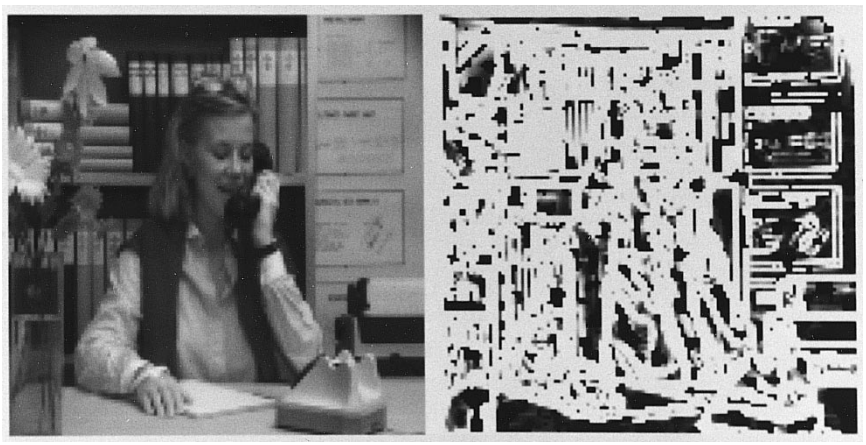
$$\text{Magnitude} = \sqrt{S_x^2 + S_y^2 + S_t^2}. \tag{7}$$

In the results shown below, the t component was omitted ($S_t^2 = 0$) because we did not want to enhance the edges along the temporal axis, which can cause an undesirable impression of false movement.

The results of applying Sobel filtering to the *Miss America* sequence are shown in Fig. 20. The image in Fig. 20(a) is frame 20 from the original sequence. The image in (b) is the same frame after Sobel filtering. For the same reasons we saw improvements in high-pass filtering, the inclusion of segmentation information in the edge-enhancement algorithm allows an improved tradeoff between edge and noise enhancement.

(a) (b)

Fig. 20. (a) Frame 20 from the original *Miss America* sequence and (b) after the Sobel filtering.



(a) (b)

Fig. 21. (a) Frame 3 from the original *Secretary* sequence and (b) after the Sobel filtering.

The results of applying Sobel filtering to the *Secretary* sequence are shown in Fig. 21. The image in Fig. 21(a) is frame 3 from the original *Secretary* sequence. The image in (b) is the same frame after Sobel filtering.

IV. Conclusions

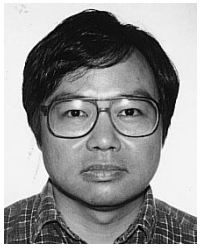
Incorporating the results of 3-D segmentation, four classes of image sequence processing techniques were considered: low-pass filtering, high-pass filtering, high-frequency emphasis, and Sobel filtering. It was demonstrated that including the information about important structure in the sequence that segmentation provides can result in relatively simple algorithms that perform much better than conventional techniques. Specifically, this allows an improved tradeoff with respect to noise. In the case of low-pass filtering, it allows for the simultaneous reduction of noise and sharpening of the image, not possible with most methods. In high-pass filtering, high-frequency emphasis, and edge enhancement, the enhancement of image structure versus noise is also much improved.

The cost of these improvements is the computational power and time required to obtain the 3-D segmentation. However, in some situations, the segmentation may already exist, such as in object analysis and 3-D segmentation-based coding. In those situations, the cost has already been paid. In addition to the techniques discussed here, 3-D segmentation information may be useful in other applications, such as deinterlacing and resolution conversion, as well.

References

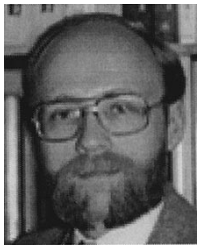
[1] M. Kunt, "Second generation image coding techniques," *Proc. IEEE*, vol. 73, pp. 549–574, Apr. 1985.
[2] Y.-Q. Zhang, F. Pereira, T. Sikora, and C. Reader, Eds., "Special issue on MPEG-4," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, Feb. 1997.
[3] I. Hussain and T. R. Reed, "Segmentation-based nonlinear image smoothing," in *Proc. 1st IEEE Int. Conf. Image Processing*, Austin, TX, Nov. 1994, pp. 507–511.
[4] T. R. Reed, "Segmentation-based image processing," in *Proc. Int. Conf. Digital Signal Processing*, Limassol, Cyprus, June 26–28 1995, vol. 1, pp. 308–313, special sessions.
[5] G. K. Wu and T. R. Reed, "A comparison of image sequence representations based on 2-D and 3-D segmentation," in *Proc. Int. Conf. Digital Signal Processing*, Limassol, Cyprus, June 26–28 1995, vol. 2, pp. 532–537.
[6] R. M. Haralick and L. G. Shapiro, "Survey, image segmentation techniques," *Comput. Vision, Graph., Image Process.*, vol. 29, pp. 100–132, 1985.
[7] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognit.*, vol. 26, no. 9, pp. 1277–1294, 1993.

[8] R. C. Dubes and A. K. Jain, "Random field models in image analysis," *J. Appl. Statist.*, vol. 16, no. 2, pp. 131–164, 1989.
[9] T. Aach and H. Dawid, "Region oriented 3D-segmentation of NMR-datasets: A statistical model-based approach," in *Proc. SPIE Visual Communications and Image Processing*, 1990, vol. 1360, pp. 690–701.
[10] C. Pudney. (Apr. 11, 1995.) 3-D Sobel operator. Available: sci.image.processing newsgroup, article 12170.
[11] D. J. Sakrison, "On the role of observer and a distortion measure in image transmission," *IEEE Trans. Commun.*, vol. COM-25, pp. 1255–1267, Nov. 1977.



Gene K. Wu received the B.S.E.E. degree from San Francisco State University, San Francisco, CA, in 1989 and the M.S. and Ph. D. degrees in electrical engineering from the University of California, Davis, in 1991 and 1997, respectively.

From 1997 to 1998, he was a Lecturer at the University of California, Davis. In 1998, he joined Ricoh Silicon Valley, CA. His research interests include image and image sequence processing and coding and digital signal processing.



Todd R. Reed (S'76–M'77–SM'90) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Minnesota, Minneapolis, in 1977, 1986, and 1988, respectively.

From 1977 to 1983, he was an Electrical Engineer with IBM (San Jose, CA; Rochester, MN; and Boulder, CO). From 1984 to 1986, he was a Senior Design Engineer with Astrocom Corp., St. Paul, MN. He was a Consultant to the MIT Lincoln Laboratory, Cambridge, MA, from 1986 to 1988. In 1988, he was a Visiting Assistant Professor in the Department of Electrical Engineering, University of Minnesota. From 1989 to 1991, he was Head of the Image Sequence Processing Research Group in the Signal Processing Laboratory, Department of Electrical Engineering, Swiss Federal Institute of Technology, Lausanne, Switzerland. From 1998 to 1999, he was a Guest Researcher in the Computer Vision Laboratory, Department of Electrical Engineering, Linköping University, Sweden. In 1991, he joined the Department of Electrical and Computer Engineering, University of California, Davis, where he is currently an Associate Professor. His research interests include image and image sequence processing and coding, multidimensional digital signal processing, and computer vision.

Prof. Reed is a member of the European Association for Signal Processing, the Association for Computing Machinery, the Society for Industrial and Applied Mathematics, Tau Beta Pi, and Eta Kappa Nu.