

Emergency Detection Method using Motion History Image for a Video-based Intelligent Security System

Jun Lee^{*}, Se-Jong Lee^{*}, Jeong-Sik Park^{*}, Yong-Ho Seo[†]

^{*,†}Dept. of Intelligent Robot Engineering, Mokwon University, Korea

Abstract

This paper proposed a method that detects emergency situations in a video stream using MHI (Motion History Image) and template matching for a video-based intelligent security system. The proposed method creates a MHI of each human object through image processing technique such as background removing based on GMM (Gaussian Mixture Model), labeling and accumulating the foreground images, then the obtained MHI is compared with the existing MHI templates for detecting an emergency situation. To evaluate the proposed emergency detection method, a set of experiments on the dataset of video clips captured from a security camera has been conducted. And we successfully detected emergency situations using the proposed method. In addition, the implemented system also provides MMS (Multimedia Message Service) so that a security manager can deal with the emergency situation appropriately.

Key words: Gaussian mixture model (GMM), Motion history image (MHI), Video-based security system, Template matching

1. INTRODUCTION

Currently there is a special need of an intelligent emergency detection technique that automatically analyzes situations recorded in a live video stream captured from a security camera. Therefore, this paper proposes a method that can detect emergency situations for a video-based intelligent security system [1].

The proposed method traces human object in the field of view of a security camera after separating the foreground from the background in the image using conventional image processing techniques such as GMM (Gaussian Mixture Model) and labeling.

We also developed a vision-based security system based on the proposed method to record images captured from an IP camera installed under the ceiling of an indoor environment similar with the existing CCTV security system.

The system also utilizes multimedia message service to acknowledge the emergency situation to a security manager

so that he can deal with the current situation appropriately.

This paper is organized as follows. The proposed emergency detection method is presented in Section 2. Experimental result is presented in Section 3. Conclusion is finally given in Section 4.

2. EMERGENCY DETECTION METHOD USING MHI

The proposed method firstly extracts a foreground which is regarded as human object from the input image, then a human object tracking is needed to create a MHI of the current object from the sequential images in a given duration triggered by a similarity threshold between the current and the next the foreground human object.

Fig. 1 shows the flowchart of the foreground extraction and the human object tracking which are used as preprocessing techniques in our emergency detection method.

A. Foreground Extraction and Human Object Tracking

A foreground should be separated from a background to detect a human object. Conventional approach to get the foreground needs to model the background in a video stream.

Manuscript received Jan. 26, 2012; revised Feb.23, 2012

[†]Corresponding Author: yhseo@mokwon.ac.kr

Tel: +82-42-829-7521, Mokwon Univ.

^{*}Dept. of Intelligent Robot Engineering, Mokwon Univ., Korea

Therefore, we use GMM which is a robust and widely used method for background modeling. It uses multi-Gaussian distributions of a model in order to limitation of preventing regular Gaussian distribution falling into the maximum local point using average and standard deviation of distributed values. It is a parametric approach for background modeling in which they maintain a mixture of Gaussians for the underlying distribution for each pixel's color values. For each new frame, the mean and covariance of each component in the mixture is updated to reflect the change of the pixel values. If the new value is far enough from the mixture in a new image, it updates the current GMM-based background model and the foreground can be extracted by calculating difference operation between the entire image and the current background model [2].

Following parameters of each Gaussian component need to be learned dynamically. Equation (1) is average, equation (2) is variance and equation (3) is weight parameter of each Gaussian component.

$$\hat{\mu}_j = \frac{\sum_{n=1}^N P(\omega_j|X_n, \theta) X_n}{\sum_{n=1}^N P(\omega_j|X_n, \theta)} \tag{1}$$

$$\hat{\sigma}_j^2 = \frac{\sum_{n=1}^N P(\omega_j|X_n) \|X_n - \hat{\mu}_j\|^2}{\sum_{n=1}^N P(\omega_j|X_n)} \tag{2}$$

$$\hat{\delta}_j = \frac{1}{N} \sum_{n=1}^N P(\omega_j|X_n) \tag{3}$$

After getting the foreground binary image, noise is removed by applying dilation and erosion operation prior to doing labeling process as well.

