Real Time Object Tracking and Identification Using a Camera

Budi Sugandi, Hyoungseop Kim, Joo Kooi Tan and Seiji Ishikawa
Graduate School of Engineering
Kyushu Institute of Technology, Japan
1-1 Sensui-cho, Tobata-ku, Kitakyushu-shi, Fukuoka, 804-8550, JAPAN
Phone: +81-93-884-3185 Fax: +81-93-861-1159
budi@kimlab.cntl.kyutech.ac.jp

Abstract: In this paper, we proposed a method for detecting and tracking of moving objects based on low resolution image employing a block matching technique and also proposed an identification method using a color and spatial information. Many tracking algorithms have better performance under static background but sometimes mistracking results are obtained under background with complex motions. Since a low resolution image has a nice property that it can remove the small size pixels, it is adopted to solve this problem due to the fact that most of the fake motions in the background have small region. In tracking the moving object, many applications have problems when objects occlude each other. The peripheral increment sign correlation is used to solve this problem. The identification object is performed using a color and spatial information of the tracked object. The experimental results prove the feasibility and usefulness of the proposed method.

1. Introduction

Automatic detection and tracking of moving object is very important task for security system, monitoring activity and surveillance application. Various imaging techniques for detecting and tracking the moving objects have been proposed by many researchers such as frame difference, background subtraction, optical flow, skin color extraction and probability based approaches. Liu *et al*. [1] proposed background subtraction to detect moving regions in an image by taking the difference between current and reference background image in a pixel-by-pixel. It is extremely sensitive to change in dynamic scenes derived from lighting and extraneous events etc. Lipton *et al*. [2] proposed frame difference that use of the pixel-wise differences between two or three successive frame images to extract the moving regions. This method is very adaptive to dynamic environments, but generally does a poor job of extracting all the relevant pixels, e.g., there may be holes left inside moving entities. Meyer *et al*. [3] proposed an optical flow method by computing the displacement vector field to initialize a contour based tracking algorithm, called active rays, for the extraction of articulated objects. The optical flow method can be used to detect independently moving objects even in the presence of camera motion. However, most flow computation methods are computationally complex and very sensitive to noise. The method proposed by Paragios *et al*. [4] considered a probability and statistics problem, and used the observed information to obtain a classification equation of probability to segment image. These detection models can detect and track moving objects, but they suffer from the high computing cost. So it is hard to be used in real-time surveillance application. Davis *et al*. [5] proposed the W4 method that designed a simplified statistics model to make real-time processing possible. First, the proposed method learned the background by the luminance of image and used a statistical cycle to observe the variance range of every pixel in the frames that have no moving objects. After the process, the method can build a background model by the observed information and compute the inter-frame difference of the current frame and background model to detect the moving pixels. The result of W4 is good and does not need much computing cost, but the background model has to be often re-established.

To overcome those problems, we propose a method to detect and track the moving object based on low resolution image employing block matching technique. The blocks are made based on peripheral increment sign correlation (PISC) [6] image. We propose also the object identification method employing color and spatial information of moving object [7].

2. Methods

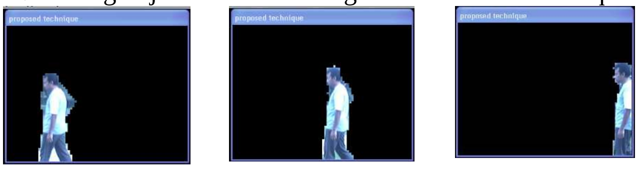
We categorize our study in three stages; object detection, object tracking and object identification. On each stage, we apply our new method and evaluate the effectiveness of our method.

The first stage is done to detect the moving object emerging in the background. In this stage, we make a low resolution image to remove the image noise having pepper and salt noise and small movement in the background such as moving leaves. The second stage is performed to track the moving object. On this stage, we propose a block matching technique made based on the PISC image that considers the brightness change in all the pixels of the block relative to the considered pixel. As a last stage, we identify the tracked object by extracting the color and spatial information. In this paper, we use mean and standard deviation value as image feature of the tracked object. The details of each technique are described as follows.

2. 1 Low resolution image

A low resolution image is made by reducing spatial resolution of the image while keeping the same image size [8-9]. In this paper, the low resolution image is made by averaging pixels value of its neighbors. The low resolution image can be used for removing the scattering noise having salt and pepper noise and the small movement in the background. These noises that have small region will be disappeared in the low resolution image.

a. moving object detection using a conventional technique



b. moving object detection using a low resolution image

Figure 1. Comparison of moving object detection

2. 2 Frame difference

Frame difference [10-11] is computing the difference image between successive frame images. In this paper, we applied a frame difference method to detect the moving objects on the three successive frames which is obtained by the low resolution image.

The comparison of moving object detection using a conventional method and a low resolution image is shown in Figure 1. As shown in that figure, using the conventional method (frame difference technique), the moving object is still greatly affected by the noise such as moving leaves. In the other hand, using a low resolution image, that kind of noise is removed completely.

