

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



A Block Matching Technique for Object Tracking Based on Peripheral Increment Sign Correlation Image

Budi Sugandi, Hyoungseop Kim., Joo Kooi Tan and Seiji Ishikawa
*Graduate School of Engineering, Kyushu Institute of Technology
Japan*

1. Introduction

Automatic detection and tracking of moving object is very important task for human-computer interface (Black & Jepson, 1998), video communication/expression (Menser & Brunig, 2000), and security and surveillance system application (Greiffenhagen et al., 2000) and so on. Various imaging techniques for detection, tracking and identification of the moving objects have been proposed by many researchers. Based on (Collins et al., 2000; Yilmaz, 2006), the object detection can be divided at least into five conventional approaches: frame difference (Lipton et al., 1998; Collins et al., 2000), background subtraction (Heikkila & Silven, 1999; Stauffer & Grimson, 1999; McIvor, 2000; Liu et al., 2001), optical flow (Meyer et al., 1998), skin color extraction (Cho et al., 2001; Phung et al., 2003) and probability based approaches (Harwood et al., 2000; Stauffer & Grimson, 2000; Paragios et al., 2000). Based on (Wu et al., 2004), the object tracking method can be categorized into four categories: region based tracking (Wren et al., 1997; McKenna, 2000), active contour based tracking (Paragois & Deriche, 2000), feature based tracking (Schiele, 2000; Coifman et al., 1998) and model based tracking (Koller, 2000). The object identification is performed to evaluate the effectiveness of the tracking object especially when the object occlusion happens. It can be done by measuring the similarity between the object model and the tracked object. Some of the researches rely on color distribution (Cheng & Chen, 2006; Czyz et al., 2007).

Regarding to our study, many of researchers have their own methods to solve the problem of object detection, object tracking and object identification.

In object detection methodology, many researchers have developed their methods. (Liu et al., 2001) 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. In another work, (Stauffer & Grimson, 1997) proposed a Gaussian mixture model based on background model to detect the object. (Lipton et al., 1998) proposed frame difference that use of the pixel-wise differences between two 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. In order to overcome disadvantage of two-frames differencing, in some cases three-frames differencing is used. For instance, (Collins et al., 2000) developed a hybrid method

that combines three-frame differencing with an adaptive background subtraction model for their VSAM (Video Surveillance and Monitoring) project. The hybrid algorithm successfully segments moving regions in video without the defects of temporal differencing and background subtraction. (Desa & Salih, 2004) proposed a combination of background subtraction and frame difference that improved the previous results of background subtraction and frame difference.

In object tracking methodology, regarding to our study, this article will describe more about the region based tracking. Region-based tracking algorithms track objects according to variations of the image regions corresponding to the moving objects. For these algorithms, the background image is maintained dynamically and motion regions are usually detected by subtracting the background from the current image. (Wren et al., 1997) explored the use of small blob features to track a single human in an indoor environment. In their work, a human body is considered as a combination of some blobs respectively representing various body parts such as head, torso and the four limbs. The pixels belonging to the human body are assigned to the different body part's blobs. By tracking each small blob, the moving human is successfully tracked. (McKenna et al., 2000) proposed an adaptive background subtraction method in which color and gradient information are combined to cope with shadows and unreliable color cues in motion segmentation. Tracking is then performed at three levels of abstraction: regions, people, and groups. Each region has a bounding box and regions can merge and split. A human is composed of one or more regions grouped together under the condition of geometric structure constraints on the human body, and a human group consists of one or more people grouped together.

Moreover, for object identification, (Cheng & Chen, 2006) proposed a color and a spatial feature of the object to identify the track object. The spatial feature is extracted from the bounding box of the object. Meanwhile, the color features extracted is mean and standard value of each object. (Czyz et al., 2007) proposed the color distribution of the object as observation model. The similarity of the object is measure using Bhattacharya distance. The low Bhattacharya distance corresponds to the high similarity.

To overcome the related problem described above, this article proposed a new technique for object detection employing frame difference on low resolution image (Sugandi et al., 2007), object tracking employing block matching algorithm based on PISC image (Satoh et al., 2001) and object identification employing color and spatial information of the tracked object (Cheng & Chen, 2006).

The study in this article includes three main building blocks for building an automated tracking system, which can be listed as object detection, object tracking and object identification as shown in Fig. 1. To cover all section, we organized our article as follows: in section 2, we give a detail of some image pre-processing techniques and filtering process to detect the moving object emerging in the scene. The tracking method is described in section 3 including the block matching algorithm. Object identification based on color and spatial feature is presented in section 4. Experimental results and conclusions are described in section 5 and 6, respectively.

Nearly, every tracking system starts with motion detection. Motion detection aims at separating the corresponding moving objects region from the background image. The first process in the motion detection is capturing the image information using a video camera. The motion detection stage includes some image preprocessing step such as; gray-scaling and smoothing, reducing image resolution using low resolution image technique, frame difference, morphological operation and labeling. The preprocessing steps are applied to

reduce the image noise in order to achieve a higher accuracy of the tracking. The smoothing technique is performed by using median filter. The low resolution image is performed in three successive frames to remove the small or fake motion in the background. Then frame difference is performed on those frames to detect the moving object emerging in the scene. The next process is applying morphological operation such as dilation and erosion as filtering to reduce the noise that is remained in the moving object. Connected component labeling is then performed to label each moving object in different label.

The second stage is tracking the moving object. In this stage, we perform a block matching technique to track only the interest moving 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 that considers the brightness change in all the pixels of the blocks relative to the considered pixel.

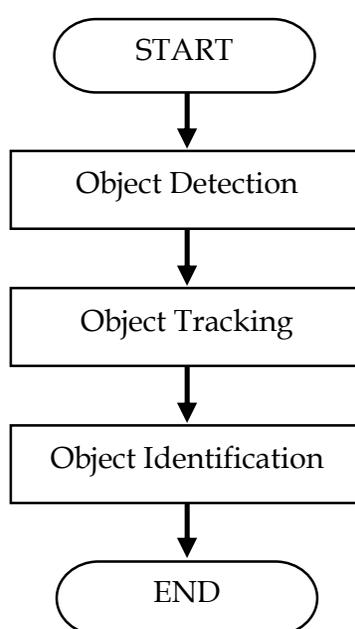


Fig. 1. Entire flow of procedure

The last stage is object identification. For this purpose we use spatial and color information of the tracked object as the image feature (Cheng & Chen, 2006). Then, a feature queue is created to save the features of the moving objects. When the new objects appear on the scene, they will be tracked and labeled, and the features of the object are extracted and recorded into the queue. Once a moving object is detected, the system will extract the features of the object and identify it from the identified objects in the queue.

The details of each stage are described as following.

2. Object detection

Performance of an automated visual surveillance system considerably depends on its ability to detect moving objects in the observed environment. A subsequent action, such as tracking, analyzing the motion or identifying objects, requires an accurate extraction of the foreground objects, making moving object detection a crucial part of the system. In order to decide on whether some regions in a frame are foreground or not, there should be a model for the background intensities. This model should also be able to capture and store

necessary background information. Any change, which is caused by a new object, should be detected by this model, whereas un-stationary background regions, such as branches and leafs of a tree or a flag waving in the wind, should be identified as a part of the background. In this thesis we propose method handle those problem related to un-stationary background such as branches and leafs of a tree by reducing the resolution of the image.

