Robust Multiple Targets Tracking Using Object Segmentation and Trajectory Estimation in Video

Ying-Tung Hsiao[*], Cheng-Long Chuang[#], Joe-Air Jiang[#], and Cheng-Chih Chien[*]

[*]Department of Electrical Engineering
Tamkang University, Taipei, Taiwan, 251, R.O.C.
Tel: +886-2-26215656 Ext. 2786
E-mail: hsiao@ee.tku.edu.tw

[#]Department of Bio-Industrial Mechatronics Engineering
National Taiwan University, Taipei, Taiwan, 106, R.O.C.
Tel: +886-2-33665341
E-mail: jajiang@ntu.edu.tw

***Abstract*—In this paper, a novel robust unsupervised video object tracking algorithm is proposed. The proposed algorithm combines several techniques: mathematical morphology, region growing, region merging, and trajectory estimation, for tracking several predetermined video objects, simultaneously. A modified mathematical morphological edge detector was employed to sketch the contour of the video frame; and an edge-based object segmentation algorithm was applied to the contour for partitioning the predetermined objects; moreover, according to the motion of the objects, the proposed algorithm can estimate and partition the objects in following video frames, automatically. The proposed algorithm is also robustness against mobile cameras. The experimental results show that the proposed algorithm can precisely partition and track multiple video objects.**

I. Introduction

TRACKING several predetermined objects in a sequence of video frames is a important research topic for computer vision and visual surveillance. The object tracking technique is applied to many kinds of applications, such as video coding standard MPEG-4 [1], highway traffic control, and security surveillance. There are several approaches has been proposed for video tracking. The most common way is background subtraction [2]; however, the background subtraction approach is only suitable for tracking objects by stationary camera and unsuitable when the camera is moving. Comaniciu *et al.* [3] use the mean-shift approach to compute the translation of a circular region. Jepson *et al.* [4] use a probabilistic appearance model to compute the affine motion of the object for capturing the stable object features, object shape, and object deals with outliers.

In this work, a morphology-based approach is presented. A modified mathematical morphological edge detector, which is proposed by Hsiao *et al.* [5], is utilized to sketch the contour of the video frame. Hsiao's edge detector is an enhanced version of morphological edge detector, which is capable of sketching out thin edges that ordinary be eliminated by other conventional edge detectors. Moreover, we utilized a novel edge-based image segmentation algorithm to extract the user predetermined objects in the video frame. According to the positions of the objects, a trajectory estimation scheme is proposed to estimate the positions of the objects in the following video frames.

This paper is organized as follows: Section II presents the proposed automatically video object tracking algorithm. Section III provides the experimental results of the proposed approach, and the conclusions are discussed in the last section.

II. The Proposed Algorithm

A. Formal Description

The primary goal of this work is to develop an algorithm that is able to automatically extract the user desired objects in a stream of video frames. The conceptual block diagram of the proposed video object tracking algorithm is as shown in Fig. 1. In this work, the video stream is processed frame by frame. At first, an edge detector is applied to the inputted video frame for sketching out the features of the video frame. Then, the edge-extracted image is processed by the proposed region growing/merging-based image segmentation algorithm for extracting the user desired objects in the video frame. According to the average coordinates of the extracted objects, the proposed trajectory estimation scheme is utilized to estimate the position, and deploy growing seeds at the appropriate position in the next video frame for extracting the same object in the following frames. The detailed information of all procedures in Fig. 1 is discussed in the following subsections:

B. Hsiao's Mathematical Morphological Edge Detector

The edge detector utilized in this work for sketching out the contour of the video frame is one of our earlier works [5]. In this subsection, the edge detector is briefly introduced. The block diagram of the edge detector is as shown in Fig. 2.

