Motion Detection Based on Improved Two-Piece Linear Approximation for Cumulative Histogram of Ratio Images

Kuo-Feng Hwang and Chih-Ching Chen
Department of Multimedia Design
National Taichung Institute of Technology
kfhwang@ntit.edu.tw, s18944201@ntit.edu.tw

摘要

本篇論文提出利用比率影像(ratio image)來進行在影像序列(image sequence)中的物體移動偵測。由於比率影像比傳統的影像減法較能排除因照度所產生之雜訊 (illumination-tolerance)，因此利用它來偵測影像中的移動物體將具有較高之準確性。除此之外，如何運用累計直方圖(cumulative histogram)之遞增特性來決定適當的門檻值，然後利用此門檻值來偵測影像序列中之移動物體是本論文的研究重點。本論文所提出的方法相較於傳統以影像序列相減之後，再從每張影像的直方圖決定門檻值的方法，具有更高的執行效率，而且從實驗結果來看，偵測結果亦維持有一定之水準。

關鍵詞：比率影像、移動偵測、影像序列、累計直方圖。

Abstract

In this paper, we use ratio images as the basis for motion detection. The ratio image is obtained from dividing two sequential images. After that the cumulative histogram of the ratio image can be calculated. We proposed a method to determine a suit threshold from the cumulative histogram. The proposed method has good performance to detect motion object in aspect of the time complexity. Comparing to the traditional background subtraction approach, the proposed method has the capability to resistant the illumination-tolerance in sequential images. The experimental results show that the proposed method is practical in motion detection especially for the application of surveillance.

Keywords: Ratio image, motion detection, image sequence, cumulative histogram.

1. Introduction

Detecting a moving object in sequential images is a very import research topic since it can be used in many applications. Although it has been studied for several decades and considerable methods have been proposed to deal with the motion detection, it is still a growing interest of researchers and needs to be robust enough when applied in various unconstrained environments. The rapid in use of digital video cameras for surveillance, traffic system, robot localization and human computer interaction (HCI) have been significant development in the computer vision system. Thanks to the motion detection, surveillance system can activate image saving or alarm signalling whenever unusual motion object is detected [1]. Traffic system can use motion detection to measure passage time of flow control [2-4]. HCI can use for gesture tracking and look after the patient [5].

A changed scene may be observed via sequence of images that may record minutes or hours. Unfortunately the task of distinguishing abnormal events from everyday activity requires the continued vigilance of operator. It is tedious and time consuming large volumes of video data. Automated motion detection through a sequence of images has remained one of the difficult problems in computer vision. Nowadays, a lot of efficient and effective motion detection algorithms had been brings out in this task but motion detection is still primitive [6]. Some algorithms perform well in restricted environments, such as tracking fixed shape, but a general solution is not currently available. A frequent problem is that when an algorithm is refined for one application, it becomes unsuitable for other applications. Therefore, a relevance of algorithm for diverse applications is necessary.

2. Literature Review

Motion detection is also called “change detection” in image processing domain [7]. There are three main approaches in motion detection: Optic flow, temporal difference and background subtraction [8]. Optic flow is also called “image flow” [9]. It can achieve success of motion detection in the presence of camera motion and separate independent motion simultaneously. In some cases, optic flow can lead to useful computation of simple discrete case, but there are too many assumptions that do not hold tight in the realistic situation. Temporal difference is simple method that uses temporal differencing for detecting moving objects in a static environment. But the object shape using this method is usually incomplete to provide the follow-up detecting [10]. Background subtraction is a traditional technique for finding moving objects in image sequences. This approach provides more feature data when compared with other motion detection approach. Although background subtraction has a lot of advantages to detect moving objects. But it is based on standard intensity, which is multiplication of illumination and reflectance. Let $F(x, y)$ denotes the standard intensity, $i(x, y)$ denotes the illumination and $r(x, y)$ denotes the reflectance ratio. Standard intensity form as $F(x, y) = i(x, y) \times r(x, y)$. Any change in illumination will affect intensity; even uniform illumination changes for all pixels will cause different intensity changes for every pixel [11]. In order to reduce the affect caused by intensity changing, we use ratio images instead of difference images in the proposed motion detection technique.