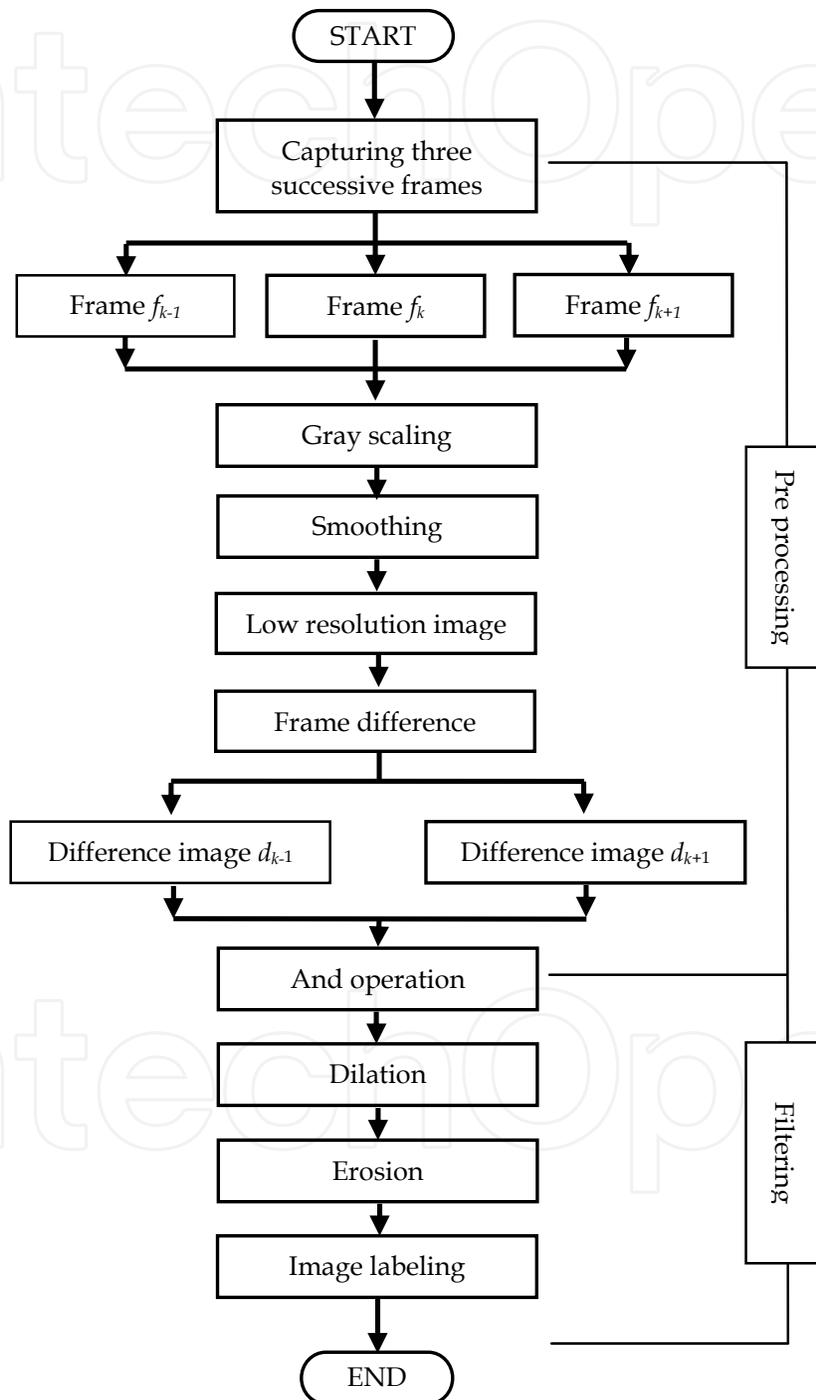


Fig. 2. Flow of object detection

Our object detection method consists of two main steps. The first step is pre-processing step including gray scaling, smoothing, and reducing image resolution and so on. The second

step is filtering to remove the image noise contained in the object. The filtering is performed by applying the morphology filter such as dilation and erosion. And finally connected component labeling is performed on the filtered image.

The entire process of moving object detection is illustrated in Fig. 2.

2.1 Pre-processing

The first step on the moving object detection process is capturing the image information using a video camera. Image is capture by a video camera as 24 bit RGB (red, green, blue) image which each color is specified using 8-bit unsigned integers (0 through 255) that representing the intensities of each color. The size of the captured image is 320x240 pixels. This RGB image is used as input image for the next stage.

In order to reduce the processing time, gray-scale image is used on entire process instead of color image. The gray-scale image only has one color channel that consists of 8 bit while RGB image has three color channels. The color conversion between gray-scale image and RGB image is defined by the following equation:

$$Y = 0.3 \times R + 0.59 \times G + 0.11 \times B \quad (1)$$

where: Y is gray-scale image, R is red, G is green and B is Blue of RGB image, respectively.

Image smoothing is performed to reduce image noise from input image in order to achieve high accuracy for detecting the moving objects. The smoothing process is performed by using a median filter with $m \times m$ pixels.

We consider un-stationary background such as branches and leaf of a tree as part of the background. The un-stationary background often considers as a fake motion other than the motion of the object interest and can cause the failure of detection of the object. To handle this problem, we reduce the resolution of the image to be a low resolution image. A low resolution image is done by reducing spatial resolution of the image with keeping the image size (Gonzales & Woods, 2001) and (Sugandi et al., 2007). In this article, the low resolution image is done by averaging pixels value of its neighbors, including itself. We use a video image with resolution 320x240 pixels. The original image size is 320x240 pixels. After applying the low resolution image, the numbers of pixels will be 160x120, 80x60, or 40x30 pixels, respectively, while the image size is still 320x240 pixels. The low resolution image can be used for reducing the scattering noise and the small fake motion in the background because of un-stationary background such as leaf of a tree. These noises that have small motion region will be disappeared in low resolution image.

To detect the moving object from the background based of image subtraction, generally there are three approaches can be performed: (i) background subtraction as discussed in (Liu et al., 2001), (ii) frame difference as discussed in (Lipton et al., 1998), and (iii) combination of background subtraction and frame difference as discussed in (Desa & Salih, 2004). Background subtraction is computing the difference between the current and the reference background image in a pixel-by-pixel. Frame difference is computing the difference image between the successive frames image. In this article, we applied frame difference method to detect the moving objects. In our case, frame difference method is performed on the three successive frames, which are between frame f_k and f_{k-1} and between frame f_k and f_{k+1} . The output image as frame difference image is two difference images d_{k-1} and d_{k+1} as expressed in Eq. (2). Threshold is performed by threshold value T on the difference image d_{k-1} and d_{k+1} as defined in Eq. (3) to distinguish between the moving object and background.

$$d_{k-1} = |f_k - f_{k-1}| \quad (2)$$

$$d_{k+1} = |f_k - f_{k+1}|$$

$$d'_k(x, y) = \begin{cases} 1, & \text{if } d'_k(x, y) > T \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

where $k' = k - 1$ and $k + 1$.