In order to increase the contrast level of the original image F, the morphological contrast enhancement process is utilized to F to increase the grey-level difference between dark objects and bright objects. Let B denotes a 3x3-structuring element shows in Fig. 3. Let F_c denote contrast enhanced image can be

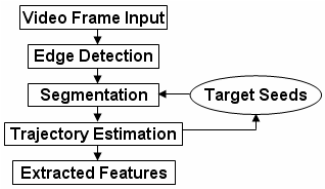


Fig. 1. Conceptual block diagram of the proposed object tracking algorithm.

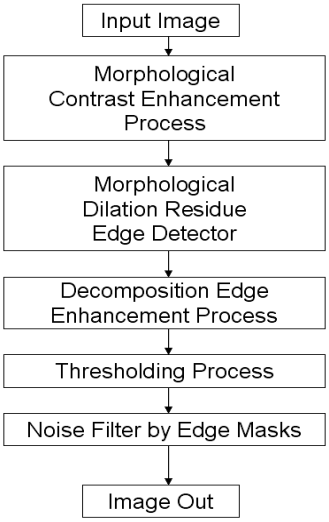


Fig. 2. Flow chart of the utilized edge detection algorithm.

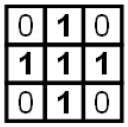


Fig. 3. Structuring element for contrast enhancement and dilation residue edge detection

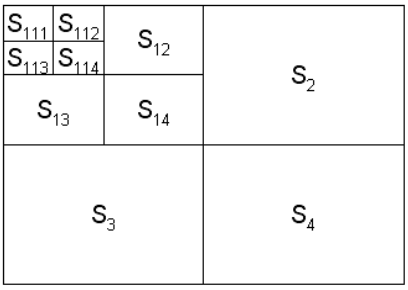


Fig 4. Illustration of quad decomposition of an image.

formulated as follow:

$$F_c = F + WTH(F) - BTH(F) \tag{1}$$

Morphological dilation residue edge detector is applied to the image F_c to extract the edge information. F_e denoted edge detected image can be formulated as follow:

$$F_e = (F_c \oplus B) - F_c \tag{2}$$

Quad decomposition edge enhancement process is applied to the image F_e to sharp thin or smooth edges. In this process, the image F_e is decomposed into four sub-images with same size, and each sub-image decomposed into four sub-images. The process is repeated until the user defined smallest sub-image size N_s is reached. This process is shown in Fig. 4. In this process, we defined four cases for enhancing the thin or smooth edges in a sub-image. Mean and standard deviation in each sub-image are calculated as follows:

$$\mu = \sum_{i=1}^{M}\sum_{j=1}^{N} F_e(i,j) \Big/ M \times N \tag{3}$$

TABLE 1
Edge point enhancement cases.

Case	Condition			Update
	#1	#2	#3	
Redundant Background	$P_{ij}<\mu_s$	$\mu_s<2$	$\delta_s<2$	$P_{ij}=0$
Thin Edge Enhancement	$P_{ij}>\mu_s$ $P_{ij}<\mu_s+1.5\times\delta_s$	$\mu_s<2$	$\delta_s<2$	$P_{ij}=P^2$
Intensive Thin Edge Enhancement	$P_{ij}>\mu_s+1.5\times\delta_s$ $P_{ij}<\mu_s+3\times\delta_s$	$\mu_s<2$	$\delta_s<2$	$P_{ij}=P^3$
Thin Edge in Complex Background	$P_{ij}>\mu_s+1.5\times\delta_s$	$\mu_s>5$	$\delta_s>6$	$P_{ij}=P^2$

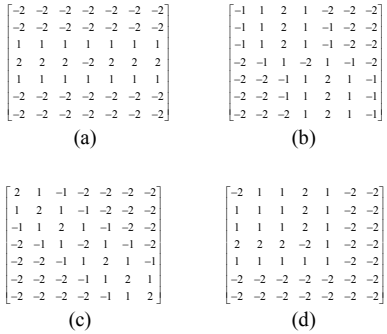


Fig. 5. Bipolar oriented edge masks. (a) for angel 0 and π/2, (b) for angel 3π/8, and 5π/8, (c) for angel π/4 and 3π/4, and (d) for corner edge

