A Non-repeated Face Extracting Scheme Using Object Location Comparison Technique for Differently Exposed video sequences

Wen-Yuan Chen, Yi-Wen Cheng, C. J. Hsu, and Chih-Ming Chang
Department of Electronic Engineering,
National Chin-Yi University of Technology, Taichung, Taiwan, R.O.C.
E-mail: cwy@ncut.edu.tw, ewen10189@hotmail.com, hsu@ncut.edu.tw, eerics@giga.net.tw

Abstract

In this paper, we based on the image processing and exclusion bitmapping techniques develop an extracting object scheme for differently exposed video sequence. In our approach, an object location comparison method is used to examine the object if repeated. We discard the repeated objects and recoding the new objects. Compare with other papers, there are three advantages in our scheme: (A) Using median filter, dilation, opening, closing, and small area discarding methods to separate objects on the image frame; (B) Using exclusion bitmap technique to correctly distinguish the objects of the image even under differently exposed; (C) Using object location comparison technique to extract only one copy human object.

In order to explore the utility and demonstrate the efficiency of the proposed scheme, several image sequences are conducted. The experimental results show our scheme can precisely extract one copy face on the differently exposed video sequence.
.

Keywords: median filter, dilation, opening, closing, face extracting, exclusion bitmapping

1. Introduction

The video recording monitor system is a kind of negative equipment; its only usability is to offer the evidence for searching after. The video recording monitor system is more powerful when it equips with some auto preventing functions. The video recording monitor systems often need to change tapes because they accumulate numerous information rapidly, which is not good for saving information. This paper process the dynamic comparison of the sequences of people face images, and save once of the same people to increase the capability of saving information. Concurrently, our scheme upgrade the functions of the video recording monitor system

In the computer vision, human face perception is currently a hot research area. Locating and tracking human face is very important for face perception. Identification and confirming are necessary in security system. In the last years, face and facial expression recognition have attracted much attention of engineers. Many commercial applications have been developed from these efforts. A pre-processing of face detection system is used to detect the locations of people faces in an image. However, face detection from an image is a challenging task owing to its variability in scale, location, orientation and pose.

Several researches were proposed for face detection. Yang and Huang [1] used a hierarchical knowledge-based method to detect faces. Three level resolutions were used in their approach. The lowest resolution image was used for searching face candidates. At the medium resolution image, local histogram equalization was performed on the face candidates and followed by edge detection. Finally, the surviving candidate regions were then examined at the highest resolution with a set of facial features. The advantage of this method is that a coarse-to-fine strategy can reduce the computation.

Leung et al. [2] based on local feature detector and random graph matching techniques propose a probabilistic method to locate a face in a scene. In his approach, five features (two eyes, two nostril, and nose/lip junction) were used to depict a typical face. Meanwhile, a facial template was defined. The relative distances of any pair of facial features of the same type were computed to test whether the testing object is a human face or not.

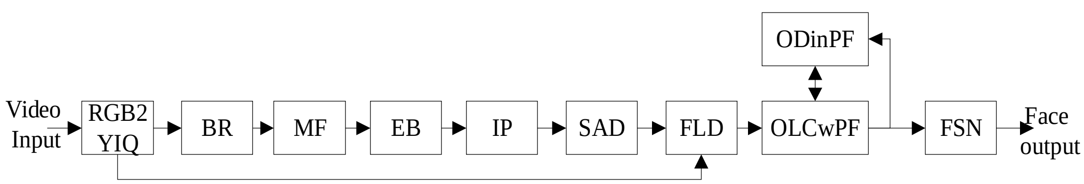
Yang and David [3] presents a survey about detecting faces in images. They provided a general and complete introduction of researchers and their researches in the face detection. Other papers about the face detection and tracking can be found in [4]-[13]

This paper gives two functional improvements for the video recording monitor system: (a) Adopt image processing techniques on object feature improve objects detection from the video recording monitor system. (b) Using exclusion bitmap technique to correctly distinguish the objects even under differently exposed images.