The process is followed by applying AND operator to d_{k-1} and d_{k+1} as expressed in Eq. (4). The output image of this operation is named as motion mask m_p .

$$m_p = d_{k-1} \cap d_{k+1} \quad (4)$$

The comparison of moving object detection using a conventional method (frame different on normal resolution) and frame different method on low resolution image is shown in Fig 3. On those figures, we use same threshold to determine the object. As shown in those figures, using the conventional method the detected moving object is still greatly affected by small noise such as moving leaves. In the other hand, by reducing the resolution of the image before taking the difference frame, that kind of noise can be removed.

2.2 Filtering

In order to fuse narrow breaks and long thin gulfs, eliminates small holes, and fills gaps in the contour, a morphological operation is applied to the image.

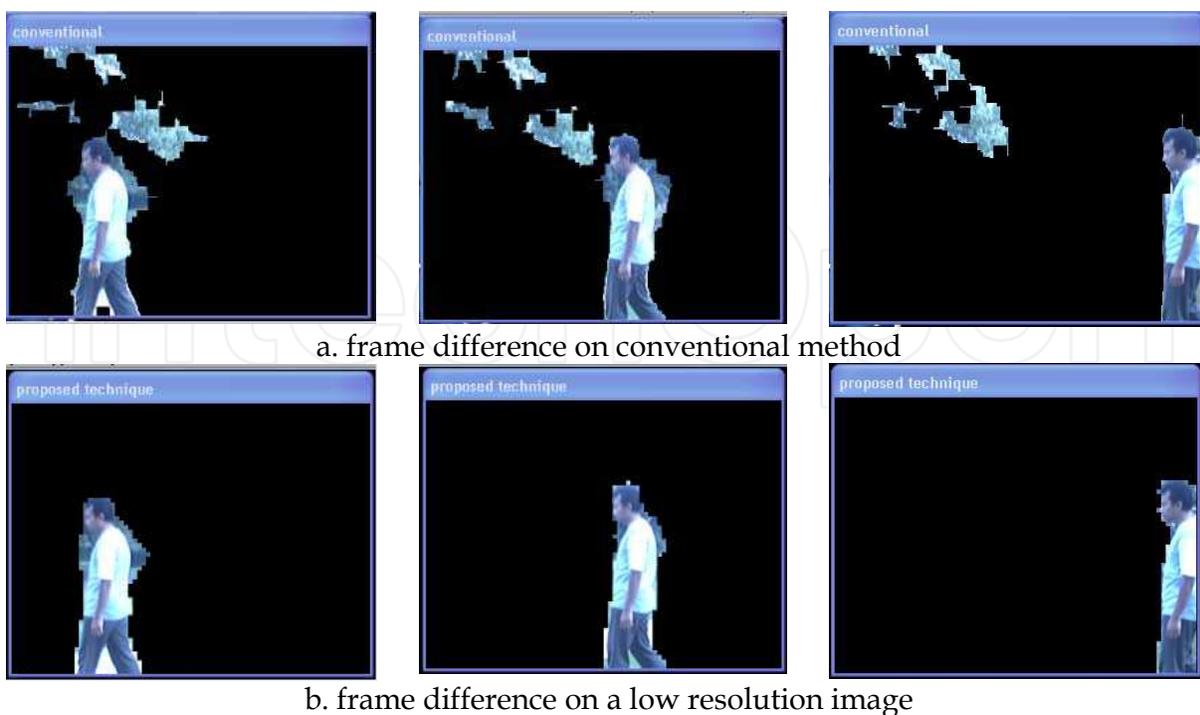


Fig. 3. Comparison of moving object detection technique

As a result, small gaps between the isolated segments are erased and the regions are merged. To extract the bounding boxes of detected objects, connected component analysis was used. We find all contours in image and draw the rectangles around corresponding contours with minimum area. Since the image may contain regions which are composed of background noise pixels and these regions are smaller than actual motion regions, we discard the region with a smaller area than the predefined threshold. As a result, the processing produces perfect bounding boxes.

Morphological operation is performed to fill small gaps inside the moving object and to reduce the noise remained in the moving objects (Stringa, 2000). The morphological operators implemented are dilation followed by erosion. In dilation, each background pixel that is touching an object pixel is changed into an object pixel. Dilation adds pixels to the boundary of the object and closes isolated background pixel. Dilation can be expressed as:

$$f(x,y) = \begin{cases} 1, & \text{if there is one or more pixels of the 8 neighbors are 1} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

In erosion, each object pixel that is touching a background pixel is changed into a background pixel. Erosion removes isolated foreground pixels. Erosion can be expressed as:

$$f(x,y) = \begin{cases} 0, & \text{if there is one or more pixels of the 8 neighbors are 0} \\ 1, & \text{otherwise} \end{cases} \quad (6)$$

Morphological operation eliminates background noise and fills small gaps inside an object. This property makes it well suited to our objective since we are interested in generating masks which preserve the object boundary.

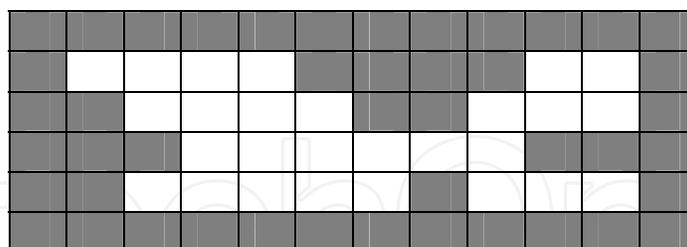


Fig. 4. Binary image

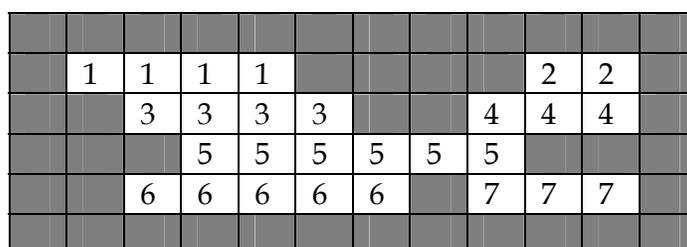


Fig. 5. Image is labeled in the same row

	1	1	1	1					2	2	
		1	1	1	1			2	2	2	
			1	1	1	1	1	1			
		1	1	1	1	1		1	1	1	

Fig. 6. Labeled image is scanned from top-left to bottom-right

	1	1	1	1					1	1	
		1	1	1	1			1	1	1	
			1	1	1	1	1	1			
		1	1	1	1	1		1	1	1	

Fig. 7. Final labeled image

Connected component labeling is performed to label each moving object emerging in the background. The connected component labeling (Gonzales & Woods, 2001) groups the pixels into components based on pixel connectivity (same intensity or gray level). In this article, connected component labeling is done by comparing the pixel with the pixel in four neighbors. If the pixel has at least one neighbor with the same label, this pixel is labeled as same as neighbor's label. The algorithm of connected component labeling algorithm is described as follows:

1. Firstly, image labeling is done on binary image as shown in Fig. 4 where object is shown as 1 (white) and background is shown as 0 (black).
2. The image is scanned from top-left to search the object pixel. The label is done by scanning the image from left to right and comparing label with the neighbor's label in the same line. If the neighbor has the same pixel value, the pixel is labeled as same as previous label as shown in Figure. 5.
3. Next, the labeled image is scanned from top-left to bottom-right by comparing with the four (or eight) neighbors pixel which have already been encountered in the scan (the neighbors (i) to the left of the pixel, (ii) above it, and (iii and iv) the two upper diagonal terms). If the pixel has at least one neighbor, then this pixel is labeled as same as neighbor's label as shown in Fig. 6.
4. On the last scanning, the image is scanned from bottom-right to top-left by comparing with the four neighbors pixel as step 3. The final labeled image is shown in Fig. 7.

