Binarization Method Based on Pixel-level Dynamic Thresholds for Change Detection in Image Sequences

HSU-YUNG CHENG, QUEN-ZONG WU*, KUO-CHIN FAN AND BOR-SHENN JENG*
Institute of Computer Science and Information Engineering
National Central University
Taoyuan, 320 Taiwan
**Chunghwa Telecommunication Laboratories*
Taoyuan, 326 Taiwan

A new change detection scheme with enhanced tolerance to noises and illumination changes is proposed in this paper. In addition to computing a global threshold for an entire target image, a dynamically adjusted cumulative histogram is computed to find the most suitable threshold for every pixel in the target image. A binarization method is designed to integrate the global threshold and the dynamic thresholds. The adoption of pixel-level dynamic thresholds makes the proposed method more robust. Experimental results show that the proposed scheme is able to locate motion pixels more accurately and work well under various kinds of illumination conditions.

Keywords: change detection, dynamic cumulative histograms, threshold selection, pixel-level dynamic thresholds, image binarization, video analysis

1. INTRODUCTION

Change detection detects moving objects in the background scenes in image sequences. It is a fundamental and crucial task in many applications, such as tracking [1], moving object segmentation [2], and traffic flow analysis [3]. The image sequences to be handled can be taken by either an ego-motion camera or a fixed camera. For ego-motion cameras, there is no stable scene in the video. In this case, moving objects are often found by estimating the ego-motion through techniques based on block matching [4, 5]. In videos taken by fixed cameras, the background scenes are quite stable. In this case, detecting changes based on background differencing is common. Although traditional non-adaptive backgrounds may accumulate errors and result in incorrect detection results, recent research has produced many background modeling techniques. Some earlier methods model background pixel intensities using minimum and maximum intensity values [6] or Kalman filters [7]. More sophisticated models, such as Gaussian models [8, 9] and multi- modal statistical models [10], have also been developed lately to provide more accurate and adaptive backgrounds. Therefore, noises caused by imaging effects or camera jitter can be reduced, and change detection based on background differencing becomes feasible [11]. In this paper, we focus on change detection in videos with stable scenes. To detect changes based on background differencing, target images are produced,

Received January 1, 2005; accepted June 30, 2005.
Communicated by C. C. Chiang, H. Y. Mark Liao and K. C. Fan.

and then thresholding is performed to binarize the target images. Target images are generated by means of background subtraction based on standard intensities or logarithmic intensities [12]. Because difference images based on logarithmic intensities have better quality for thresholding, we adopt the approach proposed in [12] as the basis of our change detection algorithm. The target images are generated according to the following formula:

$$I_{target}(p) = \exp(\log(I_{current}(p)) - \log(I_{background}(p))). \quad (1)$$

When binarization is performed, histograms are very useful for finding thresholds. Many histogram-based thresholding approaches find a threshold in a histogram by approximating the histogram with two or more distribution functions [13, 14]. However, these approaches have the drawback of slow convergence and are sometimes infeasible for finding proper thresholds of sharply fluctuating histograms.

In [15], a motion detection algorithm that applies change-point detection to cumulative histograms was proposed. Since cumulative histograms do not fluctuate, the threshold finding procedure is much simpler when cumulative histograms are involved. As Fig. 1 shows, point X_i is selected as the threshold for a target image where the distance between (X_i, Y_i) and line L is maximized, for all $2 \leq i \leq N - 1$. N is the total number of bins in the cumulative histogram. The distance between (X_i, Y_i) and line L: $a \cdot X + b \cdot Y + c = 0$ is defined as

$$D((X_i, Y_i), L) = \frac{|a \cdot X_i + b \cdot Y_i + c|}{\sqrt{a^2 + b^2}}. \quad (2)$$

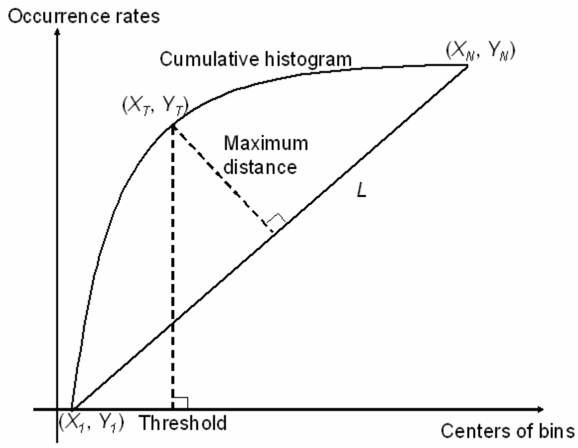


Fig. 1. Change-point detection for cumulative histograms for threshold selection.