The main goal of our work is to develop a simple and effective objects extracting system on the differently exposed video sequence. The remainder of this paper is organized as follows: Section 2 describe object extracting algorithm and its detail process. The experiment results show in Section 3. Finally we conclude this paper in Section 4.

2. Object Extracting Algorithm

This paper presents a method to extract objects and record them at only one copy from a differently exposed video recording monitor system. Fig.1 shows the flow chart of our algorithm. In the beginning of our process, the colored input frame is transformed to YIQ domain. We took the Y-component as the processing image because the Y-component occupies 93% energy of the total energy in YIQ domain. A background removal method is used to get ride of the background and reserve the objects image. Next, an exclusive bitmapping technique used to convert the image from grayscale to binary correctly even on differently exposed environments. Successively, several processing step (include, dilation, opening, and closing) is used to filter out the noise of the objects and to improve the effect of candidate object detection. Since a human object must be a large area, so the small area of pixels can be discard. In addition, according to person structure; we can easily extract the human face location by scale method of object. Furthermore, we compare the face location with the previous frame to determine the object is brand new or repeated. After the objects are all examined, we refresh the previous frame and record the new objects. Finally, we normalize the object and run the next frame to continuous the algorithm. The details of the algorithm are described in the following.



RGB2YIQ : RGB to YIQ Mapping
BR : Background Remova in Y Plane
MF : Median Filter
EB : Exclusion Bitmapping
IP : Image Processing
SAD: Small Area Discarding
ODinPF:Object Data in Previous Frame
OLCwPF:Object Location Comparison with Previous Frame
FSN : Face Size Normalization
FLD : Face Location Detection

Fig. 1 The flow chart of the object extracting algorithm

2.1. Color mapping RGB to YIQ

The red, green and blue (RGB) are three dimensions of illumination spectrum. They are enough to compose any color adequately, although the spectrum of illumination is infinite dimensional. A common alternation to the RGB representation of an image is the YIQ representation. The YIQ representation of an image is the standard model in the television transmission. The YIQ representation of an image obtained from the RGB representation of an image is given by equation (1),

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ 0.596 & -0.274 & -0.322 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.956 & 0.621 \\ 1.000 & -0.272 & -0.647 \\ 1.000 & -1.106 & 1.703 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} \tag{2}$$

Where Y is the luminance or brightness which refers to color density, I is the hue which is the dominant such as orange, red or yellow, and Q is the saturation or depth which is the amount of white light mixed with a hue of color. The equation (2) is the inverse transformation of equation (1), to transfer the image in YIQ planes back into the RGB planes.

2.2 Background Removal

Fig. 2 shows the flow chart of image processing for obtaining a grayscale image of moving objects with background removed. The background averaging procedure is used to obtain a background image by averaging exceed 50 pictures of moving objects having same background. Background subtraction is a popular and effective method for detecting moving foreground objects in the scene.

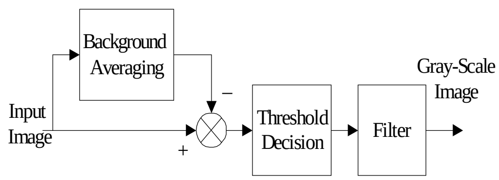


Fig. 2 The flow chart of background removal

2.3 Exclusion Bitmapping

In this paper, we use an exclusive bitmapping (EB) technique to convert the image from grayscale to binary even on differently exposed environments. In the EB process, we divided an image into several non-overlapped blocks, and take a medium value called m in a block firstly. Next, we select a range called r. And then we can correctly and easily to mapping the block image from grayscale into binary by using eq. (3).

$$f(x) = \begin{cases} 0 & if \quad m - r \le x \le m + r \\ 1 & \quad otherwise \end{cases} \tag{3}$$

where x is a pixel in the selected block image and the $f(x)$ is the corresponding pixel of the x after EB operation.

2.4 Image processing

Owing to the object image from gray scale mapping to the binary will causing some distortion, thus, several image processing techniques are used to remedy the drawback. In our scheme, we use the dilation, opening, and closing techniques to achieve the goal.