2.1 Difference image vs. Ratio image

Ratio image is a simple method in the field of image processing [4]. This method is use division that is dissimilar to difference image use subtraction. It is also a useful method to check whether any objects in the frame of camera are in motion [12]. An image captured by camera is represented as a function of $I_i(x, y)$. If $I_i(x, y)$ and $I_{i+1}(x, y)$ denote the value scalars of the pixel at coordinates (x,y) for image frame i and frame $i+1$, then the ratio image is represented as

$$I_r(\mathrm{x,y}) = \frac{\max(I_i(x,y), I_{i+1}(x,y))}{\min(I_i(x,y), I_{i+1}(x,y))} \qquad (1)$$

where $I_r(x, y)$ namely ratio image.

If the pixel $I_r(x, y)$ values is one, it means that the corresponding pixels in $I_i(x, y)$ and $I_{i+1}(x, y)$ are very similar. On the opposed case, the pixel probably belongs to the static background and motion objects respectively. The principle and its example are clearly shown in Fig. 1. In Fig. 1, it shows what is different between difference image and ratio image. If the tracked object in the frame of camera is in motion, the second image will be divided into one and non-one regions. In the aspect of finding out the tracked object, because the one's regions can be considered a background, the following operation is just concentrated and implemented in the non-one regions. It is also avoided that some lumps of pixels, which is as similar in gray as the tracked object, is segmented out from the static background. If one object in motion is as similar in gray as the background, the tracking result will be incorrect, so it is not suitable for use in this situation. It is considered over the usable range.

2.2 Histogram vs. Cumulative Histogram

Histogram manipulation can be used effectively for image processing. It is the basis for numerous motion detection techniques. A histogram is a graphical display pixel value and count of tabulated frequencies. In a digital image with gray levels in the range $[0, L-1]$ is a discrete function $H(i)=n_i$, where i is the ith gray level and n_i is the number of pixels. Fig. 2(a) is a background image which captured from surveillance camera. Fig. 2(b) is a histogram in Fig.2(a). People can obtain image characteristics that dark, light, low contrast and high contrast from Fig. 2(b).

A cumulative histogram is a mapping that counts the cumulative number of observations in all of the bins up to the specified bin. That is, the cumulative histogram $CH(i)$ of a histogram $H(i)$ is defined as:

$$CH(i) = \sum_{i=0}^{L-1} H(i) \qquad (2)$$

The cumulative histogram is a variation of the histogram in which the vertical axis gives not just the counts for a single bin, but rather gives the counts for that bin plus all bins for smaller values of the response variable [13]. Fig. 2(c) shows the cumulative histogram is an increasing or monotonic increasing function. That is, a cumulative histogram does not fluctuate as histogram.

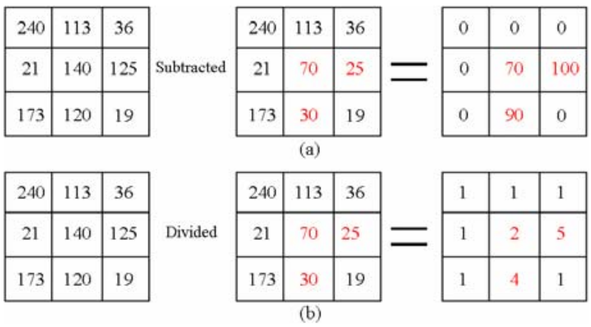

Fig. 1 A difference image and ratio image sample: (a) difference image; (b) ratio image.

2.3 Automatic threshold selection: Otsu's method

A lot of different methods have been proposed for automatic threshold selection. Most of automatic threshold selection algorithms will have to look for significant peaks and valleys in the image then separate them, e.g. Otsu [14] and Kapur [15]. In this section, we will briefly review the Otsu's method for automatic threshold selection.