3. Object tracking

After the object detection is achieved, the problem of establishing a correspondence between object masks in consecutive frames should arise. Indeed, initializing a track, updating it robustly and ending the track are important problems of object mask association during tracking. Obtaining the correct track information is crucial for subsequent actions, such as object identification and activity recognition. Tracking process can be considered as a region mask association between temporally consecutive frames and estimating the trajectory of an object in the image plane as it moves around a scene. In this article, we use block matching

technique for this purpose. The entire process of tracking the moving object is illustrated in Fig 8. The details of the tracking mechanism of the interest moving object are described in the following sections.

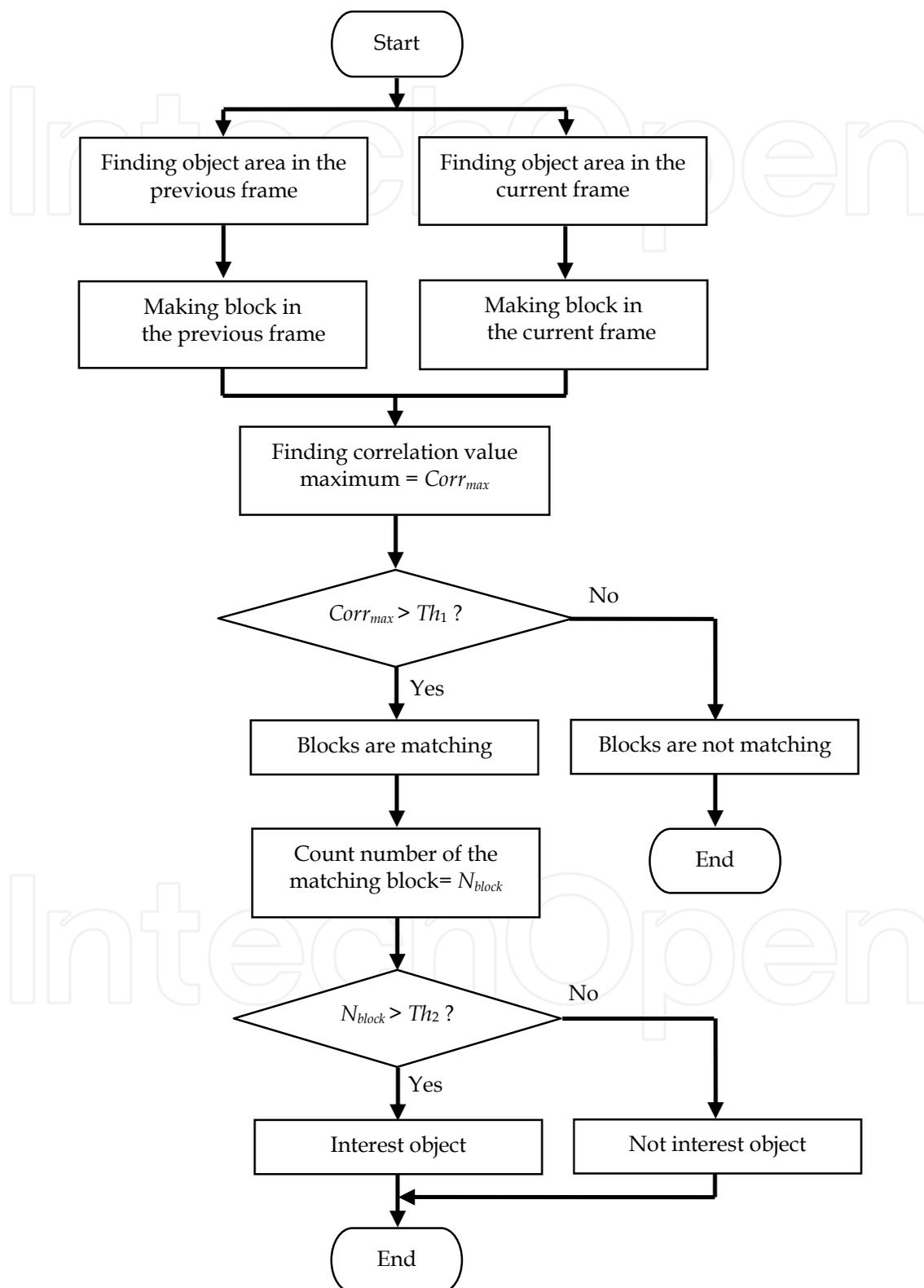


Fig. 8. Entire flow of tracking of interest object

3.1 Block matching method

Block matching is a technique for tracking the interest moving object among the moving objects emerging in the scene. In this article, the blocks are defined by dividing the image frame into non-overlapping square parts. The blocks are made based on peripheral increment sign correlation (PISC) image (Sato et al., 2001; Sugandi et al., 2007) that considers the brightness change in all the pixels of the blocks relative to the considered pixel. Fig. 9 shows the block in PISC image with block size is 5×5 pixels. Therefore, one block consists of 25 pixels. The blocks of the PISC image in the previous frame are defined as shown in Eq. (7). Similarly, the blocks of the PISC image in the current frame are defined in Eq. (8). To determine the matching criteria of the blocks in two successive frames, we evaluate using correlation value that expresses in Eq. (9). This equation calculates the correlation value between block in the previous frame and the current one for all pixels in the block. The high correlation value shows that the blocks are matched each other. The interest moving object is determined when the number of matching blocks in the previous and current frame are higher than the certain threshold value. The threshold value is obtained experimentally.

$$b_{np} = \begin{cases} 1, & \text{if } f_{np} \geq f(i, j) \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

$$b'_{np} = \begin{cases} 1, & \text{if } f_{np} \geq f(i, j) \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

$$corr_n = \sum_{p=0}^N b_{np} * b'_{np} + \sum_{p=0}^N (1 - b_{np}) * (1 - b'_{np}) \quad (9)$$

where : b and b' are the block in previous and current frame, n is number of block and N is number of pixels of block, respectively.