Thus, the computational complexity for selecting the threshold of a cumulative histogram can be significantly reduced compared to that of traditional histogram-based methods. However, this approach mostly utilizes the global information of the target image, and the local details are neglected. In fact, just as each image frame should have a different threshold value, the threshold most suitable for every pixel in the same image frame also varies. The application of pixel-level dynamic thresholds can substantially reduce the detection errors. In this paper, we propose a method that not only retains the benefits of cumulative histograms but also includes local information so that a threshold most suitable for every pixel can be selected.

2. THRESHOLD SELECTION AND BINARIZATION VIA DYNAMIC CUMULATIVE HISTOGRAMS

The procedure applied in the proposed thresholding method is illustrated in Fig. 2. This procedure can be organized into four steps. Before explaining each step in the thresholding procedure, we will first explain how a cumulative histogram of an image is obtained. A cumulative histogram is an increasing or monotonic increasing function. Each sample point in a cumulative histogram accumulates the number of occurrences of all sample points whose values are less than or equal to its value, A cumulative histogram can be represented as $\{C(i)\}$, $i = 1, 2, \ldots, N$, where N is the total number of bins. $C(i)$ is the accumulated occurrences in bins 1, 2, …, i. We can express $C(i)$ as $C(i) = \sum_{k=1}^{i}(occurrences\ \ in\ \ bin_i)$. The bins have is a range of values. For example, if $N = 16$, which means that we divide the intensity range [0, 255] into 16 bins, then the first bin has the range [0, 15], the second bin has the range [16, 31], and so on. We define the center value of a bin that is indexed i, ranging from ℓ_i to u_i, as $\left\lfloor \frac{\ell_i + u_i}{2} \right\rfloor$. For the example given here, the center value for the first bin would be 7, and the center value for the second bin would be 23. Now, we will explain the thresholding procedure step by step.

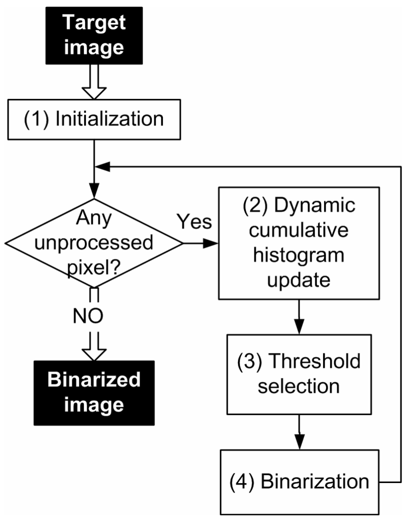


Fig. 2. Procedure applied in the proposed thresholding method.

Step 1: Initialization.

The cumulative histogram $\{C(i)\}$ is computed. All the bins are normalized so that $C(i)$ represents the occurrence rate of the bin indexed i. A global threshold is computed by using two line segments to approximate $\{C(i)\}$. The approximation method is the same as that in step 3, which will be described later. After initialization, steps 2-4 are performed for every unprocessed pixel p in the target image until every pixel is binarized.

Step 2: Dynamic cumulative histogram update.

In a cumulative histogram, a pixel p contributes to the bins whose center values are greater than or equal to the value of p. In order to focus on the occurrence of the value of p, an adjustment parameter, Δr, is added to the occurrence rate of the bins that p contributes to. In addition, an adjustment parameter, Δr, is subtracted from the occurrence rate of the bins that p makes no contribution to. To summarize, the dynamic cumulative histogram $\{DC(i)\}$ is updated according to the following equation:

$$\begin{aligned} DC(i) &= \max(0,\ C(i) - \Delta r),\ \text{if } center(i) < value(p), \\ DC(i) &= \min(1,\ C(i) + \Delta r),\ \text{if } center(i) \geq value(p), \end{aligned} \tag{3}$$

where $center(i)$ represents the center value of the bin that is indexed i. Through this step, the local information for this pixel can be included in the dynamic cumulative histogram. Note that the adjustment parameter for the occurrence rate is used to reflect the local property of the current pixel. If Δr is too big, the thresholds will fluctuate too much and lack continuity. If the scene is stable, Δr can be fixed. When selecting a suitable Δr, users may try a small Δr first and then increase Δr gradually until the thresholds fluctuate too much and cause the binarization results to degrade.