The Otsu's method is a nonparametric and unsupervised method of automatic threshold selection for image segmentation. It selects the threshold based on the minimization of the within-class variances and maximization of the between-class variance. Let $I_d(x, y)$ denote a gray-level difference image.

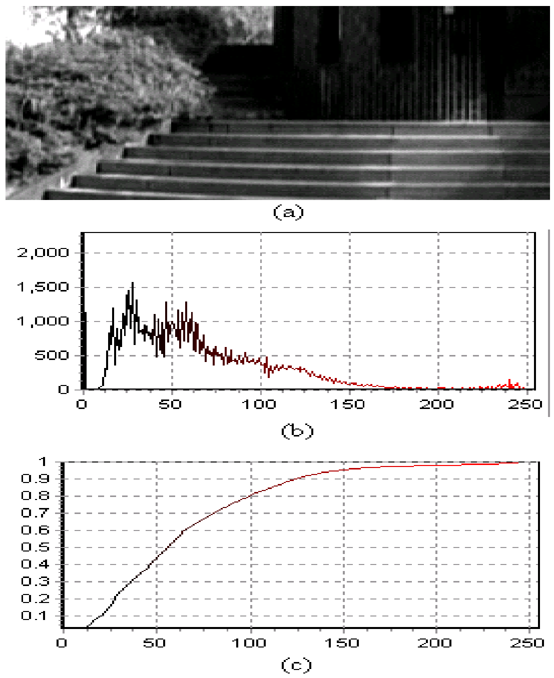

Fig. 2 (a) A background image which captured from surveillance camera. (b) Histogram. (c) Cumulative histogram.

The value of $I_d(x, y)$ is range $[0, L\text{-}1]$, where L is number of gray-level bins. The number of pixels at level i be n_i and the total number of pixels by $n = n_0 + n_1 + \ldots + n_{L-1}$. In order to simplify the discussion, we will specify the gray-level histogram is normalized as a probability distribution:

$$p_i = n_i \;/ n \tag{3}$$

The average gray-level of the entire image μ_T is computed as:

$$\mu_T = \sum_{i=0}^{L-1} i p_i \tag{4}$$

Support the image was partitioned into two classes $C_1 = \{0, 1, \ldots, t\}$ and $C_2 = \{t+1, t+2, \ldots, L\text{-}1\}$. C_1 denotes background and C_2 denotes motion objects. t is the threshold value. Then the probabilities of two classes are given by:

$$\omega_1(t) = \sum_{i=0}^{t} p_i \tag{5}$$

and

$$\omega_2(t) = \sum_{i=t+1}^{L-1} p_i = 1\text{-}\omega_1(t) \tag{6}$$

The two classes mean are:

$$\mu_1(t) = \sum_{i=0}^{t} i p_i / \omega_1(t) \tag{7}$$

and

$$\mu_2(t) = \sum_{i=t+1}^{L-1} i p_i / \omega_2(t) \tag{8}$$

The Otsu's method selects the best threshold T by maximizing the between-class variance:

$$\sigma_B^2(T) = \underset{0 \le t < L}{\arg\max} \{\sigma_B^2(t)\} \tag{9}$$

where the between-class variance σ_B^2 is defined as:

$$\sigma_B^2(t) = \omega_1(t)(\mu_1(t) - \mu_T)^2 + \omega_2(t)(\mu_2(t) - \mu_T)^2 \tag{10}$$

3. Proposed method

Our proposed method is based on ratio image. The cumulative histogram $CH(i)$ is monotonic increasing function, where $1 \le i \le N$. N is the number of bins of the cumulative histogram. The illustration is shown in Fig. 3. In function $CH(i)$, any sample point t in $2 \le t \le N\text{-}1$ can calculate $CH(t)$ that is equal to Lyi. Meanwhile, Lxi equals N - t. Let the optimal threshold T equals t as Eq. (11) shows. Where T gets the maximum length of $Lyi + Lxi$.

