MVA'90 IAPR Workshop on Machine Vision Applications Nov. 28-30,1990, Tokyo

A STUDY OF MOVING OBJECT DETECTION FROM UNSTABLE MOTION PICTURE

Hirofumi ISHII, Toshiyuki KAWARA, Isao KAWAHARA, Makoto FUJIMOTO
Central Research Laboratories,
Matsushita Electric Industrial Co., Ltd.

Yagumo-nakamachi, Moriguchi,
Osaka, 570, Japan

SUMMARY

A new method to detect and track an object having a unique movement out of general moving background comprised of a number of moving areas is reported here. After an overall picture movement is detected and normalized from local movements of an image, the moving object becomes detectable by applying a time sequential processing on the difference between frames. Thus, the detection of moving object out of a moving background can be carried out highly efficiently. As an example of applications, a detection of man walking in a panned background is demonstrated in this report.

1. INTRODUCTION

A moving object in a still background scenery can be detected easily by subjecting its difference between frames to a simple time sequential processing. Although these methods of detection are applicable to moving objects in background sceneries under a very limited condition, these have been used for practical purposes including security.

However, detection of a moving object existing in a moving background are generally very difficult to accomplish because moving backgrounds contain complicated multiple moving image regions usually, and the conventional detection methods are not applicable to these anymore.

Whereas another method wherein the detection of moving object is carried out for entire pixel units of an image, and based on this result, a clustering of these pixels is carried out, is already known, highly complicated and enormous computations have to be carried out. Therefore, this method has very little practical use.

A new method to detect a moving object having a unique movement out of a moving background which is divided into plural moving areas is presently developed. With this method, an overall image movement is detected out of local image movements before this is normalized.

After this normalization, the moving object having unnatural moving characteristics can be detected by applying a time sequential processing based on the difference between frames, and this enables an efficient detection of moving object. As an example of this, detection of a man walking in a panned background is demonstrated.

2. SYSTEM CONSTRUCTION

Figs. 1(a), 1(b) and 1(c) show positional plots of spatial characteristics of images in various states on the horizontal axis against time taken along the vertical axis. Fig.1(a) shows a case of still picture, wherein the detection of a moving object is possible by tracking the image regions where image signals are changing.

When a movement of background is produced by camera panning, for example, changes of image signals are produced over an entire picture frame as shown in Fig.1(b). Therefore, the above mentioned detection methods can not be applied.

However, by detecting a unified motion vector of background and by normalizing the background accordingly as shown in Fig. 1(c), the moving object becomes detectable by tracking only the image regions where the image signals are changing.

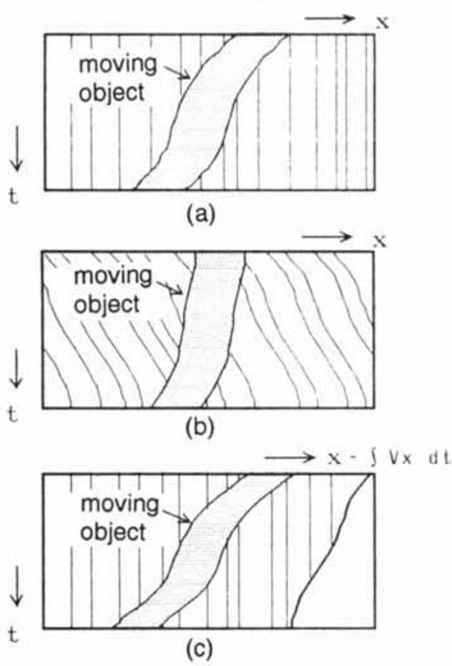


Fig. 1 Still picture and normalizing process

Fig. 2 shows a block diagram of circuit construction wherein an image signal is parallely input to a motion vector detecting block and a moving region detecting block. In the motion vector detecting block, after a proper filtering process, a

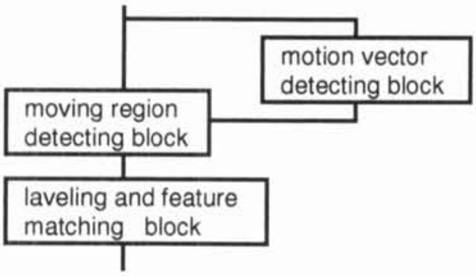


Fig. 2 Block diagram

motion vector between frames (an averaged parallel shift of background) is detected by applying a representative point matching method.