Step 3: Dynamic threshold selection.

To find the suitable dynamic threshold for pixel p, we use two line segments to approximate the dynamic cumulative histogram because the transition between two states, i.e., no-motion and motion, can be featured clearly by cumulative histograms of ratio images. As Fig. 3 shows, for any sample point j, we can obtain two line segments $L1$ and $L2$ by linearly interpolating $(1, DC(1))$ to $(j, DC(j))$ and $(j + 1, DC(j + 1))$ to $(N, DC(N))$, respectively. $L1(i)$ is the line segment used to approximate the left portion of the cumulative histogram, and $L2(i)$ is the line segment used to approximate the right portion of the cumulative histogram. $L1(i)$ and $L2(i)$ intersect at the dynamic threshold. The dynamic threshold is the center value of the index j that minimizes D_T:

$$D_T = \sum_{i=1}^{j} |DC(i) - L1(i)| + \sum_{i=j+1}^{N} |DC(i) - L2(i)|. \tag{4}$$

Step 4: Binarization.

When binarization is performed, two criteria are considered to highlight a pixel as a motion pixel. Criterion-A states that the value of the pixel should exceed its own dynamic threshold. If, for a pixel p, criterion-A is not satisfied, then criterion-B is further checked.

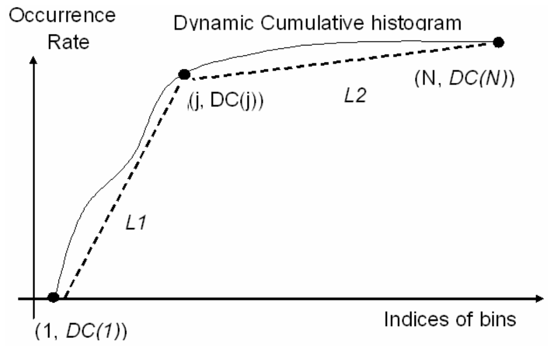


Fig. 3. Approximating the dynamic cumulative histogram by means of two linear segments in threshold selection.

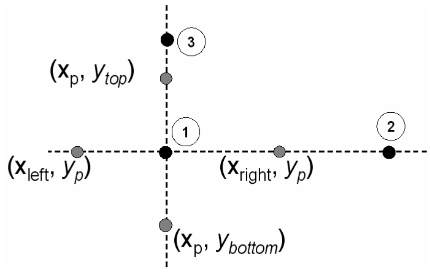


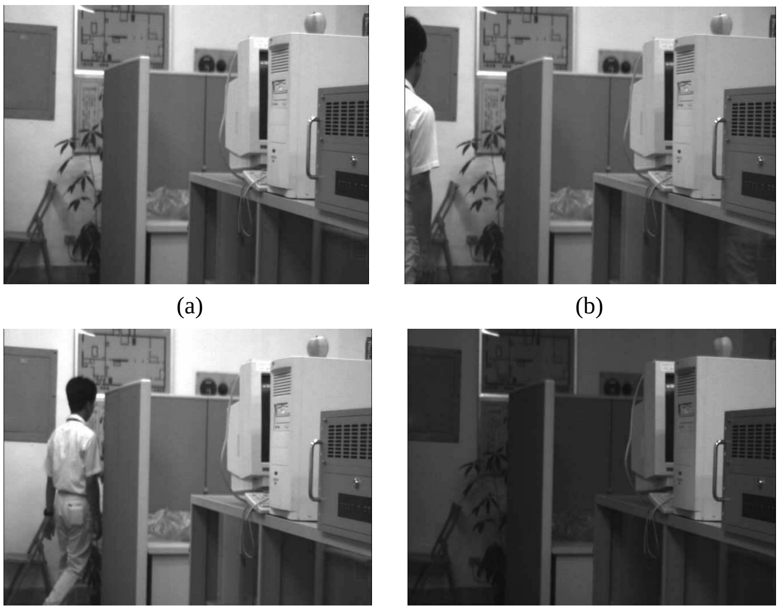
Fig. 4. An example illustrating the relations among (x_p, y_p), (x_{left}, y_p), (x_{right}, y_p), (x_p, y_{top}), and (x_p, y_{bottom}).