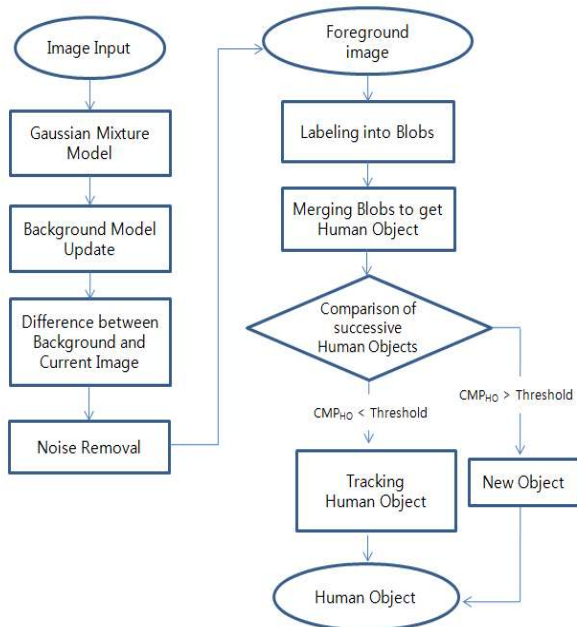


Fig. 1. Flowchart of foreground extraction and human object tracking

Tracking a human object is conducted after foreground extraction. Labeling is a process to get the blobs from the foreground image. The labeling in our code was implemented by applying the well-known blob labeling algorithm [3]. Then the near blobs need to be merged so that it can form a candidate of human object.

The tracking process also includes a comparison of the two human objects in successive image frames to determine whether the current human object is same with the previous object or new object. It can be calculated by measuring the difference of locations of center of masses the difference of sizes of two objects. In this research, we simply used a correlation-based approach by calculating the sum of pixel differences between two objects and the threshold value was defined to recognize the new human object.

In the Fig. 1, CMP_{HO} and Threshold denote the correlation value of two successive human objects and the predefined threshold value, respectively.

B. Emergency Detection using Motion History Image

Unlike the posture recognition which uses single image frame, a sequence of images should be considered to recognize human action. Given a sequence of images, we adapt a representation of motion history image (MHI) for a model of human action. The MHI accumulates an image sequence into a single image that captures spatial and temporal information about motion [4]. The MHI is known for its fast processing speed and ability to represent short-duration movement.

It is available to apply various types of MHIs, in this research we identified each human object by applying higher weight on recent frame over old frame when accumulating the human object [5]. The examples of MHI templates for emergency situations are shown in Fig. 2.



Fig. 2. Examples of MHI templates for emergency situations

Template matching was used to detect behavior of emergency situation. The size of the current MHI is

normalized to the predefined MxN gray image to be compared with the prepared MHI templates. Then we calculate the correlation coefficient using mean square error (MSE) as following equation (4).

$$Cr = 1 - \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [T(x_i, y_i) - I(x_i, y_i)]^2 \quad (4)$$

If the MHI of current situation is identical with one of the MHI templates, the value of the correlation coefficient, Cr will be close to 1. In this research, the threshold value in template matching was given as 0.4 to detect emergency situation. Fig. 3 shows flowchart of the emergency detection using MHI. In Fig. 3, Dth and TMth represent predefined threshold values.

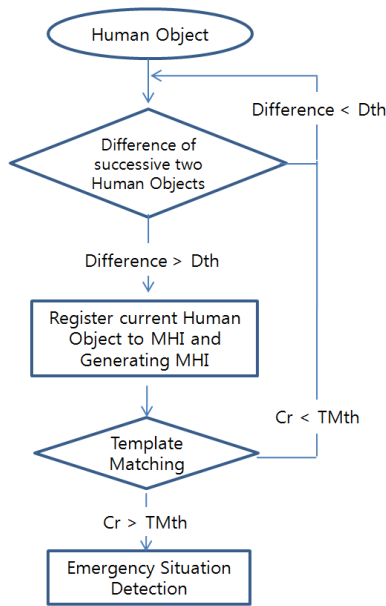


Fig. 3. Flowchart of emergency detection using MHI

3. EXPERIMENTAL RESULT

We installed the IP camera under the ceiling of a corridor in height of 2.5m and width of 1.5m for our experiment. We reproduce the situations of falling down or squatting down as emergency cases. Fig. 4 represents a result of emergency situation detection and MHI template of falling down, while Fig. 5 shows a notification to a security manager of the current emergency via MMS with the captured scene [6].

The experimental result of emergency detection is represented in the following Table I. The considered emergency situations for our experiment are squatting down and falling down.

Table I. Experimental result of emergency detection

| Emergency Case | Number of Tests | Number of Positive Alarm | Detection Rate (%) |
|----------------------------|-----------------|--------------------------|--------------------|
| Squatting down/Standing up | 50 | 28 | 56 |
| Falling down | 50 | 50 | 100 |
| Total | 100 | 78 | 78 |