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

Fig. 9. PISC image for block size 5×5 pixels

3.2 Tracking method

The tracking method used in this article can be described as following. The matching process is illustrated in Fig. 10. Firstly, blocks and the tracking area are made only in the

area of moving object to reduce the processing time. We make the block size (block A) with 9×9 pixels in the previous frame. We assume that the object coming firstly will be tracked as the interest moving object. The block A will search the matching block in each block of the current frame by using correlation value as expresses in Eq. (9). In the current frame, the interest moving object is tracked when the object has maximum number of matching blocks. When that matching criteria is not satisfied, the matching process is repeated by enlarging the tracking area (the rectangle with dash line). The blocks still are made inside the area of moving object. When the interest moving object still cannot be tracked, then the moving object is categorized as not interest moving object or another object and the tracking process is begun again from the beginning.

4. Object identification

The object identification is the last stage of our study. In this stage, firstly, we will explain about features extraction for object detection (Cheng & Chen, 2006). In this article, the extracted features are divided into the following two types; color and spatial information of the moving objects.

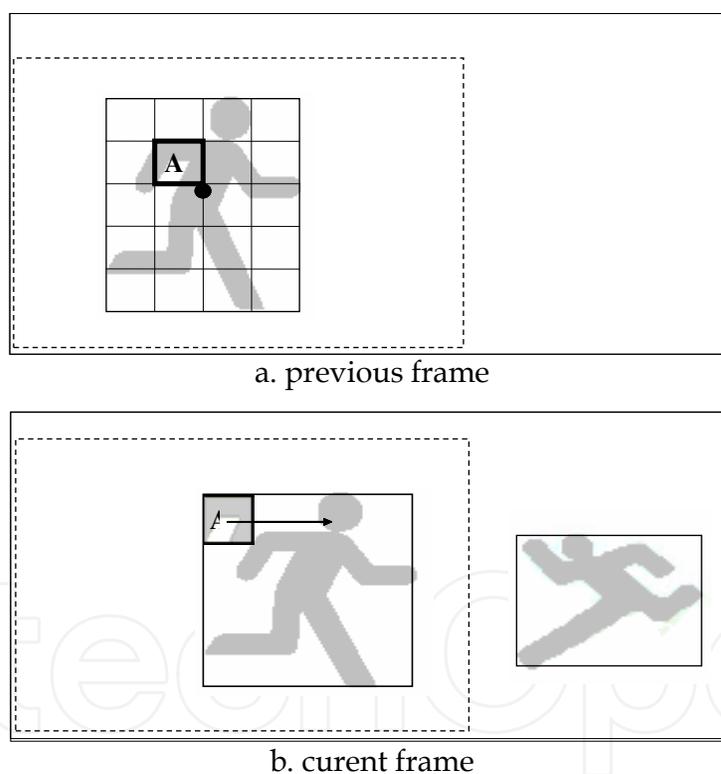


Fig. 10. Matching process

4.1 Spatial feature extraction

The feature of objects extracted in the spatial domain is the position of the tracked object. The spatial information combined with the features in time domain represents the trajectory of the tracked object, so we can estimate the movement and speed of the moving objects that we tracked. Therefore, the features of spatial domain are very important to object identification. The bounding box defined in Eq. (10) is used as spatial information of moving objects.

After getting the interest moving object, we extract the interest moving object by using a bounding box. The bounding box can be determined by computing the maximum and minimum value of x and y coordinates of the interest moving object according to the following equation:

$$B_{\min}^i = \left\{ \left(x_{\min}^i, y_{\min}^i \right) \mid x, y \in O^i \right\}$$

$$B_{\max}^i = \left\{ \left(x_{\max}^i, y_{\max}^i \right) \mid x, y \in O^i \right\} \quad (10)$$

where O^i denotes the set of the coordinate of points in the interest moving object i , B_{\min}^i is the left-top corner coordinates of the interest moving object i and B_{\max}^i is the right-bottom corner coordinates of the interest moving object i , respectively. Fig. 11 shows an example of the bounding box of the object tracking.

4.2 Color feature extraction

The color feature extracted from the object is RGB color space as the RGB color information can be obtained from video capture device directly. We extract the information from upper

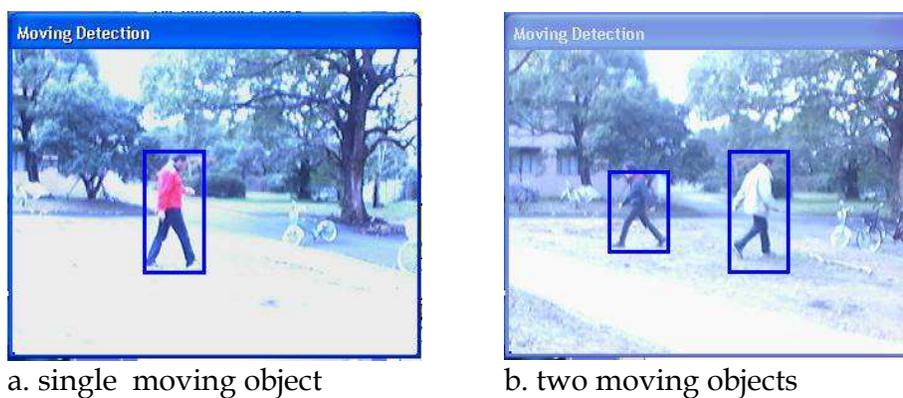


Fig. 11. Example of bounding box of moving object

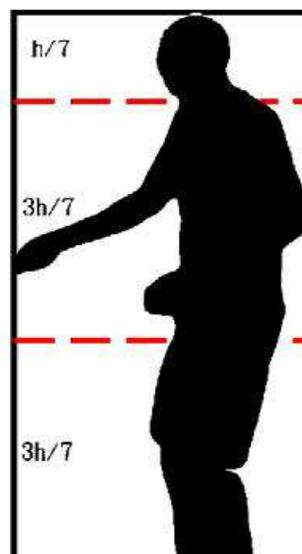


Fig. 12. Definition of a human body ratio

and lower part of the object to obtain more color information for identification. The ratio of these three parts can be defined as shown in Fig. 12. However, in this article, we only calculate the color information of the upper and lower part excluding the head part of the object. The first color information calculated is mean value of each human body part as calculated by Eq. (11) for upper part and Eq. (12) for lower part. The mean value is calculated for each color component of RGB space.

$$\mu_{f_k}^{O_U^i} = \frac{\sum_{x=x_{\min}^i}^{x_{\max}^i} \sum_{y=y_{\min}^i}^{y_{\max}^i} f_k(x, y)}{\#O_U^i} \quad (11)$$

$$\mu_{f_k}^{O_L^i} = \frac{\sum_{x=x_{\min}^i}^{x_{\max}^i} \sum_{y=y_{\min}^i}^{y_{\max}^i} f_k(x, y)}{\#O_L^i} \quad (12)$$

where i is number of the moving objects and (x, y) is the coordinate of pixels in moving object. (x_{\max}^i, y_{\max}^i) and (x_{\min}^i, y_{\min}^i) are the coordinates of the bounding box of moving object i , $f_k(x, y)$ denotes pixel value for each color component in RGB space of the current frame, O_U^i and O_L^i denote the set of coordinates of upper and lower part of human body of moving object i and $\#O_i$ is the number of pixels of moving object i .

