International Journal of Energy, Information and Communications
Vol. 2, Issue 2, May 2011

A Study on the Development of Multiple Objects Extraction System Using Difference Image Edge Information

Na-Ra Shin* and Kicheon Hong**[1]
**Dept. of Information & Telecommunications Engineering, The University of Suwon, Wau-ri, Bongdam-eup, Hwaseong-si, Gyeonggi-do, 445-743, Korea*
E-mail: nara9n@suwon.ac.kr

***Dept. of Information & Telecommunications Engineering, The University of Suwon, Wau-ri, Bongdam-eup, Hwaseong-si, Gyeonggi-do, 445-743, Korea*
E-mail: kchong@suwon.ac.kr

Abstract

This paper proposes a method of extracting motion information of a single movement object and plurality of movement objects. After area segmentation is performed using a difference image between neighboring images, a clear outline of an object and a loss part are restored by using edge information and a boosting factor δ_i *according to adjustment corresponding to a change of an input image. Whether the object is a person or a thing is determined by extracting width, height, and size information using the restored object area and setting a size of the object based on the extracted information. To obtain correct information of the object, the object is finally extracted by analyzing a shadow area of the object generated through the full process and removing the shadow. The proposed system efficiently extracts objects of various images.*

Keywords: *Difference Image, Edge Detection, Object Detection, Multi Object, Boosting Factor.*

1. Introduction

People tracking, which started from an expression of an interaction between a person and a computer, has been used in many application fields, such as robot learning, object counting, and monitoring systems, and in particular, the importance of the development of a monitoring system for automatically discovering an illegal act by recognizing and tracking people using cameras is increasing day-by-day. Examples of security systems popularly used presently are an unmanned monitoring system, an identification verification system, and a criminal search system used in department stores or shops. Since a function of real-time extracting an action of a person, tracking and recognizing the person, and processing the meaning of the action is essential to security systems, such as a subway safety-accident prevention system and a risk prevention and management system, for monitoring and tracking people, a correct people movement tracking technique may be a core technique for the security systems. Methods of acquiring and tracking people moving in an indoor environment have been proposed[1]. However, this paper proposes a method of extracting an object, perceiving a position of the object, and tracking and analyzing the object through information learning in a complex indoor and outdoor environment with a domestic system instead of an overseas system.

[1] Corresponding Author

2. Multiple Objects Extraction

2.1. Overall system flowchart

In general, people tracking is performed by extracting people from a record of a video or image and analyzing position information and a movement pattern, such as an action, of people to be tracked. Many methods are proposed to extract people; particularly, the representative ones are “model-based”, “shape-based”, and “method using statistical characteristics of blobs”.

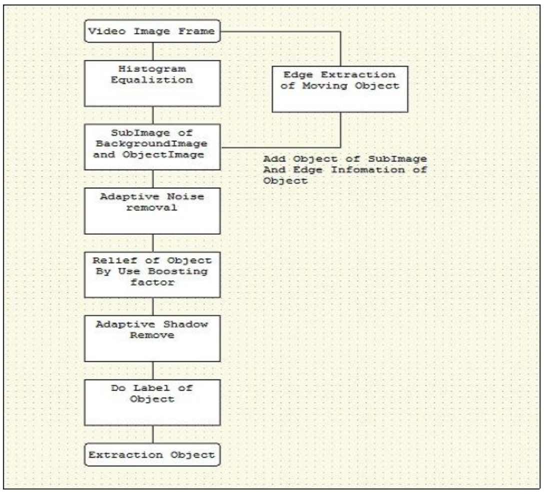


Figure 1. Multi-object extraction system flowchart

This paper proposes multi-object extraction, which is correct and robust to indoor and outdoor noise, by merging edge information of moving objects and removing shadows obstructing extraction of an original size of each object, as shown in Figure 1.