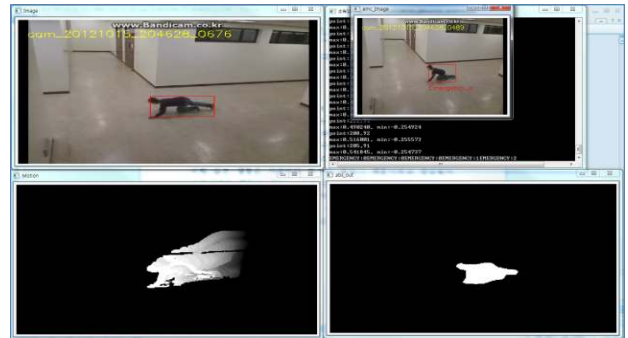


Fig. 4. Result of emergency situation detection and MHI template of falling

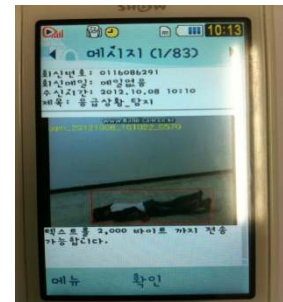


Fig. 5. Notification to a security manager via MMS

4. CONCLUSION

This paper proposes an emergency detection method using MHI for a Video-based Intelligent Security System. The developed security system using the proposed method also delivers a notification of the current emergency via MMS to a manager. The proposed method includes foreground extraction, human object tracking to create a proper MHI and template matching for detecting emergency situation.

Finally the feasibility and effectiveness of the proposed method has been successfully verified by conducting several experiments of emergency detection of squatting down and falling down under similar indoor environment with conventional CCTV security system.

For the future work, we will enhance a quality of foreground by considering external factors such as light, angle and walking speed of human object, and also enhance an accuracy of emergency detection rate. The proposed

method will be incorporated with the network based remote surveillance system with mobile robot [7].

ACKNOWLEDGMENT

This work was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(MEST) (2011-0013776). This work was also supported by the NAP(National Agenda Project) of the Korea Research Council of Fundamental Science & Technology.

REFERENCES

- [1] I. W. Jeong, J. Choi, K. Cho, Y. H. Seo, H. S. Yang, "A Vision-Based Emergency Response System with a Paramedic Mobile Robot," *IEICE TRANSACTIONS on Information and Systems*, Vol. E93-D, No. 7 pp.1745-1753, 2010
- [2] H. Samet and M. Tamminen, "Efficient Component Labeling of Images of Arbitrary Dimension Represented by Linear Bintreees," *IEEE Trans. Patt. Analy. And Mach. Intell.*, Vol. 10, No. 4, pp. 579-586, Jul, 1988.
- [3] Stauffer C., and Grimson W. E. L., "Adaptive background mixture models for real-time tracking," in *Proc. of IEEE Computer Society Conf. on Comp. Vis. and Patt. Recg.*, Vol.2, pp.246-252, 1999
- [4] A. F. Bobick and J. Davis, "The recognition of human move-ment using temporal templates," *IEEE Trans. Patt. Analy. And Mach. Intell.*, Vol. 23, No. 3, pp.257-267, 2001.
- [5] Tae-Woo Han, Yong-Ho Seo, "Emergency Situation Detection using Images from Surveillance Camera and Mobile Robot Tracking System," *Journal of the Institute of Webcasting, Internet and Telecommunication(IWIT)*, Vol. 9, No. 5, pp.101-107, May. 2009.
- [6] CTM text message delivering module offered by D&SOFT <http://open.coolsms.co.kr>
- [7] Yong-Ho Seo, "Development of Network based Remote Surveillance System Using Omni-Directional Mobile Robot," *Journal of the Institute of Webcasting, Internet and Telecommunication(IWIT)*, Vol. 10, No. 4, pp.91-97, Aug. 2010.



Jun Lee is a currently B.S. student at Department of Intelligent Robot Engineering, Mokwon University. He is currently a member of Intelligent Robotics Lab. His research interests are in the areas of Computer Vision, Activity Recognition and Vision-based Surveillance System.



Se-Jong Lee is a currently B.S. student at Department of Intelligent Robot Engineering, Mokwon University. His research interests are in the areas of Computer Vision, Activity Recognition and Vision-based Surveillance System.



Jeong-Sik Park received his B.E. degree in Computer Science from Ajou University, South Korea in 2001 and his M.E. and Ph.D. degree in Computer Science from KAIST (Korea Advanced Institute of Science and Technology) in 2003 and 2010, respectively. He is currently a Professor of the Department of Intelligent Robot Engineering, Mokwon University. His research interests include speech emotion recognition, speech recognition, speech enhancement, and voice interface for human-computer interaction. Prof. Park is a member of the Institute of Webcasting, Internet and Telecommunication(IWIT) of Korea.



Yong-Ho Seo received his BS and MS degrees from the Department of Electrical Engineering and Computer Science, KAIST, in 1999 and 2001, respectively. He also received a PhD degree at the Artificial Intelligence and Media Laboratory, KAIST, in 2007. He was an Intern Researcher at the Robotics Group, Microsoft Research, Redmond, WA in 2007. He was a consultant at Qualcomm CDMA Technologies, San Diego, CA in 2008. He is currently a Professor of the Department of Intelligent Robot Engineering, Mokwon University. His research interests include humanoid robots, human-robot interaction, robot vision and wearable computing. Prof. Seo is a member of the Institute of Webcasting, Internet and Telecommunication(IWIT) of Korea.