2.3 Morphological operation

Morphological operation is performed to fill small gaps inside the moving object and to reduce the noise remained in the moving objects [12]. The morphological operators implemented are dilation followed by erosion. Dilation adds pixels to the boundary of the object and closes isolated background pixel. Erosion removes isolated foreground pixels.

2.4 Connected component labeling

Connected component labeling is performed to label each moving object emerging in the background. The connected component labeling groups the pixels into components based on pixel connectivity (same intensity or gray level) [8]. In this paper, connected component labeling is performed by comparing a pixel with the pixels in the four neighbors from top-left to bottom-right and from bottom-right to top-left.

2.5 Block matching

Block matching is a technique for tracking the target object among the moving objects emerging in the background. The blocks are defined by dividing the image frame into non-overlapping square parts. The blocks are made based on PISC image [6] that considers the brightness change in all the pixels of the blocks relative to the considered pixel. Figure 2 shows the block in PISC image with block size 5×5 pixels. The blocks of the PISC image in the previous frame are defined as shown in equation (1). Similarly, the blocks in the current frame are defined in equation (2). To determine the matching criteria of the blocks between two successive frames, we evaluate using correlation value that express in equation (3). The high correlation value shows that the blocks are matched each other. The target moving object is determined when the numbers of matching block in the previous and current frame are higher than a certain threshold value.

$$b_{np} = \begin{cases} 1, & \text{if } f_{np} \geq f(i,j) \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

$$b'_{np} = \begin{cases} 1, & \text{if } f_{np} \geq f(i,j) \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

$$corr_n = \sum_{p=0}^{24} b_{np} * b'_{np} + \sum_{p=0}^{24} (1 - b_{np}) * (1 - b'_{np}) \tag{3}$$

where b and b' are the block in the previous and current frame, n is number of block and p is number of pixel of block, respectively.

2.6 Tracking the moving object.

To reduce the processing time, blocks are made only in the moving object area. The target moving object is tracked when the numbers of matching block is more than certain threshold. If that matching criteria is not satisfied, the matching process is repeated by enlarging the tracking area. If the target moving person still can not be tracked, the moving object is categorized as not target moving object and the tracking process is begun again from the beginning.

2.7 Feature extraction and object identification

Object identification is the last stage of our study. In this paper, the extracted features are divided into two types; color and spatial information of the moving objects.

The RGB color is used as color information of the moving objects. To obtain more color information for identification, we divide the moving object into three parts; the head, the upper and the lower part of the body. However, we only calculate the color information of upper and lower part of the human body.

The first color information calculated is mean value of each human body part as calculated by equation (4). The mean value is calculated for each color component of RGB space.

$$\mu_{f_k}^{O^i} = \frac{\sum_{x=x^i_{\min}}^{x^i_{\max}} \sum_{y=y^i_{\min}}^{y^i_{\max}} f_k(x,y)}{\#O^i} \tag{4}$$

where i is number of the moving objects and (x, y) is the coordinate of pixels in moving object. $\left(x^i_{max}, y^i_{max}\right)$ and $\left(x^i_{min}, y^i_{min}\right)$ are the maximum and minimum coordinates of moving object i, $f_k(x,y)$ denotes pixel value for each color component in RGB space of the current frame, O^i denotes

$f_{i-2,j-2}$	$f_{i-1,j-2}$	$f_{i,j-2}$	$f_{i+1,j-2}$	$f_{i+2,j-2}$
$f_{i-2,j-1}$	$f_{i-1,j-1}$	$f_{i,j-1}$	$f_{i+1,j-1}$	$f_{i+2,j-1}$
$f_{i-2,j}$	$f_{i-1,j}$	$f_{i,j}$	$f_{i+1,j}$	$f_{i+2,j}$
$f_{i-2,j+1}$	$f_{i-1,j+1}$	$f_{i,j+1}$	$f_{i+1,j+1}$	$f_{i+2,j+1}$
$f_{i-2,j+2}$	$f_{i-1,j+2}$	$f_{i,j+2}$	$f_{i+1,j+2}$	$f_{i+2,j+2}$

Figure 2. Block in PISC image

the set of the coordinate in the interested moving object i and $\#O^i$ is the number of pixels of moving object i, respectively.

We can extract more useful color features by computing the standard deviation of each human body part as shown in equation (5).

$$SD_{f_k}^{O^i} = \sqrt{\frac{\sum_{x=x_{\min}^i}^{x_{\max}^i}\sum_{y=y_{\min}^i}^{y_{\max}^i}\left(f_k(x,y)-\mu_{f_k}^{O^i}\right)^2}{\#O^i}} \qquad (5)$$

where, $\mu_{f_k}^{O_i^i}$ is the mean value and $SD_{f_k}^{O_i^i}$ is the standard deviation of each color component of the moving object, respectively.

The feature of objects extracted in the spatial domain is the position of the tracked object. The bounding box as defined in equation (6) is used as spatial information of moving objects.

$$B_{\min}^i = \left\{\left(x_{\min}^i, y_{\min}^i\right) \;\middle|\; x, y \in O^i\right\}$$
$$B_{\max}^i = \left\{\left(x_{\max}^i, y_{\max}^i\right) \;\middle|\; x, y \in O^i\right\} \qquad (6)$$

where, B^i_{min} is the left-top corner coordinates and B^i_{max} is the right-bottom corner coordinates, respectively.

After the extracted feature is obtained, we then calculate the similarity between the tracked object and the identified object as expressed in equation (7). The object with high similarity compared to certain threshold shows the similar object to the identified object, otherwise it will identify as different object.

$$S\left(F^i, F^j\right) = Mc\left(\left|\mu_{f_k}^{O_i^i} - \mu_{f_k}^{O_i^j}\right|\right) + Mc\left(\left|SD_{f_k}^{O_i^i} - SD_{f_k}^{O_i^j}\right|\right) + 0.5Mp\left(\left|B_{\min}^i - B_{\min}^j\right|\right) + 0.5Mp\left(\left|B_{\max}^i - B_{\max}^j\right|\right) \qquad (7)$$

where Mc and Mp are membership function for color and spatial information.

3. Experimental result

To evaluate our proposed methods, we did several experiments under outdoor environments with the noisy background. The experiments are implemented in the real time processing and applied on various image sequences captured using a camera. We tracked the target moving object from two or three moving objects emerging in the background. The tracking results are shown in Figure 3 - Figure 5. The identification result is shown in Table 1. In the experimental results, we can extract the moving objects on the successive frame and identification rates of 92.8% were achieved.

4. Conclusion

This paper proposed a new method for detecting and tracking the target moving person based on low resolution image employing PISC image for real time application and an identification method using color and spatial information of the moving object. Using our new method, the satisfactory results are achieved. In the experimental results, we can extract the moving objects on the successive frame and identification rates of 92.8% were achieved.

However, our proposed method still has limitations. The improvement of the methods is necesarry to improve the algorithm and speed up the processing time. These are remaining for our future works.

References

[1] Y. LIU , A. Haizho and Xu Guangyou, “Moving object detection and tracking based on background subtraction”, *Proc. Society of Photo-Optical Instrumentation Engineers (SPIE)*, Vol. 4554, pp. 62-66, 2001.

[2] A. J. Lipton, H. Fujiyoshi, and R. S. Patil, “Moving target classification and tracking from real-time video”, *Proc. IEEE Workshop Applications of Computer Vision*, pp. 8-14, 1998.

[3] D. Meyer, J. Denzler and H. Niemann, “Model based extraction of articulated objects in image sequences for gait analysis”, *Proc. IEEE Int. Conf. Image Processing*, pp. 78-81, 1998.

[4] N. Paragios and R. Deriche, “Geodesic active contours and level sets for the detection and tracking of moving object”, *IEEE Trans. Pattern Anal. Mach. Intel.*, Vol. 22, No.3, pp. 266-280, 2000.

[5] I. Haritaoglu, D. Harwood and L.S. Davis, “W4: real-time surveillance of people and their activities”, *IEEE Trans. Pattern Anal. Mach. Intel.*, Vol. 22, No. 8, pp. 809-830, 2000.

[6] Y. Satoh, S. Kaneko and S. Igarashi, “Robust object detection and segmentation by peripheral increment sign correlation (PISC) image”, *System and Computer in Japan*, Vol. J84-D-II, No.12, pp.2585-2594, 2001.

[7] F. Cheng and Y. Chen, “Real time multiple objects tracking and identification based on discrete wavelet transform”, *Journal of the pattern Recognition Society*, Vol. 39, pp 1126-1139, 2006.

[8] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA., 2001.

[9] B. Sugandi, H.S. Kim, J.K. Tan and S. Ishikawa, “Tracking of moving object by using a low resolution image”, *Proc. International Conference on Innovative Computing, Information and Control (ICICIC07)*, CDROM 4 pages, 2007.

[10] C. Stauffer and W. Grimson, “Adaptive background mixture models for real-time tracking”, *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Vol. 2, pp. 246-252, 1999.

[11] S. M. Desa and Q. A. Salih, “Image subtraction for real time moving object extraction”, *Proc. International Conference on Computer Graphics, Imaging and Visualization* (CGIV’04), pp. 41-45, 2004.

[12] E. Stringa, “Morphological change detection algorithms for surveillance applications”, *Proc. British Machine Vision Conf.*, pp. 402-412, 2000.

Figure 4. Tracking the target from two moving persons move in the same direction

Figure 5. Tracking the target from two moving persons move in the different direction

Figure 6. Tracking the target object from three moving person emerging in the scene

Table 1. Object identification results

Experiment	Object detected	Correct identification	Identification rates [%]
1	126	117	92.8
2	148	136	91.9
3	154	141	91.6