2.2. Histogram equalization

A camera and an image sensor must manage well not only a brightness and darkness contrast of a scene but also exposure of an image. In general, a brightness range of light is too wide for a sensor to perfectly express. Once a camera captures an image, it is impossible to change the image. However, it is still possible to change a dynamic area of a pixel value of the image. To do this, the most popularly used scheme is histogram equalization[2], which is based on a mathematical scheme of transforming a predetermined distribution function to another type of distribution function, wherein a brightness value histogram of an input image is an input distribution function, and an ideal uniform distribution function is an output function. That is, histogram smoothing is performed for a histogram distribution of the input image to be as uniformly spread as possible[3]. To uniformly spread a predetermined distribution, an accumulated distribution function may be used. While an application of this scheme to a continuous function results in a transform to a perfectly uniform distribution, an application of histogram equalization to a discrete image results in a type of distribution that is more or less different from a uniform distribution.

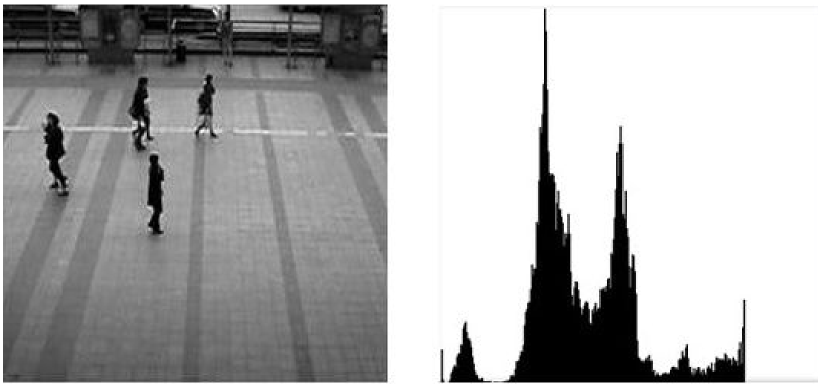


Figure 2. Histogram distribution diagram of experimental image

Figure 2 shows a histogram distribution diagram of an experimental image for extracting multi-objects and also shows that a brightness distribution of the image is mainly distributed in the middle. If the brightness distribution is not uniform, an error or a problem that is hard to solve may occur in image processing, thus histogram equalization is necessary. Figure 3 shows an image obtained by performing histogram equalization on the image shown in Figure 2.

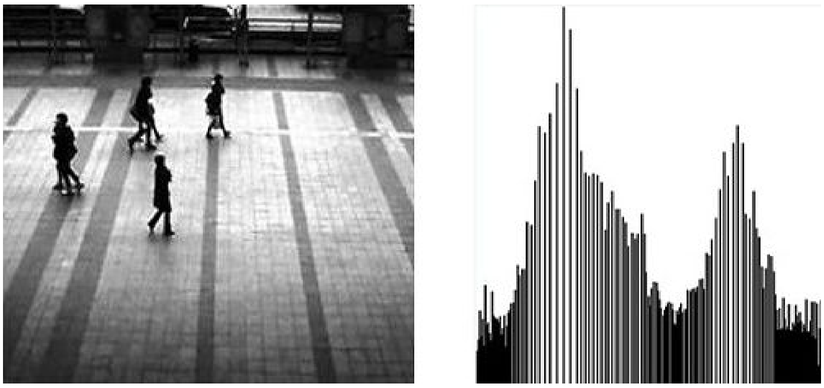


Figure 3. Image and distribution diagram after histogram equalization

2.3. Object extraction

In the image after histogram equalization, clear outlines of moving objects are obtained by merging partial outlines obtained from a difference image between a background image and the moving objects and edge information extracted from a color image. A more robust object area is extracted by expanding an area of the moving objects using a Boosting factor.

2.3.1. Object area extraction: A difference image indicates a brightness value difference between two pixels corresponding to the same position in two images. An object extraction result shown in Figure 4 may be obtained using Equation (1).

$$\delta I(x,y) = |I_t(x,y) - I_{t-1}(x,y)| \tag{1}$$

$\delta I(x,y)$ indicates a brightness value difference between pixels located at coordinates x and y[4][5].