Criterion-*B* is satisfied when the value of *p* exceeds the global threshold and the position of *p* is within a proper range relative to the some specific pixels that satisfy criterion-*A*; more specifically, $x_{left} < x_p < x_{right}$ and $y_{top} < y_p < y_{bottom}$. For $p(x_p, y_p)$, x_{left} and x_{right} are defined as the *x*-coordinate values of the leftmost and rightmost pixels in row y_p such that their values exceed their dynamic thresholds; y_{top} and y_{bottom} are similarly defined. An example that illustrates criterion-*B* and the definition of x_{left}, x_{right}, y_{top}, and y_{bottom} is shown in Fig. 4. Suppose that the value of $p(x_p, y_p)$ is larger than the global threshold. If $p(x_p, y_p)$ is at position-1, then criterion-*B* is satisfied. If *p* is at position-2 ($x_p > x_{right}$) or at position-3 ($y_p < y_{top}$), then criterion-*B* is not satisfied. The value of *p* is set to 255 if either criterion-*A* or criterion-*B* is satisfied for *p*; otherwise, it is set to 0. A pixel is marked as processed after it has been binarized.

3. EXPERIMENTAL RESULTS

In order to demonstrate the robustness of the proposed method, experimental image sequences with different illumination conditions were tested. Selected image frames from three experimental videos will be discussed in this section. In the first experiment, we used an image sequence taken under normal illumination conditions in an indoor setting. In the second and third experiments, videos with especially bright and dark scenes were tested.

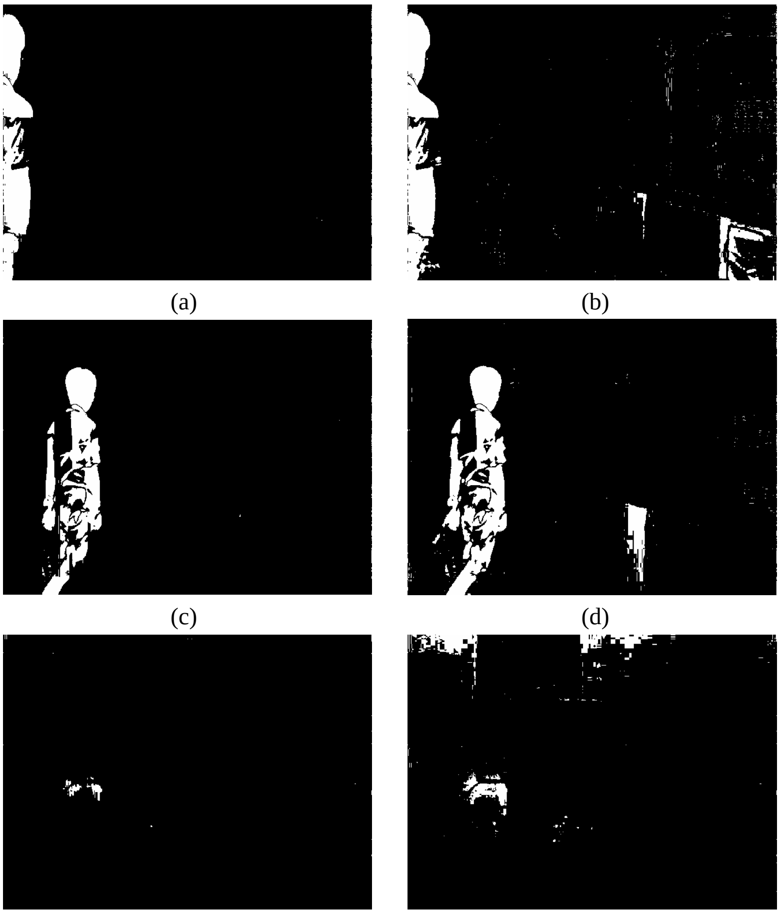
Fig. 5 shows four selected image frames from the video tested in the first experiment. Fig. 5 (a) shows the background image. Figs. 5 (b), (c) and (d) are three image frames selected from a sequence showing a person who passed through an aisle and then turned off a light. There is no moving object in Fig. 5 (d), yet its illumination is quite different from that of the background image. The change detection results obtained with the proposed method are displayed in Figs. 6 (a), (c) and (e). The results obtained via change- point detection for cumulative histograms [15] are displayed in Figs. 6 (b), (d) and (f). Compared with the results shown in Figs. 6 (b), (d) and (f), those shown in Figs. 6 (a), (c) and (e) are obviously more accurate. Almost no non-motion pixels are misclassified as motion pixels in Figs. 6 (a) and (c). And there are many fewer misclassified pixels in Fig. 6 (e). Note that the adjustment parameter, Δr, was set to 0.06 in all of the experiments.