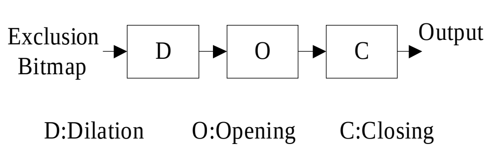


Fig. 3 The flow chart of the image pre-processing

2.5 Small Area Discard and Face Location Detection

To distinguish an object in the image is the core work of the algorithm. Fig. 4 shows the flow chart of the object detection in this paper. At the beginning, the median filter is used to filter out the noise of the image. Next, an EB technique can correctly map the image from grayscale into binary. After EB operation, a dilation operation, an opening operation, and a closing operation are used to connect the two parts of the image that separated causing by noise. Finally, a small area discard step is used to throw those objects which are belonging to noise images because their areas are so small. In our scheme, we use the value 3800 as a criterion to discard those noise object images. Fig. 4 shows the images of the object detection processing, where Fig. 4(a) denotes the image after median filter, Fig. 4(b) is the image after EB, Fig. 4(c) is the image after closing operation, and Fig. 4(d) shows the image after small area discard. From the Fig. 4, we see the objects are detected clearly and correctly.

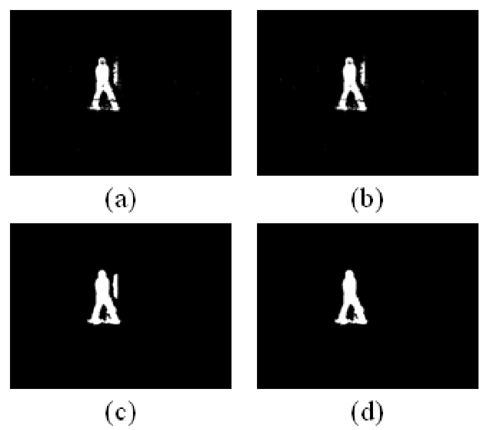


Fig. 4 The images under object detection processing, (a) the image after median filter, (b) the image after EB, (c) the image after closing operation, (d) the image after small area discard.

2.6 Object Location Comparison with Previous Frame

In order to obtain only one copy of the face image from several same objects on a video sequence, an object location comparison with previous frame (OLCwPF) step is needed. In Fig. 5 shows the schematic of the OLCwPF. Fig. 5(a) is the detected previous object frames, there are three objects, denote B, C, and D. And Fig. 5(b) is the current frame, there are four objects, denote A, B, C, and D. At the beginning of OLCwPF, the object A is used to compare with previous frame. When object comparing, a search range is defined, shown as the dash rectangular in

Fig. 5. In the other words, the distance dx and dy around the center of the object A is the search range. In our approach, if the object is appearing in the search range of the previous frame then the object is repeated, will be discard. By the rule, we check all the objects in the current frame, and obtained only one copy of face objects on the video sequence.

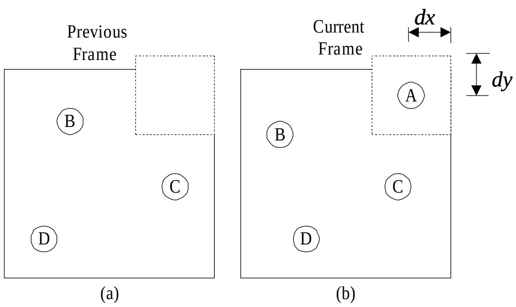


Fig. 5 The schematic of the object location comparison with previous frame method

2.7 Processing Steps

Fig. 6 shows a test sequence to express our approach processing. The Fig. 6(a) – Fig. 6(p) is the order of the frames on the video sequence. The number on the frame denotes the objects that are appear in the sequence. The processing steps of the approach are shown in Fig. 7. Fig. 7(a) denotes the one of frames on the video sequence in Y component; Fig. 7(b) express the background image; Fig. 7(c) shows the result images after background removed; Fig. 7(d)-Fig. 7(h) express the image after median filter, EB mapping, image dilation, image opening, and image closing operation, respectively. Finally, Fig. 7(i) shows the result of image pre-processing after discard small area pixels operation. The extracted face image from the test sequence in the Fig. 6 is shown in Fig. 8. From the Fig. 6, there are four peoples appear on the sequence, and processing result of our scheme is also four peoples as the Fig. 8 shown. This demonstrate our scheme is an effective non-repeated object extracting method. Beside, our scheme normalizes the extracted face so that the face is all have the same size.