Figure 4. Object extraction image obtained through difference image

2.3.2. Object outline enhancement: In this paper, information regarding a position, a shape, and a size of an object may be obtained by detecting an edge through a merge of edge information detected from the object and an image obtained from a difference image, in a method of enhancing outlines of moving objects[6]. In the edge detection method used in this paper, edge detection is performed by using a 3×3 Laplacian mask of Equation (2).

$$\mathcal{L}\,(f) = \frac{\sigma^2 f}{\sigma x^2} + \frac{\sigma^2 f}{\sigma y^2} \tag{2}$$

An edge image shown in Figure 5 is obtained by replacing pixels in which zero-crossings occur with a white color and the remaining pixels with a black color in an experimental image through Equation (2)[7].

The picture, which is obtained by magnifying an object in the detected edge image, shows that a background and the object are separated from each other along an outline.

2.3.3. Adaptive object area expansion: As shown in Figure 6, a divided object may be extracted case-by-case with only a moving object area obtained from the difference image and the detected edge information.

Because of the above problem, a Boosting factor is applied to the experimental image through Equation (3) so that a desired object and divided area can be adaptively extracted according to an environment by weighting the object area with the Boosting factor and setting a predetermined threshold.

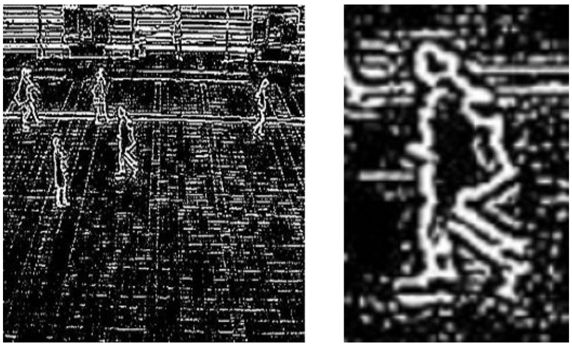


Figure 5. Edge image detected from experimental image

Figure 6. Normally detected object area (left) and abnormally extracted object area (right)

$$B(x,y) \simeq \begin{cases} Object, & f(x,y)*\delta_i \geq T \\ Background, & Otherwise \end{cases} \tag{3}$$

Figure 7, which is obtained using Equation (3), shows a result image obtained by restoring the object divided into two by weighting an empty part, which was not extracted as an object, as an object area.

Figure 7. Restoration of object area

As shown in Figure 8, since an image for δ_2 is better for object area expansion than for δ_1 in the environment of the experimental image, the image for δ_2 is used for the experiment.