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

Fig. 5. Selected image frames from experimental video 1.

Fig. 7 displays four selected image frames from the second experimental video. This image sequence was taken in a corridor with sufficient light. The light reflected by the glass brightens the entire scene. In addition, the reflection of the moving person also affects the change detection results. In the results obtained via change-point detection for cumulative histograms [15] displayed in Figs. 8 (b), (d) and (f), we can see that the interference caused by the light reflected by the glass is quite obvious. When we applied the proposed method, the unwanted effects caused by the reflections were significantly suppressed, as shown in Figs. 8 (a), (c) and (e).



(a) (b) (c) (d) (e) (f)

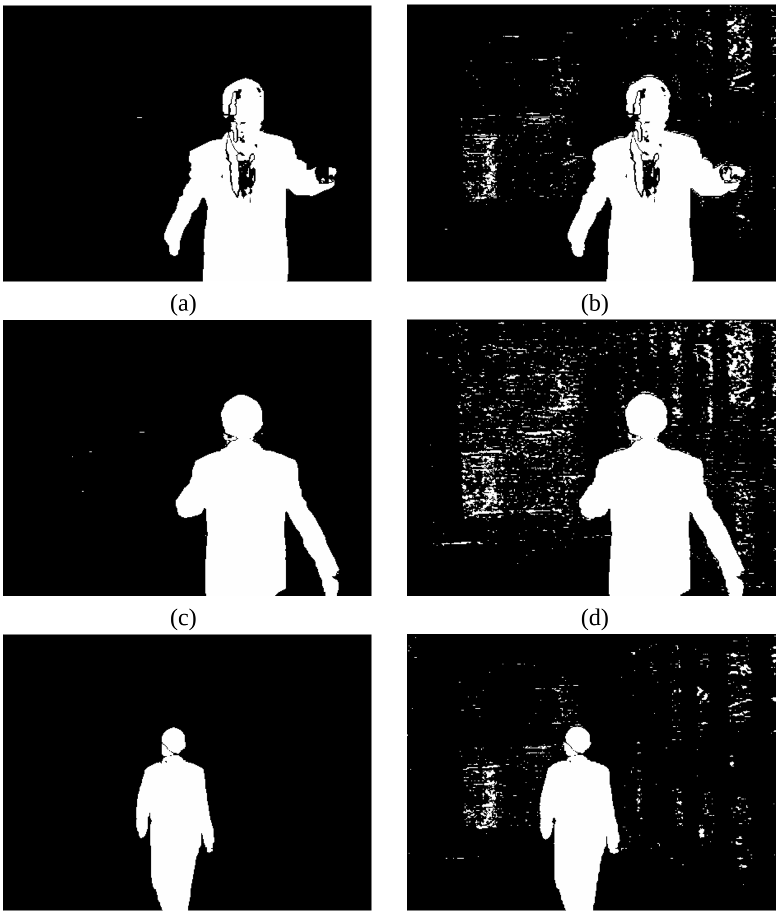
Fig. 6. Change detection results for the image frames shown in Fig. 5.

(a) (b)

Fig. 7. Selected image frames from experimental video 2.

(c) (d)

Fig. 7. (Cont'd) Selected image frames from experimental video 2.



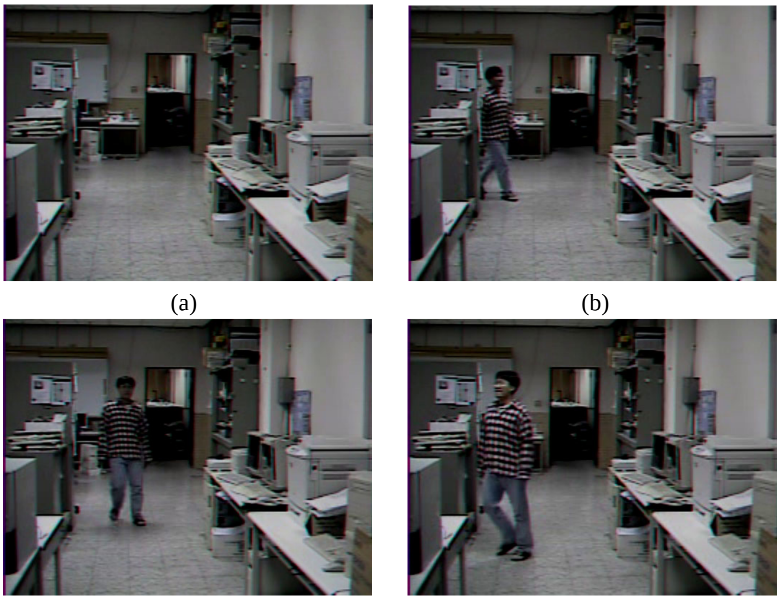
(a) (b)