$$T = \underset{t=1}{\overset{N-1}{\arg\max}} \; Lyi + Lxi \tag{11}$$

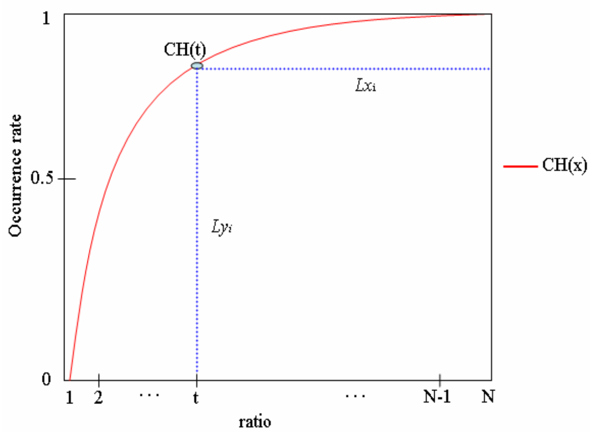

Fig. 3 The illustration of looking for optimal threshold *T*

The proposed procedure is illustrated in Fig. 4. We briefly describe the procedures as follows.

Step 1. Setup background image $I_b(x, y)$.

Step 2. Input current image $I_c(x, y)$.

Step 3. Produce ratio image $I_r(x, y)$.

$$I_r(x, y) = \frac{\max(I_b(x,y), I_c(x,y))}{\min(I_b(x,y), I_c(x,y))} \quad (12)$$

After this step the cumulative histogram has been established.

Step 4. To find a threshold *T* from ratio image $I_r(x, y)$ by Eq. (11).

Step 5. The optimal threshold *T* had been obtained from previous step. The motion object is extracted by *T* as follows.

$$I_r(x,y) = \begin{cases} 0 & \text{if } I_r(x,y) \le T \\ 1 & \text{if } I_r(x,y) > T \end{cases} \quad (13)$$

Where $I_r(x, y) = 1$ denotes the pixel at coordinates (x, y) is a motion pixel. In other case, $I_r(x, y)$ is a background pixel.

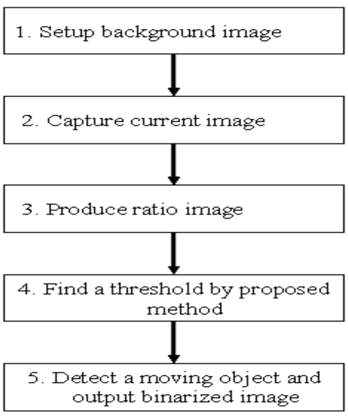

Fig. 4 The Proposed procedures.

4. Experimental Results

In order to demonstrate the efficient of our proposed method, experimental image sequences with different situations were tested. Furthermore, judge a qualitative of algorithm's result by manually. We tested the proposed method and Otsu's method over than 40 cases. Selected image frames from three experimental sequences will be discussed in this section. In the first experiment, we used an image sequence under normal illumination conditions in an indoor scene. In the second experiment, the image sequence had been captured in outdoor scene. In the third, a darker scene had been tested.

Fig. 5 shows the images of 352x240 pixels used for the experiments at normal illumination conditions of surveillance in the bookstore. Fig. 5(a) is the background image. Fig. 5(b) and Fig. 5(c) are two image frames taken from a sequence showing a girl walking into the bookstore. The motion detection results obtained with the proposed method are displayed in Fig. 6(a) and Fig. 6(b). The results obtained via Otsu's method are displayed in Fig. 6(c) and Fig. 6(d). The threshold of Fig. 6(a) is 4.5 and Fig. 6(b) is 6.2. The threshold of Fig. 6(c) is 53 and Fig. 6(d) is 55. Note that the proposed method obtained motion objects from the ratio image; on the other hand, Otsu's method is based on difference image to obtain motion objects. Compared with the results shown in Fig. 6(c) and Fig. 6(d), those shown in Fig. 6(a) and Fig. 6(b) are obviously more accurate of moving object's position but features are not very clear. Although Otsu's method is quite complete of moving object's features. But there are some misclassification errors that we can observe Fig. 6(c) and Fig. 6(d) obviously influence by illumination near the back and knees.