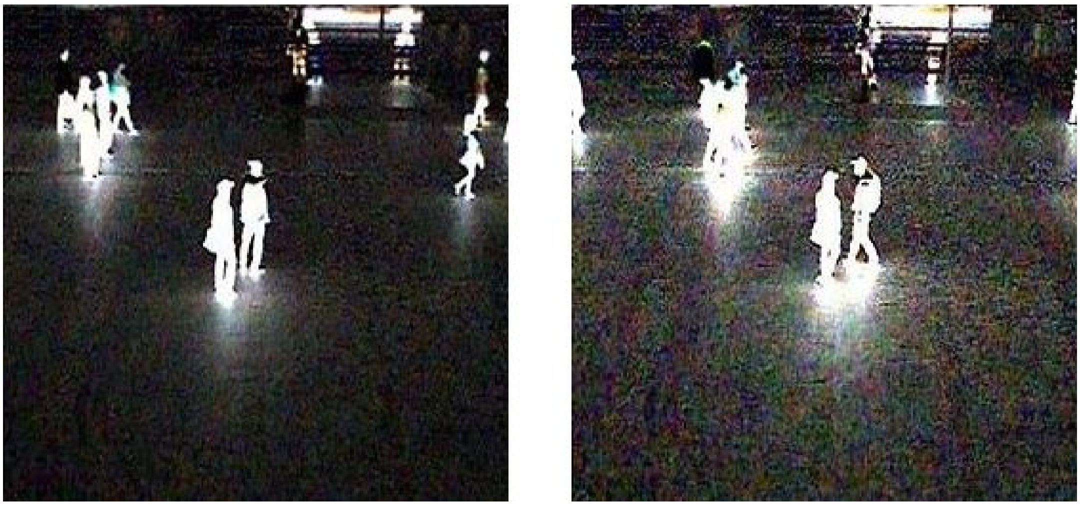


Figure 8. Images for i=1 (left) and for i=2 (right) in the Boosting factor δ_i

3. Adaptive Labeling

Labeling is used to apply various functions, such as position information of objects, sizes of the objects, and discrimination of a human movement from a vehicle movement. In detail, labels having the same number are allocated to all neighboring pixels, and labels having different numbers are allocated to other components not connected to each other.

To do labeling, all pixels of an image are examined i.e., whether a corresponding pixel was visited before, the number of neighboring pixels, and examining of allocation of a number to a corresponding blob.

$$N(x,y) = \int_{j=0}^{h} \int_{i=0}^{w} f(x_i, y_j) = 255 \ f(x_{i+1} y_{j+1}) = 255 \tag{4}$$

Equation (4) determines neighboring pixels based on two conditions while examining all pixels of an input image. Here, h denotes a height of the input image, and w denotes a width of the input image. Neighboring pixels are determined from pixels (0, 0) to (x_{width}, y_{height}) using Equation (4). If a predetermined pixel is recognized as a pixel of an object, whether a next pixel at a next column, i.e., x_{i+1}, is also a pixel of the object is examined, and if the next pixel is a pixel of the object, a next pixel is continuously examined. Otherwise, the examination continues from a pixel at a next row, i.e., y_{i+1}, of the input image.

Figure 9 shows images which labeled objects are marked with a rectangle, wherein the left one is an input image and the right one is an image obtained by binarizing the input image with a threshold suitable for an environment.

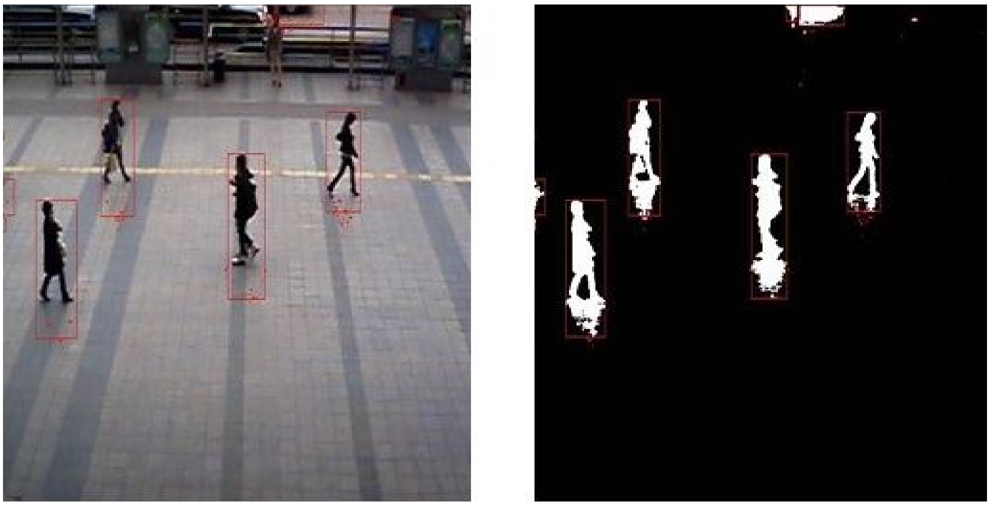


Figure 9. Simple labeling result

Comparing the images shown in Figure 10, obtained by magnifying the images of Figure 9 with each other, since even neighboring pixels having a small size are recognized as an object, these neighboring pixels may become noise in extraction of moving objects, thereby resulting in a problem in the number of objects.