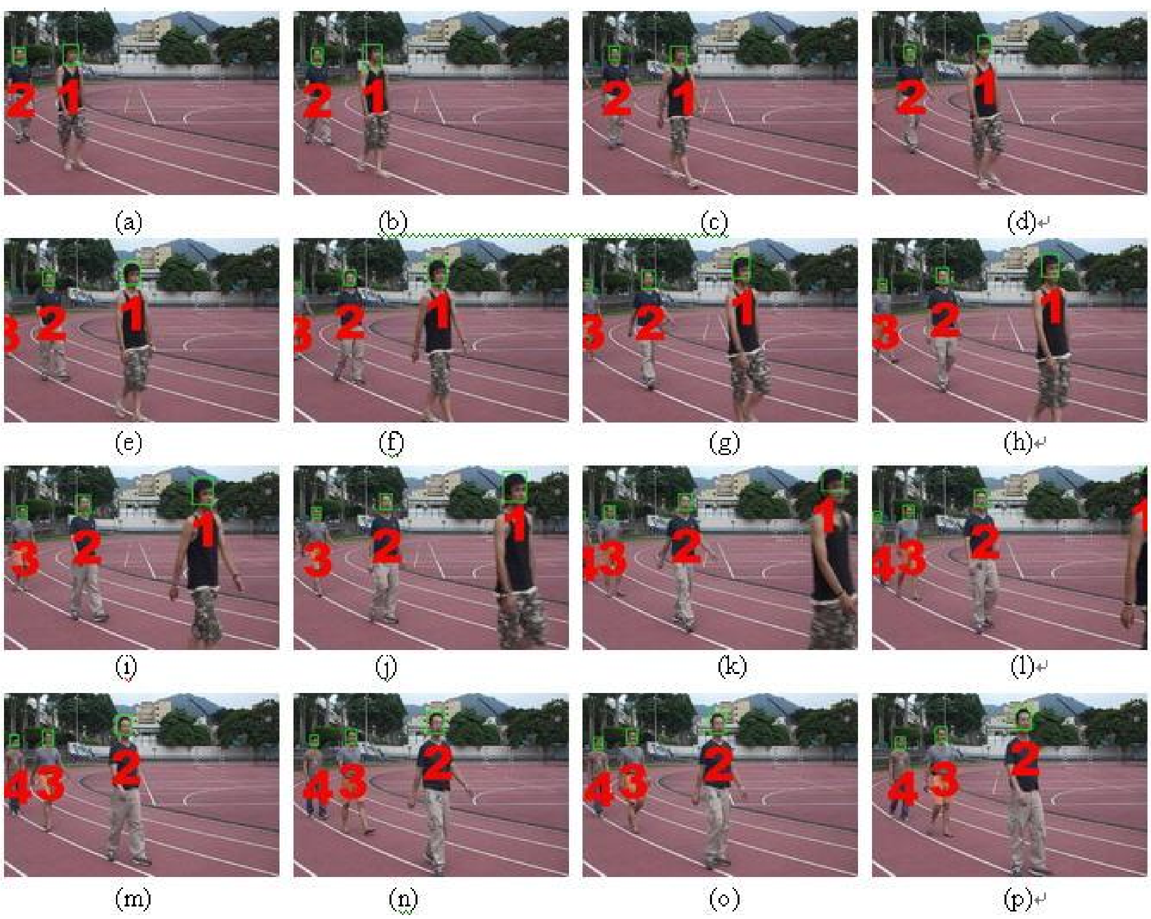


Fig. 6 The test sequence-1; (a)-(p) express the order on the frames on the video sequence. The number of the frame denotes the objects that are appear in the sequence.