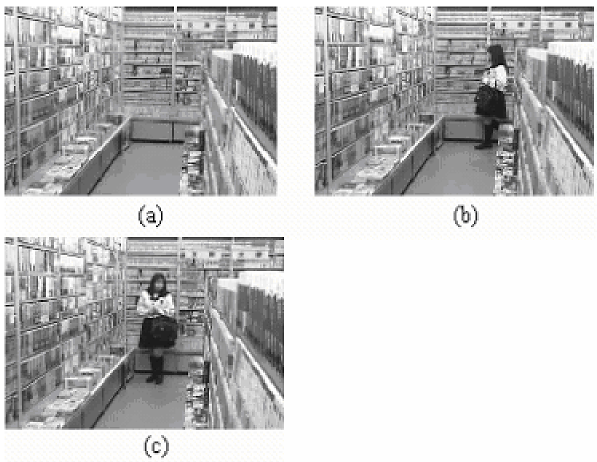

Fig. 5 Normal illumination conditions in the indoor scene: (a) the background image; (b) a girl entered the aisle in bookstore; (c) the girl turned around.

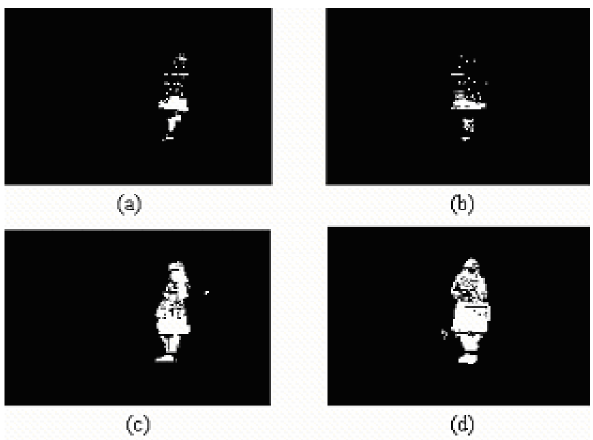

Fig. 6 Motion detection results for the image frames shown in Fig. 5.

Fig. 7 shows the selected image frames from home surveillance in the outdoor. Fig 7(a) is the background image. Fig. 7(b) and Fig. 7(c) are two image frames taken from a sequence showing a woman out of house. The motion detection results obtained with proposed method are displayed in Fig. 8(a) and Fig. 8(b) that threshold are 4.8 and 4.0. The results obtained via Otsu's method are displayed in Fig. 8(c) and (d) that threshold is 20 and 80 respectively. Fig. 8(a) has shown the moving object accurately. Fig. 8(c) has shown the moving object in a clutter. It is very obviously that proposed method is better than Otsu's method. Although Fig. 8(b) is very similar with Fig. 8(d), but we still might distinguish that Fig. 8(d) has some misclassification in the upper left region.

Fig. 9 shows a dark scene in home surveillance to monitor visitor. Fig 9(a) is the background image. Fig. 9(b) is showing a lady had dropped into the house. Fig. 9(c) is showing the lady went to stairs. In this case, surveillance video lacked sufficient lighting. Due to the poor illumination, noises were induced into the video, affecting the correctness of the motion detection results. Many algorithms cannot work well in this situation but our proposed method and Otsu's method can do it. The motion detection results obtained with proposed method are displayed in Fig. 10(a) and Fig. 10(b) that threshold is 2.5 and 2.2. The results obtained via Otsu's method are displayed in Fig. 10(c) and Fig. 10(d) that threshold is 17 and 6 respectively. In Fig. 10(a) and Fig. 10(b), we can see the moving object position had been detection accurate by proposed method. In Fig. 10(c), moving object's had more complete detection by Otsu's method. But in Fig. 10(c), it had been influenced by illumination change on the upper right region. A lot of non-motion pixels have been detection in Fig. 10(d) by Otsu's method. In this case, we can know that Otsu's method does not work well at the illumination change, but our method can eliminate the noise by illumination-tolerance.