In the moving region detecting block, after relative positions between frames are adjusted by using detected motion vector, the moving regions including moving object are detected by applying a time-sequential processing to the difference between frames.

In the labeling and feature matching block, labeling of plural moving image regions produced by moving objects is carried out when plural moving objects exist in a background. An identification and tracking process of each moving object is then conducted according to the image characteristics such as color and size and location associated with each of the moving image regions.

By conducting the operations of above, moving objects can be detected out of a background which is in a parallel shift movement.

3. DETECTION OF MOTION VECTOR OF BACKGROUND

The motion vector of entire background can be detected by using a representative point matching method. The necessary number of representative points depends directly on the number of computations required for the processing, scale of the circuit, and the accuracy of detection.

Although an adequate detection accuracy could be secured if representative points of several thousand were used, this is highly impractical from a standpoint of circuit scale. An adequate accuracy, however, is realized by using fairly smaller number of representative points, or approximately thirty points by applying a proper filtering in the preprocess[1]) for improving its S/N.

Fig.3 shows a relationship between S/N and standard deviation of motion vectors, and this shows an adequate detection accuracy can be secured at a region where S/N is more than 32 dB and the standard deviation is less than 0.5 pixel.

In order to make the accurate detection of motion vector of background possible, the effects of a moving object to the background have to be care-ully examined. Thus, a picture frame is divided into several image regions, and the motion vectors are derived there-from. The image regions affected by a moving object are excluded next, and the motion vector of the background is detected by using the motion vectors derived from the remaining image regions.

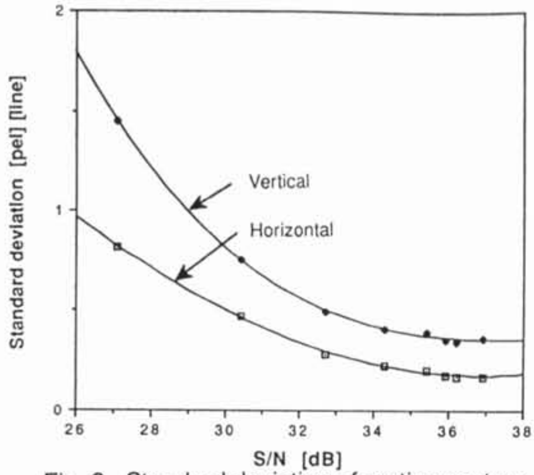


Fig. 3 Standard deviation of motion vectors

The reliability of motion vector is examined in terms of shape of a correlation function obtained by a representative point matching method. Fig. 4 shows a change of correlation measure against the coordinate values. As shown by Curve (a), a normal correlation measure shows a sharp minimum dip point, and the position of this minimum point gives a motion vector.

On the other hand, in a case where a moving object exists in a moving background, the correlation function is represented by Curve (c), and the detected motion vector is affected by a moving object. Beside this, when the image signal level is low, that is, the image pattern signal level is low, or when the change of background takes place along only one direction, the accuracy of motion vector detection is lower. In order to differentiate such conditions, provision of conditions shown below can be effectively used.

$$\text{Average value} > \alpha \quad (1)$$

$$\text{Average value - minimum value} \times \beta > 0 \quad (2)$$

$$\text{Gradient} > \gamma \quad (3)$$

wherein β is a predetermined factor, and α,γ are predetermined threshold values.

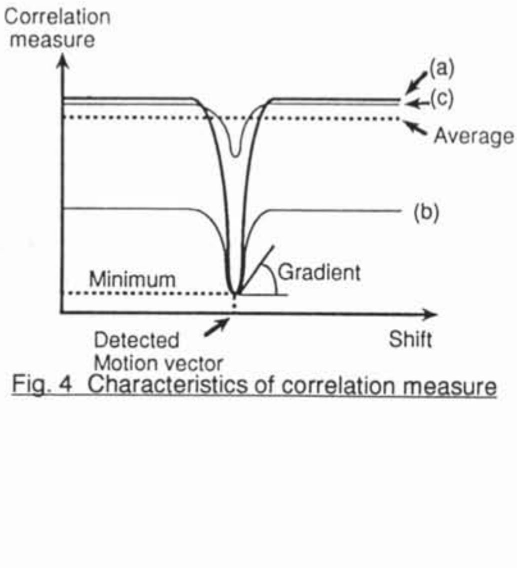


Fig. 4 Characteristics of correlation measure

In a case where a moving object occupies a most of the detection regions, the above shown conditional equations of correlation function is not adequate to give a satisfactory judgment. Therefore, judgement is carried out by examining a correspondence of these to the motion vector derived from different image regions[2].