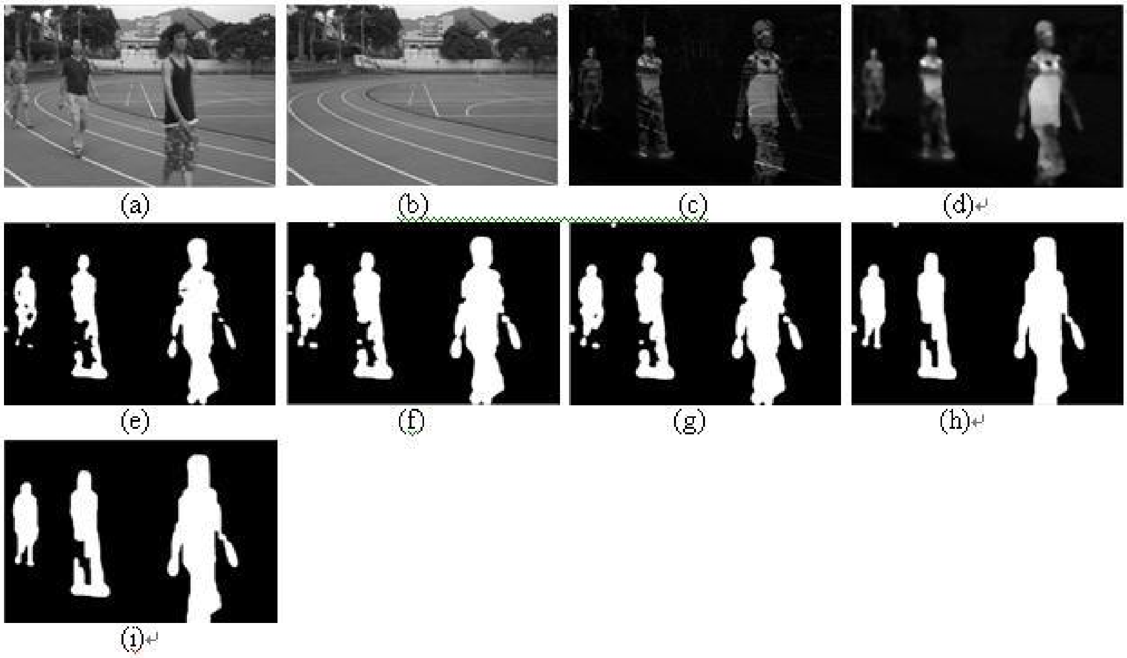


Fig. 7 The test image processing procedure; (a) the one of frames on the video sequence in Y component, (b) the background image, (c) the result image of (a)-(b), (d) after median filter of (c), (e) after EB of (d), (f) after dilation operation of (e), (g) after opening operation of (f), (h) after closing operation of (g), (i) after discard small area pixel of (h).

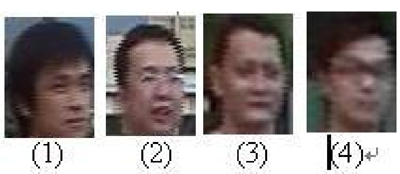


Fig. 8 The extracted face images from the test sequence in Fig. 6.