(c) (d)

(e) (f)

Fig. 8. Change detection results for the image frames shown Fig. 7

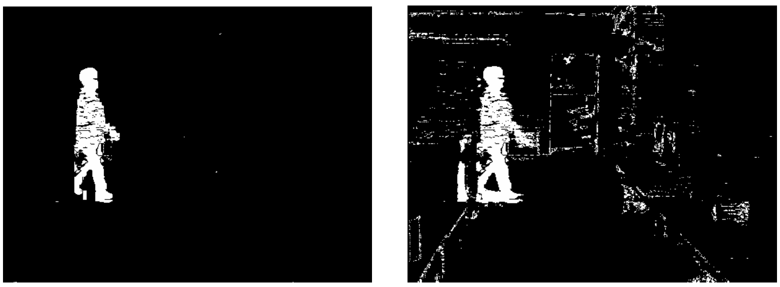
Fig. 9 shows the selected image frames from the third experimental video. The scene in the third experimental video lacked sufficient lighting. Due to the poor illumination, noises were induced into the video, affecting the correctness of the change detection results. Comparing the change detection results shown in Figs. 10 (a), (c) and (e) with the results shown in Figs. 10 (b), (d) and (f), which were obtained using the proposed method and the method described in [15] respectively, we can observe that the misclassified pixels caused by noises in Figs. 10 (b), (d) and (f) are eliminated in Figs. 10 (a), (c) and (e). Moreover, the proposed scheme was able to reduce the effect of the shadows by analyzing the dynamic cumulative histogram on the pixel level.



(a) (b)

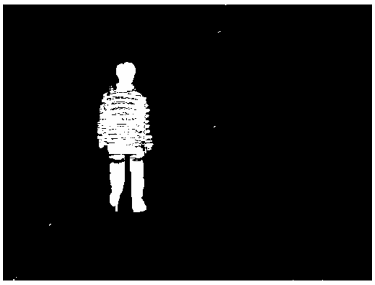
(c) (d)

Fig. 9. Selected image frames from experimental video 3.

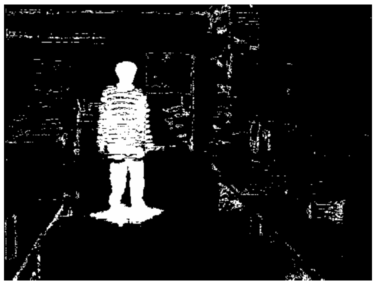


(a) (b)

Fig. 10. Change detection results for the image frames shown in Fig. 5.



(c)



(d)

(e)

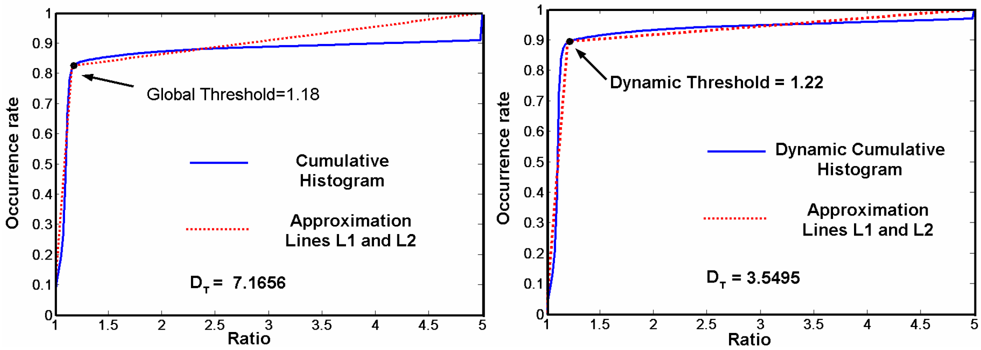
(f)

Fig. 10. (Cont'd) Change detection results for the image frames shown in Fig. 5.