4. NORMALIZATION OF ENTIRE IMAGE MOVEMENTS AND DIFFERENCE

Once a motion vector of entire picture frame is detected, by using this, the motion vector is normalized. Formula (4) is a conditional equation utilized to detect moving regions, and a region which satisfies Formula (4) in the coordinate (x,y) is detected as a moving region. The normalization of motion vector is accomplished, as shown in the first term in Formula (4), by shifting the relative positions of two frames according to the detected vector.

$$|S(t,x,y) - S(t-1,x+Vx,y+Vy)| - a \times (|dS/dx| + |dS/dy|) > H \qquad (4)$$

wherein S is an image signal, t, t-1 are two frames, x,y are positions of an image, Vx,Vy are detected motion vector, dS/dx, dS/dy are spatial gradients of image signals, a is a predetermined factor, and H is a predetermined threshold value.

As the difference between frames can be taken as a product of spatial gradient and movements of an image, the difference between frames can be larger when the spatial gradient is larger even if the movement remains at a constant rate. Therefore, in a background where spatial gradient is large, there would be more possibility that this region could be detected as a moving region because of a small detection error in motion vector. The second term in Formula (4) is provided to suppress this possibility.

By doing this, as shown in Fig. 1, the moving region produced by a moving object can be detected as a case of still background. A postprocessing such as an elimination of small region from the result obtained by the above is then carried out.

5. TIME-SEQUENTIAL PROCESSING

A moving region detected only by the difference between two continuous frames includes a part of background appeared by a movement of moving object, and a part of moving object may not be detected as a moving region if it had a small spatial gradient. Therefore, an accumulation of moving regions derived from the differences between two frames is performed in order to identify a region of moving object more accurately.

In this case, the following operations shown below are conducted as a time accumulation processing. As shown in Formula (5-1), synthesized image data (synthesized background data S'(t-1, x,y)) is employed instead of image data (S(t-1,x,y)) of the preceding frame to obtain the difference between two frames employed in Formula (4). The image data (S'(t,x,y)) of synthesized background is synthesized from the image of present frame and the synthesized background of preceding frame according to the results of detection of motion region ,as shown in Formula (5-2).

$$|S(t,x,y) - S'(t-1,x+Vx,y+Vy)| - a \times (|dS/dx| + |dS/dy|) > H \qquad (5\text{-}1)$$

When (x,y) are related to a moving region,
$S'(t,x,y) = S'(t-1, x+Vx, y+Vy)$,
and if (x,y) are unrelated,
$$S'(t,x,y) = S'(t,x,y) \qquad (5\text{-}2)$$

By repeating this process time-sequentially, an initial picture and a new background appeared by panning are added to the image data(S'(t,x,y)) to be synthesized. Furthermore, by taking a difference between this background scenery and the input image to the present frame, a moving object and a new background appeared by movement of moving object are detected as moving regions. The discrimination between a moving object and a new back-ground appeared by the movement of moving object is possible by examining the motion vector in the corresponding region.

6. LABELLING AND CHARACTERISTICS MATCHING

When plural moving objects are presented in a picture frame, each moving object has to be identified and tracked. In this case, labelling of connected regions are conducted according to their image data (x,y) of motion regions, and after this identification and tracking of each moving object are conducted based on the characteristics (color-distribution, size, location) of each labelled moving region.

7. EXPERIMENT

According to the above explained method, an experiment to detect a man walking in a background which is being panned is conducted. Detection regions are provided at four corners of frame, and thirty representative points are provided for each of these detection regions. The detection of motion vector is carried out by a representative point matching method of these points.

After the motion vector is detected, the number of pixels are compressed into one-fourth of that of original in both horizontal and vertical directions for reducing the volume of arithmetic processes.

Figs. 5(a) to 5(c) show input images, and Figs. 6(a) to 6(c) show background sceneries synthesized by a time sequential processing at each time. It should be noticed that the image of man is wiped off from the input image, and background without showing the man is synthesized as shown in Fig. 6(b) and after.