3. Empirical Results

Several video sequence of the test image with frame size (640×480) are used in simulation for demonstrating the performance of the proposed scheme. In our work, we use block size 3×3 for median filter operation. We use the value 3800 as a criterion to discard those small object images of the candidate face area. Finally, we set the extracted face image size is 64×48. Owing to paper size confined, we only show three cases from several test examples. Fig. 9 shows the test video sequence, it is an outdoor bright luminance environment, have four persons walk through the scene taken from the recording monitor system. Fig. 9(a) ~ 9(p) express the order of the frames on the video sequence, where the number of the frames denote the objects that is appear in the sequence. Fig. 10 shows the test image processing procedure corresponding to Fig. 9; Fig. 10(a) denotes the one of frames on the video sequence in Y component, Fig. 10(b) is the background image, Fig. 10(c) shows the result image after background removed, Fig. 10(d)-Fig. 10(h) express the image after median filter, EB mapping, image dilation, image opening, and image closing operation, respectively. Finally, Fig. 10(i) shows the results of the image pre-processing after discard small area pixel operation. The extracted face images from the test sequence-2 in Fig. 9 are shown in Fig. 11. Fig. 12 – Fig. 14 are another test cases.

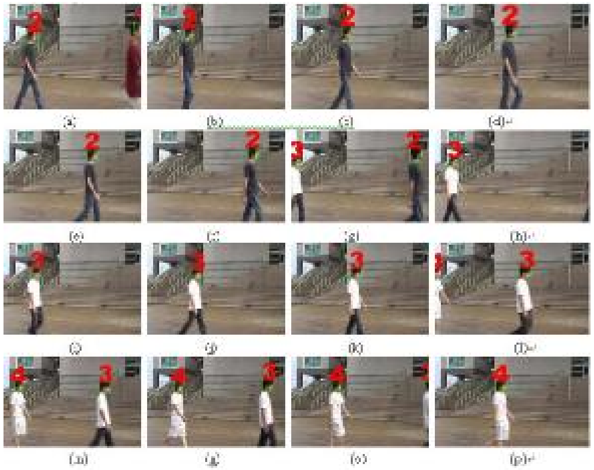


Fig. 9 The test sequence-2; (a)-(p) express the order on the frames on the video sequence. The number of the frame denotes the objects that are appear in the sequence.

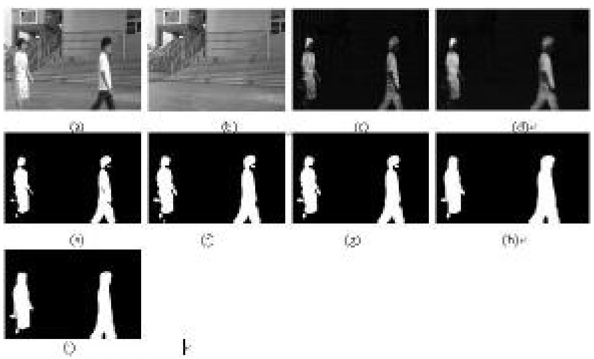


Fig. 10 The test image processing procedure correspond Fig. 9; (a) the one of frames on the video sequence in Y component, (b) the background image, (c) the result image of (a)-(b), (d) after median filter of (c), (e) after EB of (d), (f) after dilation operation of (e), (g) after opening operation of (f), (h) after closing operation of (g), (i) after discard small area pixel of (h).

Fig. 11 The extracted face images from the test sequence-2 in Fig. 9.

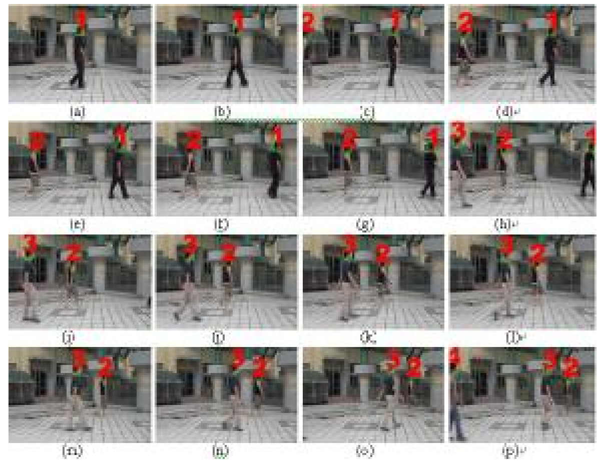


Fig. 12 The test sequence-3; (a)-(p) express the order on the frames on the video sequence. The number of the frame denotes the objects that are appear in the sequence.