Figure 10. Noise in simple labeling

Unlike general labeling, since the adaptive labeling includes much information, such as a size, a width, and a height of each detected object, the number of detected pixels, and the number of labels in an image, to be applicable to an indoor and outdoor environment, the noise shown in Figure 10 can be canceled using Equation (5) based on this information.

$$Object = \begin{cases} Object, & PixelSize > T \\ Noise, & Otherwise \end{cases} \tag{5}$$

Using Equation (5), pixels having a pixel size equal to or less than a threshold T are recognized and removed as noise not included in an object, by checking sizes of pixels in each object area. A result image is shown in Figure 11.

Figure 11. Noise canceling result

Besides, a vehicle, a cart, or the like recognized together with a person can be excluded from object recognition by using height and width information of each recognized object.

4. Shadow Removing

According to a binary image shown in Figure 11, a size of a moving object is changed because of a shadow of the object. Shadow removing is required to remove the shadow and recognize the original size of the object. A shadow removing algorithm is based on the brightness and darkness of a shadow. If light is strong, a shadow according to the light is dark on the contrary. A brightness distribution diagram of a shadow area is obtained by performing a histogram analysis of V for outputting a brightness value in an HSV region[8].

For a shadow of a moving object, a histogram of a shadow area in a moving object area is analyzed, an intensively distributed part is set as the shadow area, and the shadow is removed by removing the shadow area. To apply the shadow removing algorithm to outdoor environments, the minimum brightness value and the maximum brightness value are adjusted by setting two brightness thresholds as defined in Equation (6) so that the shadow removing algorithm is adaptable to outdoor environments.

$$Shadow(x,y) = T_{\min} < Object(x,y) < T_{\max} \tag{6}$$

$Object(x,y)$ denotes an image of a moving object before a shadow thereof is removed, and a value greater than the threshold T_{min} and less than the threshold T_{max} in the image $Object(x,y)$ indicates a shadow in a brightness area. That is, a brightness part of a shadow according to light can be determined with the two thresholds. Figure 12 shows an image before shadows therein are removed and an image obtained by removing shadows therefrom.

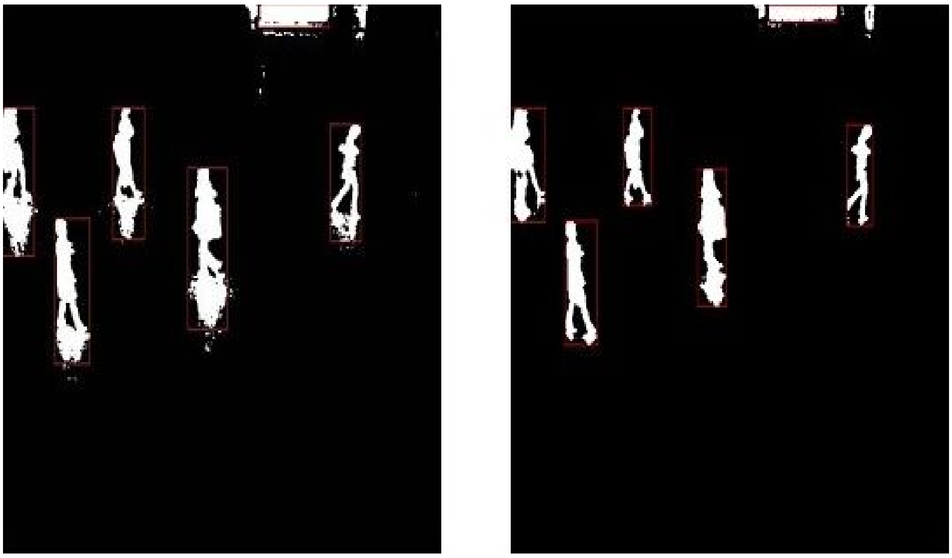


Figure 12. Image before removing shadows (left) and image after removing shadows (right)

Figure 13 shows that a size of a moving object is restored.