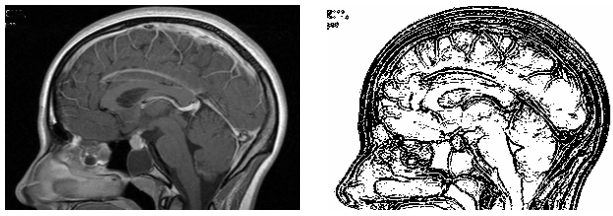


Fig. 6. Extraction result of the utilized edge detector.

$$\delta = \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N} (F_e(i,j)-\mu)^2} \Big/ M \times N \tag{4}$$

where M and N are width and height of the sub-image, espectively. By these mean and standard deviation values of each sub-image, we defined four edge point enhancement cases as Table 1. Let F_q denotes the image combined from all processed sub-image.

A threshold value ranging from 20 to 30 is applied to the image F_q and producing the binary image F_t with the clearly thin edges and without causing too many noise points.

Moreover, there are five bipolar oriented edge masks shown in Fig. 5 are applied to the image F_t for removing the unwanted noise pixels, bit the mask only applied to the pixel that satisfies the following terms:

1) The pixel is thresholded to be an edge point.

2) The sub-image contains the pixel and its mean and standard deviation obtained in *Step 4* are $\mu<3$ and $\delta<3$, respectively.

Then, since the processes are done, a high performance edge extracted image is produced as demonstrated in Fig. 6.

C. *Object Segmentation in Video Frame*

In this subsection, a novel image segmentation algorithm is

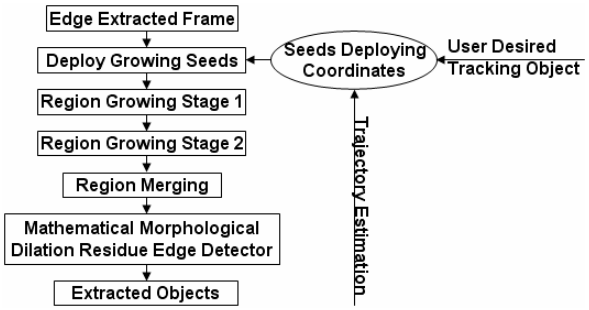


Fig. 7. Flow Chart of the proposed object segmentation algorithm.

briefly introduced. The goal of region growing and region merging is to divide the domain R of the image F into regions $\{R_i,\ i = 1\ldots\ m\}$ so that $R = \cup_{i=1}^{m} R_i$, and $R_i \cap R_j = \varnothing$ if $i \neq j$, and F satisfy the homogeneity criterion on each R_i.

The block diagram of the proposed object segmentation technique is as shown in Fig. 7. The coordinates of the initial growing seeds for each frame must be defined by user or the trajectory estimation scheme to extract desired objects in a series of video frames. Each growing seed owns a unique identification number. Since the seeds are deployed, the stage one of the region growing process is then started. These seeds shall start to extend its area and add the empty pixel (non-edge pixel) near to the seed to form their own regions. However, because the edge points in the edge-extracted image might not always be continuously extracted by morphological edge detector. Therefore, some gaps may exist that may cause the growing process produces a useless result. In this work, a boundary condition, similar to the snake model, is introduced to control the region growing process. The proposed boundary condition is different from the snake model approach because the first one controls the growing process of each region, and the last one controls the reaching process of an edge vector. The boundary condition E_{Cost} consists of two performance indexes given by:

$$E_{Cost} = \sum_{i=1}^{N} \left(E_e(i) + E_r(i) \right) \tag{5}$$

where $E_e(i)$ and $E_r(i)$ are edge shape cost and reached edge cost at the ith position of the contour of the region, respectively, which can be defined as follow:

$$E_e(i) = \alpha_i \left| s_i - s_{i-1} \right|^2 + \beta_i \left| s_{i-1} - 2s_i + s_{i+1} \right|^2 \tag{6}$$

$$E_r(i) = -\left| \left(255 - F(s_i) \right) / 255 \right| \tag{7}$$

where $F(s_i)$ is the pixel value of the edge-extracted image in position s_i, N represents the number of boundary point, and the order of the boundary points is counterclockwise along the contour of the region. s_i is the coordinate of the ith boundary. The weighting factors α_i and β_i are defined to control the importance of the membrane and thin-plate term for the ith boundary point, respectively. By minimizing the boundary

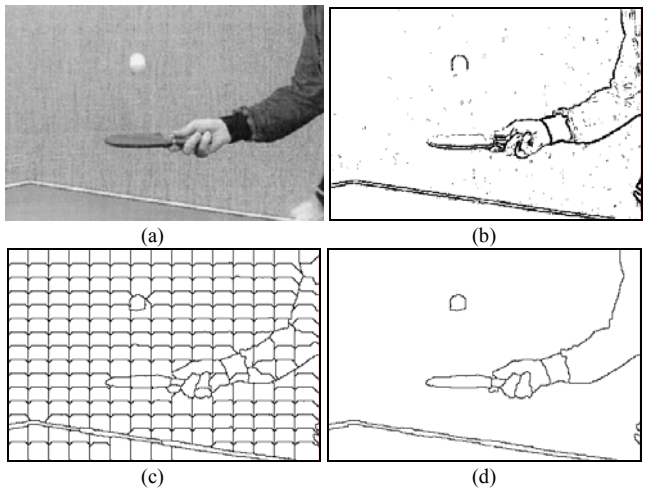


Fig. 8. (a) A frame of Table Tennis video clip. (b) Edge detection result by Hsiao's mathematical morphological edge detector. (c) Boundary of the regions after region growing. (d) Final result after region merging process.

condition E_{Cost}, the region is attracted to grow to the boundary with the lowest cost, such as edge lines. Therefore, the boundary of each region converged on a set of meaningful edge points instead of growing through gaps of the non-continuous edges lines.

Since the first stage of region growing process is completed, the second stage also lets the regions keep growing, but the second stage allows the regions add the 8-neighborhood pixels (edge points) without under controlling of boundary condition.

The process of region merging is aimed to reduce redundant regions that do not have enough edge points between them. For a pair of regions, that the number of the edge points reached by both of them is smaller than 10 percents of the boundary length, these two regions will merge into single region. Since the process of region merging is completed, the segmentation line can be extracted as shown in Fig. 8(d).

D. Trajectory Estimation Scheme

A trajectory estimation scheme is proposed in this section. As mentioned in last subsection, we can deploy growing seeds to extract desired objects. Therefore, when we want to track the desired objects in a sequence of frames, it is necessary to develop a trajectory estimation scheme for automatically deploying seeds into the incoming frames.

The principle of the trajectory estimation scheme is quite simple. First, we calculate the average coordinate $X_O(t)$ and $Y_O(t)$ of an object O on x axis and y axis in frame t is defined as follow

$$X_O(t) = \left(\sum_{p \subset S_o} x_p \right) \Big/ \left\| p \subset S_o \right\| \tag{8}$$

$$Y_O(t) = \left(\sum_{p \subset S_o} y_p \right) \Big/ \left\| p \subset S_o \right\| \tag{9}$$

where p is the pixel in the frame, S_O is the pixel set of object O, and x_p, y_p is the coordinate of pixel p on x axis and y axis. And

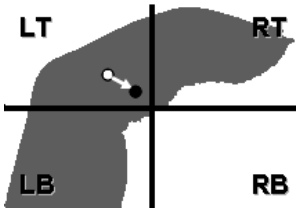


Fig. 9. Illustration of trajectory estimation scheme. The area of extracted object is represented in grey color.