Fig. 7 shows a result of detection of moving region (region of man). Figs. 7(a) to 7(c) show detections of moving regions detected at the succeeding frame by using background shown in Fig. 6. Fig. 7(a) show a detection of regions produced by adding a man to a new background appeared by movement.

The results shown since (b) show exact detection of position of man. Fig. 7(c) shows a case where the number of men is plural. color distribution of images in the moving region is utilized to identify two moving objects in this case.

8. CONCLUSION

The advantages of proposed motion vector detection method are summarized as follows.

(1) Detection of moving objects out of general pictures including a panned background became possible by normalizing the movement of background.

(2) Motion vector of background can be detected by using a matching method of representative-points of a limited number.

(3) Precise detection of motion vector of background became possible by eliminating the effects of a moving object existing in a picture frame.

(4) Synthesis of background excluding moving objects, and thus the exact detection of moving object became possible by applying a time sequential processing to the difference between frames.

(5) Identification and tracking of plural moving objects became possible by memorizing their characteristic features such as color distribution and size and location of each region.

The effectiveness of presently proposed motion vector detection method is experimentally confirmed. Since no complex recursive processes have to be employed, realtime image processes can be executed. Beside a panned background problem mentioned, the detection of moving object out of general moving background including zoomed and rotated background is a next problem to be solved.

REFERENCES

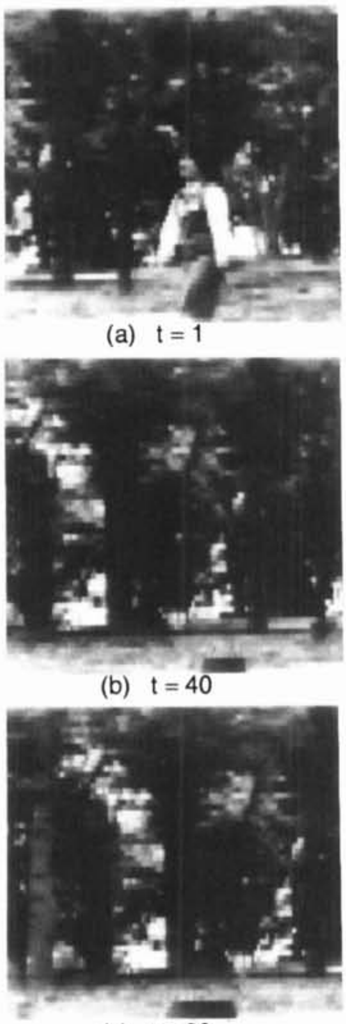
1. H. Ishii et al.: "A motion vector detection method on video image", Spring Annual Conf., Inst. Electronics, Inf. and Com. Eng. of Japan, 1989.
2. K. Uomori et al.: "Intelligent digital image stabilizer" The 20th Image Engineering Conf. 1989.

(a) t = 1

(b) t = 40

(c) t = 60

Fig.5 Input image

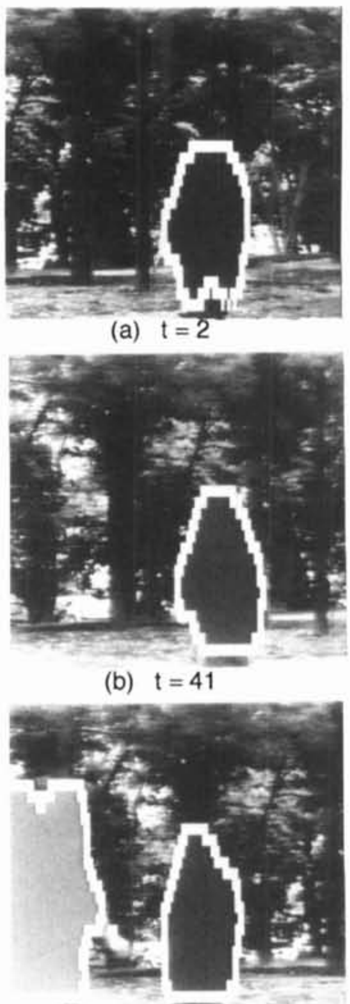


(a) t = 1

(b) t = 40

(c) t = 60

Fig.6 Synthesized background



(a) t = 2

(b) t = 41

(c) t = 61

Fig.7 Result of detection