Figure 13. Size of restored object

5. Experimental Result

5.1. Boosting factor threshold

The suggested scheme was experimented with image samples captured outside, and to perform excellent multi-object extraction in outdoor environments, settings optimized to each image are obtained by adjusting a total of four thresholds. Figure 14 shows a δ_i value of a Boosting factor suitable for 5 images.

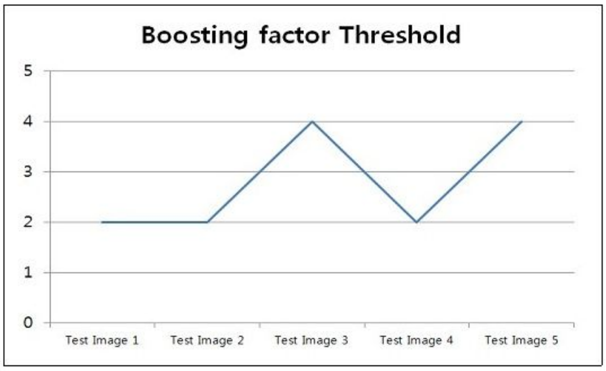


Figure 14. δ_i value according to images

As shown in Table 1, 2 and 4 are allocated as the δ_i value, thereby expressing an object area clearly in the experimental images. The δ_i value is adjusted high in most dark images and adjusted low in most bright images. In general, since the Boosting factor is constant for the dark images and the bright images, the δ_i value adaptively input for the images is valid.

5.2. Noise Remove threshold

Figure 15 shows a threshold for canceling noise of a result obtained from a relief of an object area and a merge of edge information.

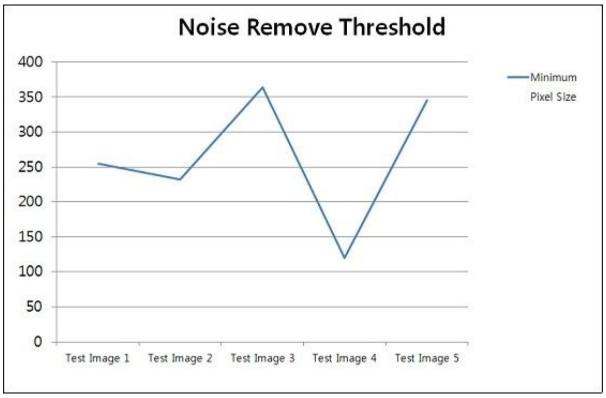


Figure 15. Cancellation of noise according to image

The minimum pixel size depends on a size of an object in an image. That is, a Test Image 3 is an image in which the largest objects are shown, and a Test Image 4 is an image in which the smallest objects are shown.

5.3. Shadow remove threshold

Figure 16 shows two thresholds used to cancel a shadow. The maximum threshold indicates the maximum value of an area in which a shadow exists among V values indicating a brightness area, and the minimum threshold indicates the minimum value of an area in which a shadow exists among the V values. A shadow can be removed by removing an area having values between the two thresholds.

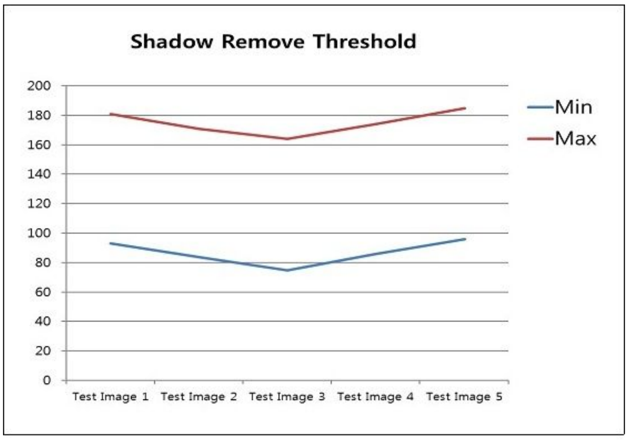


Figure 16. Distribution of shadows according to image.