From the experimental results displayed in Figs. 5-10, we can see that global thresholding can not reflect the local property of a pixel, such as its illumination. A systematic local thresholding method based on a computed global threshold can be used to reduce the number of mis-binarized pixels. Since the method described in [15] utilizes only global threshold, the algorithm proposed in this paper can clearly outperform it. Fig. 11 (a) depicts the cumulative histogram and the global threshold of the image shown in Fig. 7 (b). Fig. 11 (b) shows the dynamic cumulative histogram and the pixel-level dynamic threshold for a pixel (5, 125) of the image shown in Fig. 7 (b). We can see that the dynamic cumulative histogram in Fig. 11 (b) achieves a better approximation. More specifically, the distance D_T between the approximation lines and the dynamic cumulative histogram, which is what we want to minimize, is smaller in Fig. 11 (b) than in Fig. 11 (a). Therefore, the statistics from this experiment also confirm the superiority of the dynamic cumulative histogram. As for the time complexity, the computation time needed in [15] was equal to the time needed to compute a global threshold. It is proportional to the number of bins in the histogram and does not depend on the size of the image frame. For the proposed algorithm, the time complexity is proportional to the size of the image frame.

4. CONCLUSIONS

In this paper, a new thresholding method based on pixel-level dynamic thresholds has been proposed. In addition to finding a global threshold for every image frame, we adjust the cumulative histogram of a target image dynamically and use two line segments



(a) Cumulative histogram. (b) Dynamic cumulative histogram.

Fig. 11. Comparison between a cumulative histogram and a dynamic cumulative histogram.

to approximate the dynamic cumulative histogram in order to obtain a dynamic threshold for every pixel. Both pixel-level dynamic thresholds and the global threshold are considered when performing binarization. We preserve the simplicity of utilizing cumulative histograms while including more local information. The proposed method can select the most suitable threshold for every pixel in the target image and substantially reduce the number of detection errors. The experimental results confirm the superiority of the proposed method, showing that it is able to work well under various illumination conditions, including especially bright and dark scenes.

REFERENCES

1. D. R. Magee, “Tracking multiple vehicles using foreground, background and motion models,” *Image and Vision Computing*, Vol. 22, 2004, pp. 143-155.
2. C. Kim and J. N. Hwang, “A fast and robust moving object segmentation in video sequences,” in *Proceedings of the IEEE International Conference on Image Processing*, Vol. 2, 1999, pp. 131-134.
3. S. C. Liu, C. W. Fu, and S. Chang, “Statistical change detection with moments under time-varying illumination,” *IEEE Transactions on Image Processing*, Vol. 7, 1998, pp. 1258-1268.
4. J. H. Lee and K. W. Lim, “A fast multi-resolution block matching algorithm and its LSI architecture for low bit-rate video coding,” *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11, 2001, pp. 1289-1301.
5. S. Zhu and K. K. Ma, “A new diamond search algorithm for fast block-matching motion estimation,” *IEEE Transactions on Image Processing*, Vol. 9, 2000, pp. 287-290.
6. I. Haritaoglu, D. Harwood, and L. Davis, “W4: who? when? where? a real time system for detecting and tracking people,” in *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, 1998, pp. 222-227.
7. C. Ridder, O. Munkelt, and H. Kirchner, “Adaptive background estimation and forground detection using kalman-filtering,” in *Proceedings of the Recent Advances in Mechatronics*, 1995, pp. 193-199.