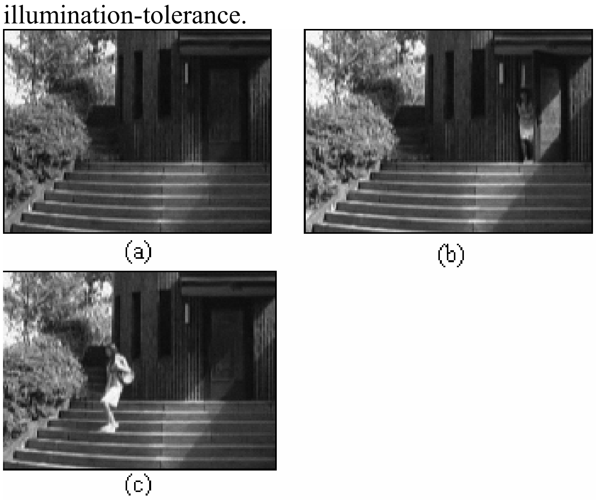

Fig. 7 Motion detection in the outdoor scene: (a) the background image; (b) a lady opened the door; (c) the lady left home.

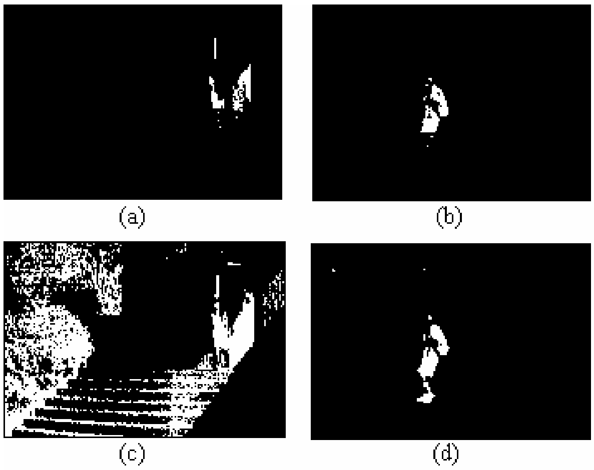

Fig. 8 Motion detection results for the image frames shown in Fig. 7.

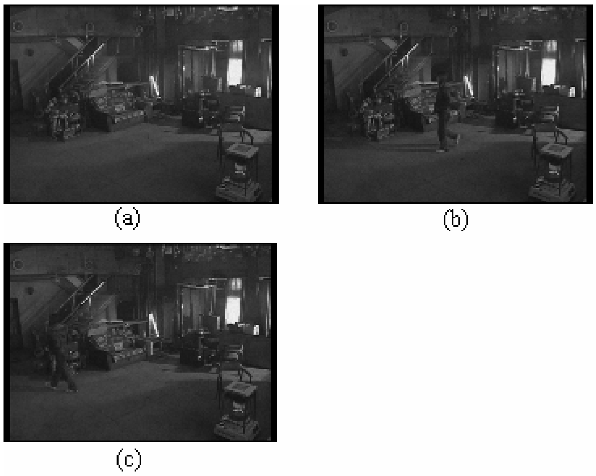

Fig. 9 Motion detection in dark scene: (a) the background image; (b) a lady dropped into the house; (c) the lady went to stairs.