5.4. Experiments with image samples

The suggested scheme was experimented with image samples captured in many different kinds of environment. As shown in Figure 17, the scheme shows the excellent multi-object extraction by using the optimized thresholds.

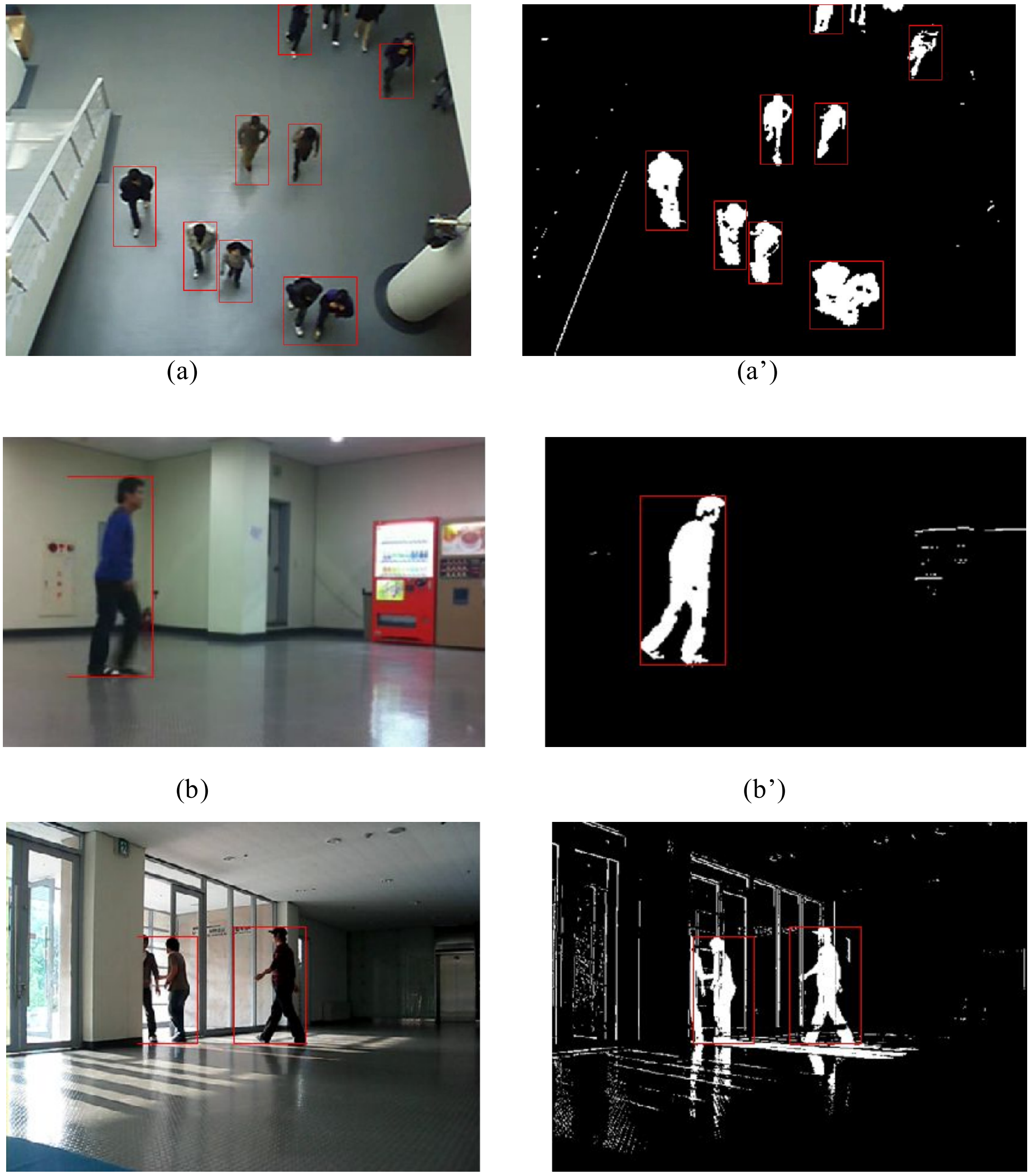
6. Conclusion

In the research, an algorithm adaptable to various outdoor environments is proposed, wherein a movement of a moving object is estimated by using a difference image scheme, a desired object area is expanded by weighting objects in an image with Laplacian edge detection and a Boosting factor for unclear objects, a threshold is set based on height and width information of each object through labeling, and if a moving object corresponds to a value less than or greater than the threshold, the moving object is removed as an undesired one.

In addition, to remove shadows in outdoor environments, a system for searching for shadows by a user setting the maximum and minimum thresholds to adjust a brightness area of the shadows has been studied.

Multi-objects are efficiently extracted by using a system developed through this research.

Beyond the object tracking using a center point based on an object probability distribution of an input image, the development of an intelligent multi-object tracking algorithm using various unique feature points of objects extracted by the system still remains a task for future development.



(a) (a')

(b) (b')

(c) (c')

Figure 17. Multi-object extraction

Acknowledgments

This work was supported by the GRRC program of Gyeonggi Province [(GRRC SUWON 2010-B4), Development of Event Extraction System for Intelligent Surveillance]

References

[1] N. Ahmed, T. Natarajan, and K. R. Rao, “Discrete cosin transform,” IEEE Transactions on Computers, Vol. 23, pp. 90-93, 1974.
[2] W. T. Freeman and M. Roth, “Orientation histograms for hand gesture recognition,” International WorkShop on Automatic Face and Gesture Recognition, pp.296-301, June 1995.