8. C. Wren, A. Azarbyejani, T. Darrell, and A. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, 1997, pp. 780-785.
9. S. McKenna, S. Jabri, Z. Duric, and H. Wechsler, "Tracking interacting people," in *Proceedings of the IEEE International Conference on Automatic Face and Recognition*, 2000, pp. 348-353.
10. D. R. Magee, "Tracking multiple vehicles using foreground, background and motion models," *Image and Vision Computing*, Vol. 22, 2004, pp. 143-155.
11. M. Fettke, K. Sammut, M. Naylor, and H. Fangpo, "Evaluation of motion detection techniques for video surveillance," in *Information*, *Decision and Control*, 2002, pp. 247-252.
12. Q. Z. Wu and B. S. Jeng, "Background subtraction based on logarithmic intensities," *Pattern Recognition Letters*, Vol. 23, 2002, pp. 1529-1536.
13. M. I. Sezan, "A peak detection algorithm and its application to histogram-based image data reduction," *Graphical Models and Image Processing*, Vol. 29, 1985, pp. 47-59.
14. D. M. Tsai, "A fast thresholding selection procedure for multimodal and unimodal histograms," *Pattern Recognition Letters*, Vol. 16, 1995, pp. 653-666.
15. Q. Z. Wu, H. Y. Cheng, and B. S. Jeng, "Motion detection via change-point detection for cumulative histograms of ratio images," *Pattern Recognition Letters*, Vol. 26, 2005, pp. 555-563.

Hsu-Yung Cheng (鄭旭詠) received her B.S. degree in Computer Science and Information Engineering from National Chiao Tung University in 2000 and her M.S. degree from the same department in 2002. She is now a Ph.D. candidate in the Institute of Computer Science and Information Engineering, National Central University. She was elected as a member of the Phi Tau Phi Scholastic Honor Society of the R.O.C. in 2002 and honored with the MediaTek Foundation Fellowship in 2004. Her current research interests include pattern recognition, image processing, and video analysis.

Quen-Zong Wu (吳坤榮) received his B.S. degree from the Department of Electrical Engineering, Tatung Institute of Technology, Taiwan, in 1987, his M.S. degree in Computer Science and Information Engineering from National Chiao Tung University, Taiwan, in 1989, and his Ph.D. degree from the Department of Computer Science and Information Engineering, National Chiao Tung University, Taiwan, in 1997. Presently, Dr. Wu is with Chunghwa Telecommunication Laboratories, Taiwan. His current research interests include pattern recognition, image processing, and neural networks.

Kuo-Chin Fan (范國清) was born in Hsinchu, Taiwan, on June 21, 1959. He received the B.S. degree in Electrical Engineering from National Tsing Hua University, Hsinchu, Taiwan, in 1981. He received the M.S. and Ph.D. degrees from the University of Florida in 1985 and 1989, respectively. In 1983, he joined the Electronic Research and Service Organization (ERSO), Taiwan, as a computer engineer. From 1984 to 1989, he was a Research Assistant with the Center for Information Research, the University of Florida. In 1989, he joined the Institute of Computer Science and Information Engineering, National Central University, Chungli, Taiwan, where he became a Professor in 1994. From 1994 to 1997, he was chairman of the department. Currently, he is the Director of the Computer Center. His current research interests include image analysis, optical character recognition, and document analysis. Prof. Fan is a member of SPIE.

Bor-Shenn Jeng (鄭伯順) received the B.S. degree from the Department of Physics, National Taiwan Normal University, Taipei, Taiwan in 1969, and the M.S. degree from the Department of Geophysics and the Ph.D. degree from the Institute of Optical Sciences, National Central University (NCU), Taiwan, in 1973 and 1991, respectively. He has been with the Research Institute of Chunghwa Telecom (formerly Telecommunication Laboratories) in TaoYuan, Taiwan, R.O.C., since 1974. Currently, he is the vice president of the Research Institute of Chunghwa Telecom, where he has been working on multimedia, intelligent human/ machine interfaces, Chinese character recognition, generation and compression, signal processing, intelligent transportation systems, etc. From 1991 to 1997, he was an Associate Professor in the Institute of Optical Sciences, NCU, Taiwan. He has been a Professor in the Information Engineering Department of NCU, Taiwan, since 1997. He has published over 100 technical papers and has over 100 patents in the areas of image processing, intelligent transportation systems, pattern recognition, etc. His significant achievements include: leadership in the research and development of the first electronic-toll- collection system in Taiwan, contribution to the design and deployment of the first government certificate authority in Taiwan, contribution to the design and deployment of a text-to-speech system for the blind in Taiwan, leadership in the design of the first ADSL- based multimedia on demand system in Taiwan, contribution to the design of the first commercial ADSL system in Taiwan, etc. He received the Distinguished Performance in Information Sciences Award, conferred by the National Science Council of Executive Yuan of Taiwan, R.O.C., in 1988, and the Distinguished Performance in Technology Award, conferred by the Executive Yuan, in 1989, which is the highest technology award conferred by the government of the R.O.C. He is a senior member of the IEEE.