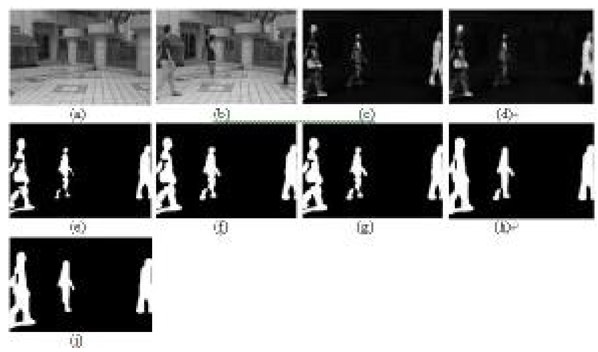


Fig. 13 The test image processing procedure correspond Fig. 12; (a) the one of frames on the video sequence in Y component, (b) the background image, (c) the result image of (a)-(b), (d) after median filter of (c), (e) after EB of (d), (f) after dilation operation of (e), (g) after opening operation of (f), (h) after closing operation of (g), (i) after discard small area pixel of (h).

Fig. 14 The extracted face images from the test sequence-3 in Fig. 12.

4. Conclusion

At the beginning of our approach, we use Y component of the YIQ representation of an image to simplify the processing. On the object extracting stage, we use image subtraction method to remove the background image. On the other hand, we use the medium filter, image dilation, image opening, image closing and small area discarding operation to detect object correctly. Meanwhile, the EB mapping technique is used to correctly convert the image from grayscale to binary even on different exposed environments. Finally, in order to obtain only one copy of the face image on the video sequence, an object location comparison with the previous frame method used to achieve the goal. To explore the utility and demonstrate the efficiency of the proposed scheme, several image sequences are conducted. The experiment results show that our proposed scheme can precisely extract the faces under duplicated objects on the sequence taken from recording monitor system.

Reference

[1] G. Yang and T. S. Huang, "Human face Detection in Complex Background," **Pattern Recognition**, vol. 27, no.1, pp53-63, 1994.

[2] T. K. Leung, M. C. Burl, and P. Perona, "Finding Face in Cluttered Scenes Using Random Label Graph Matching," **Proc. Fifth IEEE int'l Conf. Computer Vision**, pp. 637-644, 1995.

[3] M. H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting Faces in Images: A Survey," **IEEE Trans. on Pattern Anal. Mach. Intell**., Vol. 24, pp. 34-58, 2002

[4] http://www.cs.nthu.edu.tw/

[5] http://www.ee.ccu.edu.tw/

[6] http://www.cs.nccu.edu.tw/

[7] http://www.itri.org.tw/index.jsp.

[8] http://www.csie.ncu.edu.tw/Csie/

[9] C. Garcia, G. Zikos, and G. Tziritas, "Wavelet Packet Analysis for Face Recognition," **Image and Vision Computing**, Vol. 18, pp. 289-297, 2000.

[10] J. Wu and Z. H. Zhou, "Efficient Face Candidates Selector for Face Detection," **Pattern Recognition**, Vol. 36, pp. 1175-1186, 2003.

[11] J. Badenas, J. M. Sanchiz, and F. Pla, "Motion-based Segmentation and Region Tracking in Image Sequence," **Pattern Recognition**, Vol. 34, pp. 661-670, 2001.

[12] J. Wu and Z. H. Zhou, "Efficient Face Candidates Selector for Face Detection," **Pattern Recognition**, Vol. 36, pp. 1175-1186, 2003.

[13] Y. Ishii, H. Hongo, K. Yamamoto, and Y. Niwa, "Real-Time Face and Head Detection Using Four Directional Features, "**Proceedings of the Sixth IEEE int'l Conf**. on Automatic Face and Gestuer Recognition, pp. 1-6, 2004.

[14] A. Sad, W. A. Pearlman, "A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees", **IEEE Trans. on Circuit and Systems for video Technology**, Vol. 6, No.3, pp. 243-250, 1996.