The standard deviation has proven to be an extremely useful measure of spread in part because it is mathematically tractable. Standard deviation is a statistical term that provides a good indication of volatility. It measures how widely the values are dispersed from the average. Dispersion is the difference between the actual value and the average value. The larger the difference between the actual color and the average color is, the higher the standard deviation will be, and the higher the volatility. We can extract more useful color features by computing the dispersed color information from upper and lower part of body as shown in Eq. (13) for the upper part and Eq. (14) for the lower part.

$$SD_{f_k}^{O_U^i} = \sqrt{\frac{\sum_{x=x_{\min}^i}^{x_{\max}^i} \sum_{y=y_{\min}^i}^{y_{\max}^i} (f_k(x, y) - \mu_{f_k}^{O_U^i})^2}{\#O_U^i}} \quad (13)$$

$$SD_{f_k}^{O_L^i} = \sqrt{\frac{\sum_{x=x_{\min}^i}^{x_{\max}^i} \sum_{y=y_{\min}^i}^{y_{\max}^i} (f_k(x, y) - \mu_{f_k}^{O_L^i})^2}{\#O_L^i}} \quad (14)$$

where $SD_{f_k}^{O_U^i}$ and $SD_{f_k}^{O_L^i}$ denote the standard deviation of each color component of RGB space for upper and lower part of the human body of moving object i , respectively.

4.3 Identification process

After the feature extraction, we can represent the moving object i by the following feature vectors;

$$F^i = \left(\mu_{f_k}^{O_U^i}, \mu_{f_k}^{O_L^i}, SD_{f_k}^{O_U^i}, SD_{f_k}^{O_L^i}, B_{\min}^i, B_{\max}^i \right) \quad (15)$$

To identify a moving object, a feature queue is created to save the features of the moving objects. When a new object enters the system, it will be tracked and labeled, and the features of the object are extracted and recorded into the queue. Once a moving object i is detected, the system will extract the features F^i of the object i and identify it from the identified objects in the queue by computing the similarity $S(F^i, F^j)$, $j=1 \dots n$, where j is one of the n identified objects. The similarity, $S(F^i, F^j)$ is computed as in Eq. (16),

$$\begin{aligned}
 S(F^i, F^j) = & Mc\left(\left|\mu_{f_k}^{O_u^i} - \mu_{f_k}^{O_u^j}\right|\right) + Mc\left(\left|\mu_{f_k}^{O_L^i} - \mu_{f_k}^{O_L^j}\right|\right) \\
 & + Msd\left(\left|SD_{f_k}^{O_u^i} - SD_{f_k}^{O_u^j}\right|\right) + Msd\left(\left|SD_{f_k}^{O_L^i} - SD_{f_k}^{O_L^j}\right|\right) \\
 & + 0.5Mp\left(\left|B_{\min}^i - B_{\min}^j\right|\right) + 0.5Mp\left(\left|B_{\max}^i - B_{\max}^j\right|\right)
 \end{aligned} \tag{16}$$

where Mc and Msd are the membership function of color information and standard deviation as defined in Eq. (17) and Eq. (18), Mp is the membership function of spatial information as defined in Eq. (19).

$$Mc(x) = \begin{cases} 1 - x / Thr & \text{if } x < Thr \\ 0 & \text{if } x \geq Thr \end{cases} \tag{17}$$

$$Msd(x) = \begin{cases} 1 - x / Thr & \text{if } x < Thr \\ 0 & \text{if } x \geq Thr \end{cases} \tag{18}$$

$$Mp(x) = \begin{cases} 1 - 3x / W & \text{if } x < W / 3 \\ 0 & \text{if } x \geq W / 3 \end{cases} \tag{19}$$

where Thr is threshold which is obtained experimentally and W is the width or height of the image frames. We compare the features of detected object with those of the objects in the feature queue. The one with the maximum similarity is identified as the same object. The whole object identification flow is shown in Fig. 13.

5. Experimental results

We have done the experiments using a video camera in outdoor environment and real time condition. The experiment is performed in 2.54 [GHz] Pentium 4 PC with 512 MB memory. The image resolution is 320×240 [pixels]. The size of each block is 9×9 [pixels]. The experimental results are shown in Fig. 14–Fig. 16. The rectangle area on the object shows the tracked object. The identification result is shown in Table 1. In the experimental results, we can extract the moving objects on the successive frame successfully and identification rates of 92.8% were achieved.

In our experiment, we tracked the interest object from two and three moving objects that occluded between each other when objects move in the same and different direction. We assumed that the first moving object emerging in the scene as the interest moving object. We did the experiments in three conditions. In the first condition, we tracked the interest