[3] [OpenCV] Open Source Computer Vision Library(OpenCV), http://sourceforge.net/ projexts/opencvlibrary/.
[4] M. Hu “Visual pattern recognition by moment invariants,” IRE Transactions on Information Theory, Vol. 8, pp. 179-187, 1962.
[5] J. Philbin, O. Chum, M. Isard, J.Sivic, and A.Zisserman, “Object retrieval with large vocabularies and fast spatial matching” Proceedings of the IEEE Conference on Computer Vision and pattern Recognition, 2007.
[6] J. Canny, “A computational approach to edge detection,” IEEE Transactions on Pattern Analysis and Machine Inteligence, Vol. 8, pp.679-714, 1986.
[7] Jianping Fan, David. K. Y. Uai, Ahmed K. Elmagamid, and Walid G. Aref, “Automatic Image Segmentation by Intergrating Color-Edge Extraction and Seeded Region Growing,” IEEE Transaction On Image Processing, Vol. 10, No. 10, Oct 2001.
[8] W. N. Martin, J. K. Aggarwal, “Survey: Dynamic Scene Analysis,” Computer Graphics and Image Processing, Vol. 7, pp.356-374, 1978.

Authors

Na-Ra Sin received the B.E. degree in Information and Telecommunications Engineering from The University of Suwon, Seoul, Korea in 2011 and is currently studying for a M.S. degree. His research interests include image processing, pattern recognition and multiple objects tracking.

Kicheon Hong received the B.E. degree in electronics engineering from Sungkyunkwan University, Seoul, Korea in 1985 and M.S. and Ph.D. degrees in Electrical and Computer Science from Stevens Institute of Technology in 1988 and 1994, respectively. He was with Bell Communications Research Center from 1992 to 1993 and Samsung Semiconductor Co. Ltd. from 1994 to 1998. Since 1998, he has been with the Department of Information and Telecommunications Engineering, The University of Suwon, Seoul, Korea, where he is currently an Associate Professor. His research interests include image processing and compression, multimedia processor, embedded system and pattern recognition.