the velocities $Vx_O(t)$ and $Vy_O(t)$ of object O on x axis and y axis is obtained by the difference of average coordinates between two continuous frames t-1 and t as follow

$$Vx_O(t) = X_O(t-1) - X_O(t) \tag{10}$$

$$Vy_O(t) = Y_O(t-1) - Y_O(t) \tag{11}$$

We assume that the object O is as the grey area in Fig. 9. After we obtained the moving velocity of the object O in frame t, the velocity information can be used to determine the position of the initial growing seed for segmenting the same object O in frame t+1. The object O, which is segmented in frame t, is partitioned into four parts: left-top (LT) part, left-bottom (LB) part, right-top (RT) part, and right-bottom (RB) part, respectively. We select three parts among the four partitioned parts, which can achieve the maximum summation of the object O, and the corner part of these pre-selected three parts is used to calculate the basis coordinate for determining the position of the initial growing seed for segmenting the same object O in next frame t+1. The basis coordinate $X_{B\_O}(t)$ and $Y_{B\_O}(t)$ of an object O on x axis and y axis in frame t is defined as follow

$$X_{B\_O}(t) = \left(\sum_{p \subset S_o \cap \text{Part}} x_p \right) \Big/ \left\| p \subset S_o \cap \text{Part} \right\| \tag{12}$$

$$Y_{B\_O}(t) = \left(\sum_{p \subset S_o \cap \text{Part}} y_p \right) \Big/ \left\| p \subset S_o \cap \text{Part} \right\| \tag{13}$$

As shown in Fig. 9, maximum area summation of the object O can be achieved by combining with RT, LT, and LB. And the LT part plays a corner role in the combination. Therefore, the "Part" defined in (12) and (13) is the LT part, and where $(X_{B\_O}(t), Y_{B\_O}(t))$ is the basis coordinate, which is as the white spot in Fig. 9.

When we obtained the velocity of the desired objects and their basis coordinate, the position $x_{N\_O}(t+1)$ and $y_{N\_O}(t+1)$ of the next initial growing seeds for segmenting the same objects O in next frame can be obtained as follows

$$x_{N\_O}(t+1) = X_{B\_O}(t) + Vx_O(t) \tag{14}$$

$$y_{N\_O}(t+1) = Y_{B\_O}(t) + Vy_O(t) \tag{15}$$

Therefore, when the frame t+1 is been obtained, the proposed scheme can automatically deploy growing seeds for extracting the same desired objects in the frame t+1, as the black spot in Fig. 9, and the white arrow is the moving velocity of the object.

III. Experimental Results

A. Tracking by Mobile Camera

A sequence of video frames that acquired by a non-stationary camera is used to evaluate the tracking performance of the proposed algorithm. In the video stream, as shown in Fig. 10, a bottle is placed on a desk, and the white cap of the bottle is the target object. Because the video stream is acquired by a non-stationary camera, the background is constantly changing during the tracking process. The tracking result shows that the proposed tracking algorithm is capable of precisely tracking the complete region of a predetermined object, and tracking performance isn't influenced by the changing background.

B. Tracking with Complex Background 1

A video clip, as shown in Fig. 11, records a finger slide over a bottle, is used to evaluate the tracking performance of the proposed algorithm. The tracking target is the complete region of the finger. Although the bottle has a lot of Chinese words printed and light reflection on it, the proposed algorithm still tracks the finger very precisely.

C. Tracking with Complex Background 2

Fig. 12 shows another color video clip that records a hand taking off a bottle from the desk, which is used to evaluate the tracking performance of the proposed algorithm. The tracking target in this video clip is the cap of the bottle. As shown in the figure, the region of the cap moves through the complex background that consists of many instruments. In this test, the proposed algorithm is still able to successfully track the cap, and the tracking performance has not been influenced by the complex background in the video stream.

D. Multiple Targets Tracking by Mobile Camera

In this paragraph, the multiple objects tracking ability of the proposed algorithm is evaluated. A video clip depicted in Fig. 13 records that a hand takes a bottle moving over a silver box. The tracking targets in this example are the cap of the bottle and the silver box. In this video clip, the silver box is placed in a fixed position on the desk, the hand takes the bottle from left side to right side in the sense of the camera, and the sense of the camera is moving from right to left. Therefore, not only the background is constantly changing, but also the two tracking targets are overlapping with each other during the tracking process. The tracking result shows that the proposed algorithm can produce good tracking performance on tracking multiple objects in the video that acquired by a non-stationary camera.

E. Video from opening ceremony of Games of the XXVIII Olympiad Athens 2004

A video stream in Fig. 14 is used to evaluate the tracking performance of the proposed algorithm. The video stream is a part of the TV program that was playing the opening ceremony of Games of the XXVIII Olympiad Athens 2004. The tracking target is the face of Gianna Angelopoulos-Daskalaki, who was the president of the Athens Olympic Organizing Committee (ATHOC). While the Olympiad song was singing, the sense of the camera shifts through Gianna. The tracking result, as depicted in Fig. 14, shows that the proposed tracking algorithm is

able to successfully produce the region mask for extracting the face of Gianna.

F. Video clips 1 from movie: I Robot

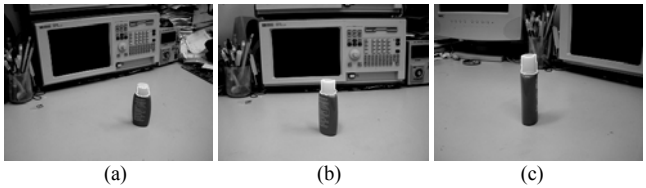
A video stream, captured from the movie called "I Robot", is used to demonstrate the performance of the proposed tracking algorithm. In this video stream, as shown in Fig. 15, the tracking target is the face of Will Smith. His face was not moving, but because Will was speaking, the shape of his face is changing at all time. The tracking result shows that the proposed algorithm can successfully track the face of Will Smith.

G. Video clips 2 from movie: I Robot

The last video stream, captured from the movie called "I Robot", is used to evaluate the performance of the proposed tracking algorithm. In this video stream, as shown in Fig. 16, the tracking target is the head of a robot. Because both of the robots are fighting to each other, the position of the head of the robot was changing at high speed. The tracking result shows that the proposed algorithm is robust against objects movement with high velocity.

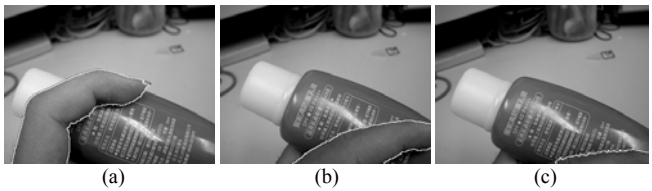
IV. CONCLUSION

A hybrid of mathematical morphology, region growing and merging, and trajectory estimation scheme approach is proposed. At first, user can predetermine the desired objects for tracking, and then the video frame is processed by the Hsiao's edge detector [4] to sketch the contour the frame. The contour information is then processed by the proposed segmentation algorithm, which consists of region growing and region merging processes. According the extracted object, we utilized the proposed trajectory estimation scheme for automatically deploying the growing seeds for tracking the same object in further incoming video frames. The experimental results show that the tracking performance is guaranteed.



(a) (b) (c)

Fig. 10. Mobile camera tracking. (a) frame 1, (b) frame 30, and (c) frame 60 in the frames sequence A.



(a) (b) (c)

Fig. 11. Finger moving on complex background. (a) frame 1, (b) frame 20, and (c) frame 30 in the frames sequence B.

V. ACKNOWLEDGMENTS

The authors are grateful to the National Science Council of the Republic of China for financially supporting this research under contract no. NSC 93-2213-E-032-029.

(a) (b) (c)

Fig. 12. Bottle moving on complex background. (a) frame 1, (b) frame 20, and (c) frame 30 in the frames sequence C.

(a) (b) (c)

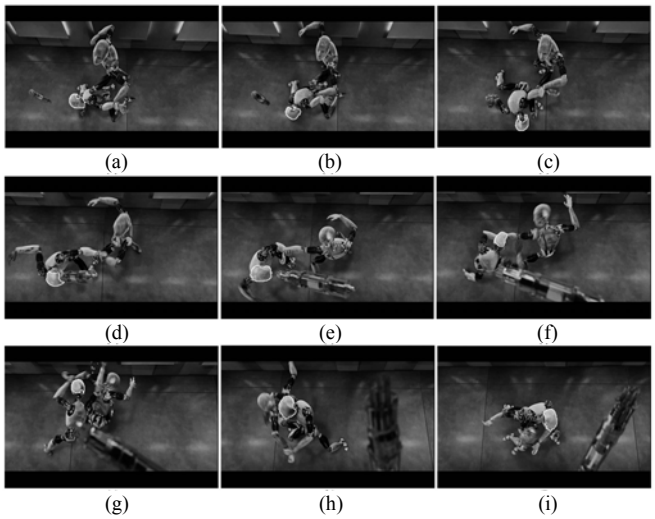
Fig. 13. Multiple objects tracking by a moving camera. (a) frame 1, (b) frame 30, and (c) frame 45 in the frames sequence D.

(a) (b) (c) (d)

Fig. 14. Tracking result of the horizontal moving face. (a) frame 1, (b) frame 40, (c) frame 85, and (d) frame 135 in the sequence of frames.

(a) (b) (c) (d)

Fig. 15. Tracking result of the stationary face tracking I with greater changing on the face features. (a) frame 1, (b) frame 8, (c) frame 19, and (d) frame 26 in the sequence of frames.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Fig. 16. Tracking an object at high moving velocity. (a) frame 1, (b) frame 16, (c) frame 32, (e) frame 48, (f) frame 64, (g) frame 80, (h) frame 96, (i) frame 112, and (j) frame 128 in the sequence of frames.

REFERENCES

[1] SIKORA, T.: "The MPEG-4 Video Standard verification model." *IEEE Trans. Circuits Syst. Video Technol*., vol. 1, July, 1997, pp. 19-31.

[2] C. Stauffer and W. Grimson, "Learning Patterns of Activity Using Real Time Tracking," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747-767, Aug. 2000.

[3] D. Comaniciu, V. Ramesh, and P. Meer, “Kernel-Based Object Tracking,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 5, pp. 564-575, May 2003.
[4] A.D. Jepson, D.J. Fleet, and T.F. El-Maraghi, “Robust Online Appearance Models for Visual Tracking,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, Oct. 2003.
[5] Y.T. Hsiao, C.L. Chuang, S.H. Yen, and H.J. Lin, "A Mathematical Morphological Approach to Thin Edge Detection in Dark Region," *Proceedings of the 4th IEEE International Symposium on Signal Processing and Information Technology*, Rome, Italy, Dec. 2004.
[6] J. Serra, *Image Analysis and Mathematical Morphology*, New York: Academic. 1982.
[7] P. Soille, *Morphological Image Analysis: Principles and Applications*, Springer: Berlin Heidelberg, 1999.
[8] J. Haddon and J. Boyce, “Image segmentation by unifying region and boundary information,” *IEEE Trans. Pattern Anal. Machine Intell*., vol. 12, pp. 929–948, 1990.
[9] T. Pavlidis and Y. Liow, “Integrating region growing and edge detection,” *IEEE Trans. Pattern Anal. Machine Intell*., vol. 12, pp. 225–233, 1990.
[10] T. Chen, Q. H. Wu, R. Rahmani-Torkaman, and J. Hughes, "A pseudo top-hat mathematical morphological approach to edge detection in dark region," Pattern *Recognition 35*, pp.199-210, 2002.
[11] J. H. Bosworth and S. T. Acton, "Morphological Image Segmentation by Local Monotonicity," *Proc. Conf. on Signals, Systems, and Computers*, Vol 1 , pp.53-57, Oct. 1999.
[12] M. Basu, "Gaussian-Based Edge-Detection Methods-A Survey," *IEEE Trans. System, Man, and Cybernetics-Part C: Application and Reviews*, Vol. 32, No. 3, pp.252-260, Aug. 2002.
[13] A. Trujillo, "Thin Edge Detector," *Proc. Int. Conf. on Image Analysis and Processing*, pp. 1051-1054, Sept. 1999.
[14] M. M. Ratnam, "Automatic Edge Detection in Noisy Fringe Patterns," *Proc. Int. Symp. Signal Processing and its Applications*, Vol. 1, pp. 140-143, Aug. 2001.
[15] S. Shah and P. S. Sastry, "Fingerprint Classification Using a Feed-back-Based Line Detector," *IEEE Trans. Systems, Man and Cybernetics, Part B*, Vol. 34, No. 1, pp. 85-94, Feb. 2004.
[16] J. E. Abdou, "Quantitative methods of edge detection," *USC Report, 830*, July, 1978.
[17] Soille, *Morphological image analysis: principles and applications*, Springer, 1999.
[18] J. Yoo, C.A. Bouman, E.J. Delp, E.J. Coyle, "The nonlinear prefltering and difference of estimates approaches to edge detection," *CVGIP: Graph. Models Image Process*, pp. 140-159, 1993.
[19] R.C. Hardie, C. Boncelet, "LUM filters: a class of rank-order-based filters for smoothing and sharpening," *IEEE Trans. Signal Processing*, vol. 41, No. 3, pp. 1061-1076, March 1993.
[20] J. Huddleston and J. Ben-Arie, "Grouping edgels into structural entities using circular symmetry, The distributed Hough transform and probabilistic nonaccidentalness," *J. Comput. Vis., Graph., Image Process: Image Understand*., vol. 57, o. 2, pp. 227–242, Mar. 1993.
[21] V. Murino, C. S. Regazzoni, and G. L. Foresti, "Grouping as a searching process for minimum-energy configurations of labeled random fields," *Computer Vision Image Understanding. New York: Academic*, 1996, pp. 157–174.
[22] T. Kanungo, B. Dom, W. Niblac, and D. Steele, "A fast algorithm for MDL-based multi-band image segmentation," *Proc. IEEE Computer Vision Pattern Recognition*, Seattle, WA, June 1994, pp. 609–616.
[23] J. Shi and J. Malik, "Normalized cuts and image segmentation," *Proc. 1997 IEEE Computer Soc. Conf. Computer Vision Pattern Recognition*, San Juan, Puerto Rico, June 1997, pp. 731–737.
[24] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 696–710, July 1997.
[25] Y.-L. Chang and X. Li, "Adaptive image region-growing," *IEEE Trans. Image Processing*, vol. 3, pp. 868–872, 1994.
[26] C. Chu and J. K. Aggarwal, "The integration of image segmentation maps using region and edge information," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 1241–1252, 1993.
[27] J. Haddon and J. Boyce, "Image segmentation by unifying region and boundary information," IEEE *Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 929–948, 1990.
[28] L. Leonardis, A. Gupta, and R. Bajcsy, "Segmentation of Range Images as the Search for Geometric Parametric Models," *Int. J. Computer Vision*, vol. 14, no. 3, pp. 253-277, 1995.
[29] R.V. Babu, K.R. Ramakrishnan, and S.H. Srinivasan, "Video Object Segmentation: A Compressed Domain Approach," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 14, no. 4, pp. 462-474 , April, 2004.