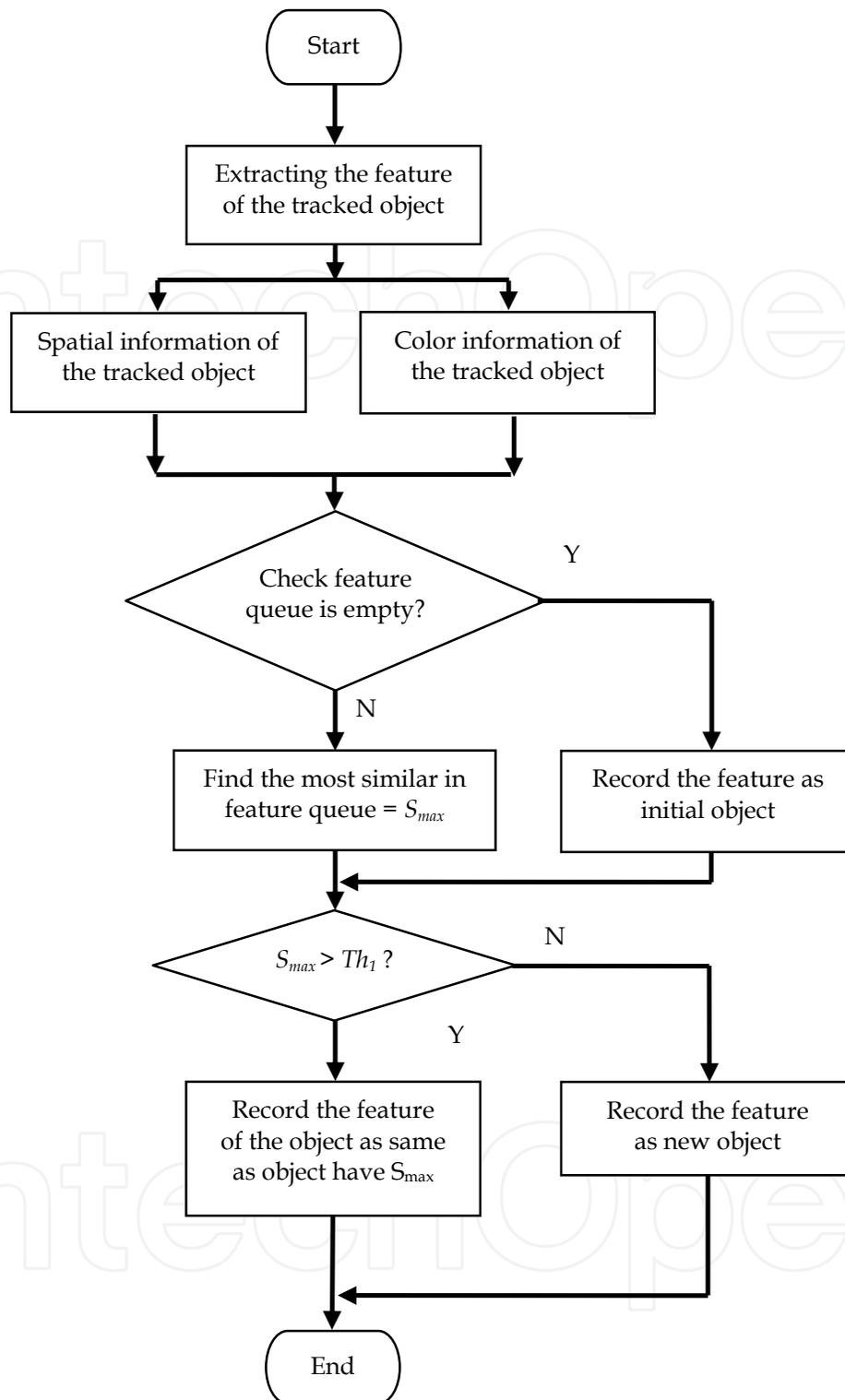


Fig. 13. Flow chart of object identification

moving object when it is occluded by another moving object when they move in the same direction as shown in Fig. 14. In the second condition, we tracked the interest moving object when it is occluded by another moving object when they move in different direction as shown in Fig. 15. In the third condition, we tracked the interest moving object among three moving object appear in the scene as shown in Fig. 16.

On the first case (Fig. 14), at first, the man wearing the white shirt enters the scene from the left side. This object is successfully detected as the interest moving object. While the first object is being tracked, another object (man wearing the blue shirt) enters the scene from the right side. They move in the different direction and overlap each other in the middle of the scene. We successfully track the first moving object as the interest moving object as our assumptions while the other moving object is not tracked.



Fig. 14. Two moving objects move in different direction

On the second case (Fig. 15), at first, the man wearing the blue shirt enters the scene from the left side. This object is successfully detected as the interest moving object. Then on the next



Fig. 15. Two moving objects move in same direction

frame, the man wearing the white shirt enters the scene from the same side. They move in the same direction and occlude each other in the middle of the scene. We successfully track the first moving object as the interest moving object as our assumptions while the other moving object is not tracked.



Fig. 16. Three moving objects occlude in the middle of the scene

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

Table 1. Object identification results



a. an object entering scene b. an object moves slow down

Fig. 17. Error in object identification

On the third case (Fig.16), at first, the man wearing the white shirt enters the scene from the right side. This person will be tracked as the interest moving object. Then on the next frame, another man enters the scene from right side. They move in the same direction. The third

man enters the scene from the left side. They occlude each other in the middle on the scene. We successfully track the first moving object as interest moving object as our assumptions while the other moving object is not tracked.

To evaluate our method, we performed the object identification method based on spatial and color information of the tracked object. Table 1 shows the result of the object identification. From the table, we notice that some objects are not correctly identified in some frames of each experiment. The wrong identification occurs in two types. First, it occurs when the moving object just enters or leaves the scene. When the moving object is just entering the scene, the system can detect the moving object. However, the extracted features of the moving object in this case cannot represent the moving object very well, because only partial of the features of the moving objects are extracted. The second error occurs when the moving object is slowing down. In this situation, the frame difference of the object becomes smaller. Therefore, only smaller bounding box and less moving pixels are obtained. On this condition, the extracted features will lose its representative. Both of errors are illustrated in Fig. 17.

6. Conclusion

This article proposed a new method for detecting the moving object employing frame difference on low resolution image, tracking the moving object employing block matching technique based on peripheral increment sign correlation image for tracking the interest moving object among the moving objects emerging in the background and identifying the moving objects employing color and spatial information of the tracked object. The experiment results and data show the effectiveness and the satisfaction of the proposed methods. Using our method, we can achieve the identification rate of 92.1[%] in average.

However, the proposed method still has limitations. The limitations can be investigated as followings. Firstly, the detection method based on frame difference on low resolution image has a limitation when the moving object is too small to be detected. It is occurred because the low resolution image removes the small moving objects emerging in the background. To overcome this limitation, we can add another method such as skin color detection. By using this method, even if the moving object is too small, it can still be detected based on the skin color of the object. Secondly, the block matching technique has successfully tracked the interest moving object in the occlude condition. However, when the moving objects appear in the same time, we cannot judge any object to be an interest object. Moreover, when the interest moving object is covered by the occluded object, the image information of the interest moving object cannot be read by the camera. This condition cause the system cannot recognize the interest moving object. Those limitations can be solved by adding other information to the interest moving object such as flow of moving object based on optical flow, dimension or another feature and also we can add the color information to each object. So whenever the objects appear, they have their own model that different from each other. And we can track them based on the model.

Thirdly, color and spatial information method show the high correct identification rate. However, the system still cannot identify the objects sometimes when they are just entering or leaving the scene. The extracted features in this case are not enough to be used to identify the moving objects. The system also has limitation when the object is moving slowly. In this condition, the inter-frame difference image of the object will become smaller and we will get smaller bounding box and less moving pixels. Therefore, the extracted features will lose its

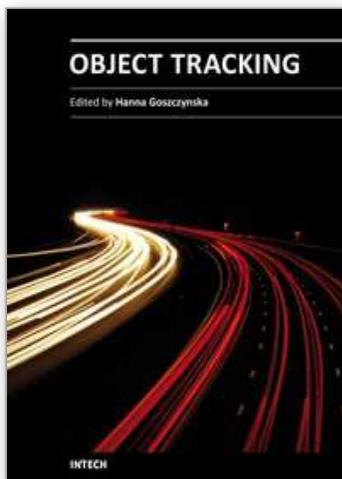
representative. The correct identification rate highly depends on the correctness of the moving object detection and feature representation. This problem can be improved by a better feature selection method and moving object detection method.

By considering those limitations and implement some improvements to our method including speed up the processing time, they could lead to some improvements in the tracking system. These are remaining for future work.