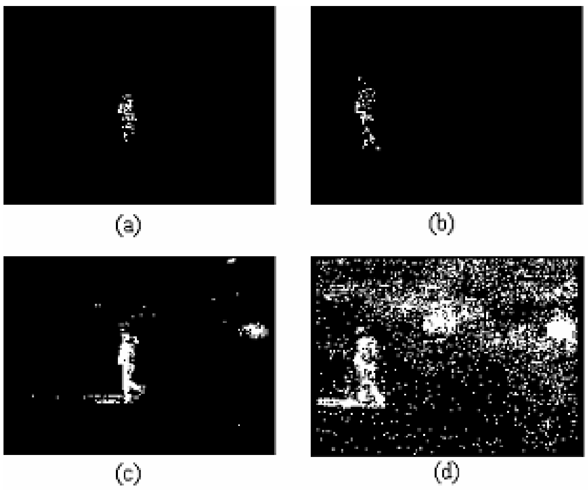

Fig. 10 Motion detection results for the image frames shown in Fig. 9.

5. Conclusions and Future Works

In this paper, we have proposed an efficient approach to facilitate the motion detection by using ratio images and cumulative histogram. This proposed has improved Wu et al. [4] specially in small object. The object/background is smaller than 30%. Experimental results shown that the proposed method has the capability to resist the illumination noises and high performance. Furthermore, it is useful in varied scenes including the outdoor, indoor even in dark environment. In our future works, we will focus on developing a more robust algorithm to get better results and more complete features in many different environments.

References

[1] L. M. Fuentes and S. A. Velastin, "People Tracking in Surveillance Applications," Image and Vision Computing, vol. 24, pp. 1165-1171, 2006.

[2] J. Aguilera, H. Wildenauer, M. Kampel, M. Borg, D. Thirde, and J. Ferryman, "Evaluation of Motion Segmentation Quality for Aircraft Activity Surveillance," IEEE International Workshop on VS-PETS, Beijing, 2005.

[3] T. Alexandropoulos, S. Boutas, V. Loumos, and E. Kayafas, "Real-time Change Detection for Surveillance in Public Transportation," IEEE Conference on Advanced Video and Signal Based Surveillance, 2005.

[4] Q. Z. Wu, H. Y. Cheng and K. C. Fan, "Motion Detection Based on Two-piece Linear Approximation for Cumulative Histograms of Ratio Images in Intelligent Transportation Systems," IEEE International Conference on Networking, Sensing & Control, Taipei, 2004.

[5] T. B. Moeslund, A. Hilton, and V. Kruger, "A Survey of Advances in Vision-based Human Motion Capture and Analysis," Computer Vision and Image Understanding, vol. 104, pp. 90-126, 2006.

[6] P. L. Rosin and E. Ioannidis, "Evaluation of Global Image Thresholding for Change Detection," Pattern Recognition Letters, vol. 24, pp. 2345-2356, 2003.

[7] H. Y. Cheng, Q. Z. Wu, K. C. Fan and B. S. Jeng, "Binarization Method Based on Pixel-level Dynamic Thresholds for Change Detection in Image Sequences," Journal of Information Science and Engineering, vol. 22, pp. 545-557, 2006.

[8] Y. Ren, C. S. Chua, and Y. K. Ho, "Statistical Background Modeling for Non-stationary Camera," Pattern Recognition Letters, vol. 24, pp. 183-196, 2003.

[9] L. G. Shapiro and G. C. Stockman, Computer Vision: Prentice-Hall, Inc., 2001.

[10] P. Spagnolo, T. D. Orazio, M. Leo, and A. Distante, "Moving Object Segmentation by Background Subtraction and Temporal analysis," Image and Vision Computing, vol. 24, pp. 411-423, 2006.

[11] R. C. Gonzalez and R. E. Woods, Digital Image Processing, second ed. New Jersey: Prentice-Hall, Inc., 2002.

[12] 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, pp. 555-563, 2005.

[13] J. T. McClave and T. Sincich, Statistics, 9 ed: Prentice-Hall, Inc., 2002.

[14] N. Otsu, "A Threshold Selection Method from Gray-level Histograms," IEEE Transations on Systems, Man and Cybernet. 9, pp. 62-66, 1979.

[15] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A New Method for Gray Level Picture Thresholding Using the Entropy of the Histogram," Computer Vision, Graphics and Image Processing, vol. 29, pp. 273-285, 1985.