7. References

- Lipton, A; Fujiyoshi, H. & Patil, R.(1998) .Moving target classification and tracking from real-time video, *Proceeding of IEEE Workshop Applications of Computer Vision*, pp. 8-14.
- Stauffer, C. & Grimson, W. (1999). Adaptive background mixture models for real-time tracking, *Proceeding of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 246-252.
- LIU, Y.; Haizho, A. & Xu Guangyou.(2001). Moving object detection and tracking based on background subtraction, *Proceeding of Society of Photo-Optical Instrument Engineers (SPIE)*, Vol. 4554, pp. 62-66.
- Cheng, F. & Chen, Y. (2006). Real time multiple objects tracking and identification based on discrete wavelet transform, *Journal of the pattern Recognition Society*, pp 1126-1139.
- Davis, J. W. & Bobick, A. F. (1997). The representation and recognition of action using temporal templates, *Proceeding of Comp Vision and Pattern Recognition*, pp. 928-934.
- Meyer, D.; Denzler, J. & Niemann, H. (1998). Model based extraction of articulated objects in image sequences for gait analysis, *Proceeding of IEEE Int. Conf. Image Processing*, pp. 78-81.
- Paragios, N. & Deriche, R. (2000). Geodesic active contours and level sets for the detection and tracking of moving object, *IEEE Trans. on Pattern Analysis and Machine Intelligent*, pp. 266-280.
- Satoh, Y.; Kaneko, S. & Igarashi, S. (2001). Robust Object Detection and Segmentation by Peripheral Increment Sign Correlation Image, *System and Computer in Japan*, pp.2585-2594.
- Gonzalez, R. C. & Woods, R. E. (2001). *Digital Image Processing*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA.
- Sugandi, B.; Kim, H.S.; Tan, J.K. & Ishikawa, S. (2007). Tracking of moving object by using low resolution image, *Proceeding of Int. Conf. on Innovative Computing, Information and Control (ICICIC07)*, Kumamoto, Japan, (4 pages).
- Desa, S. M. & Salih, Q. A. (2004). Image subtraction for real time moving object extraction, *Proceeding of Int. Conf. on Computer Graphics, Imaging and Visualization (CGIV'04)*, pp. 41-45.
- Stringa, E. (200). Morphological change detection algorithms for surveillance applications, *Proceeding of British Machine Vision Conf.*, pp. 402-412.
- Sugandi, B.; Kim, H.S.; Tan, J.K. & Ishikawa, S. (2007).Tracking of moving persons using multi camera employing peripheral increment sign correlation image, *ICIC Express Letters, An International Journal of Research and Surveys*, Vol. 1, No. 2, pp. 177-184.
- Black, M. & Jepson, A. (1998). A probabilistic framework for matching temporal trajectories: condensation-based recognition of gestures and expressions, *Proceeding of the European Conference on Computer Vision*, pp. 909-924.

- Menser, B. & Brunig, V. (2000). Face detection and tracking for video coding applications, *Asilomar Conference on Signals, Systems, and Computers*, pp. 49-53.
- Greiffenhagen, M.; Ramesh, V.; Comaniciu, D. & Niemann, H. (2000). Statistical modeling and performance characterization of a real-time dual camera surveillance system, *Proceeding of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 335-342.
- McIvor, A. M. (2000). Background subtraction techniques, *Proceeding of Image and Vision Computing*, 6 pages.
- Heikkila, J. & Silven, O. (1999). A real-time system for monitoring of cyclists and pedestrians, *Proceeding of Second IEEE Workshop on Visual Surveillance*, pp 74-81.
- Collins, R. ; Lipton, A.; Kanade, T.; Fujiyoshi, H.; Duggins, D.; Tsin, Y.; Tolliver, D.; Enomoto, N. & Hasegawa. (2000). System for video surveillance and monitoring, *Technical Report CMU-RI-TR-00-12*, Robotics Institute, Carnegie Mellon University.
- Meyer, D.; Denzler, J. & Niemann, H. (1998). Model based extraction of articulated objects in image sequences for gait analysis, *Proceeding of the IEEE Int. Conf. on Image Processing*, pp. 78-81.
- Phung, S.; Chai, D. & Bouzerdoum, A. (2003). Adaptive skin segmentation in color images, *Proceeding of IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 3, pp. 353-356.
- Cho, K. M.; Jang, J. H. & Hong, K. S. (2001). Adaptive skin-color filter, *Pattern Recognition*, pp. 1067-1073.
- Harwood, D. ; Haritaoglu, I. & Davis, L. S. (2000). W4: Real-time surveillance of people and their activities, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 8, pp. 809-830.
- Stauffer, C. & Grimson, W. (2000). Learning patterns of activity using real-time tracking, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 747-757.
- Hu, W. ; Tan, T.; Wang, L. & Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors, *IEEE Trans. On System, Man, and Cybernetics*, Vol. 34, No. 3, pp. 334-352.
- Wren, C. R.; Azarbayejani, A.; Darrell, T. & Pentland, A. P. (1997). Pfunder: real-time tracking of the human body, *IEEE Trans. Pattern Analysis. Machine Inteligent.*, Vol. 19, pp. 780-785.
- McKenna, S.; Jabri, S.; Duric, Z.; Rosenfeld, A. & Wechsler, H. (2000). Tracking groups of people, *Computer Vision: Image Understanding*, Vol. 80, No. 1, pp. 42-56.
- Paragios, N. & Deriche, R. (2000). Geodesic active contours and level sets for the detection and tracking of moving objects, *IEEE Trans. Pattern Analysis and Machine Intelligent*, Vol. 22, pp. 266-280.
- Schiele, B. (2000). Voxel-free tracking of cars and people based on color regions, *Proceeding of IEEE Int. Workshop Performance Evaluation of Tracking and Surveillance*, Grenoble, France, pp. 61-71.
- Coifman, B.; Beymer, D.; McLauchlan, P. & Malik, J. (1998). A real-time computer vision system for vehicle tracking and traffic surveillance, *Transportation Research: Part C*, Vol. 6, No. 4, pp. 271-288.
- Koller, D.; Danilidis, K. & Nagel, H. (2000). Model-based object tracking in monocular image sequences of road traffic scenes, *Int. Journal of Computer Vision*, Vol.10, No.3, pp.257-281.
- Czyz, J.; Ristic, B. & Macq, B. (2007). A particle filter for joint detection and tracking of color objects", *Image and Vision Computing*, Vol. 25, No. 8, pp 1271-1281.
- Yilmaz, A.; Javed, O. & Shah, M. (2006). Object tracking: a survey, *ACM Computing Survey*, Vol.38, No.13, pp.1-45.



Object Tracking

Edited by Dr. Hanna Goszczynska

ISBN 978-953-307-360-6

Hard cover, 284 pages

Publisher InTech

Published online 28, February, 2011

Published in print edition February, 2011

Object tracking consists in estimation of trajectory of moving objects in the sequence of images. Automation of the computer object tracking is a difficult task. Dynamics of multiple parameters changes representing features and motion of the objects, and temporary partial or full occlusion of the tracked objects have to be considered. This monograph presents the development of object tracking algorithms, methods and systems. Both, state of the art of object tracking methods and also the new trends in research are described in this book. Fourteen chapters are split into two sections. Section 1 presents new theoretical ideas whereas Section 2 presents real-life applications. Despite the variety of topics contained in this monograph it constitutes a consisted knowledge in the field of computer object tracking. The intention of editor was to follow up the very quick progress in the developing of methods as well as extension of the application.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Budi Sugandi, Hyoungseop Kim., Joo Kooi Tan and Seiji Ishikawa (2011). A Block Matching Technique for Object Tracking Based on Peripheral Increment Sign Correlation Image, Object Tracking, Dr. Hanna Goszczynska (Ed.), ISBN: 978-953-307-360-6, InTech, Available from:
<http://www.intechopen.com/books/object-tracking/a-block-matching-technique-for-object-tracking-based-on-peripheral-increment-sign-correlation-image>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen