MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11 - 13, 2002, Nara- ken New Public Hall, Nara, Japan

8—21

Extraction of Moving Objects from Video Sequence Using Estimated Background Brightness

Tomiyasu Ishikawa*
Graduate School of Engineering
Nagoya Institute of Technology

Shinji Fukui†
Programs in Education for Information Sciences
Aichi University of Education

Yuji Iwahori‡
Center for Information and Media Studies
Nagoya Institute of Technology

Hidenori Itoh§
Department of Intelligence and Computer Science
Nagoya Institute of Technology

Abstract

This paper proposes a new method to extract the moving objects from the video sequence. The proposed method is robust to the noise and the illumination change effect of the observed image. It is basically the background subtraction with the background image which can be estimated from the global illumination change in the observed image. A mapping table between the present image and the original background taken in advance, is generated to estimate the present background image. After the background subtraction for each pixel, the mean and the variance for the difference is calculated for each block to get higher robustness for the segmentation. It is also applicable to the illumination change by estimating and using the background brightness which is occluded by the moving object itself. The performance of the real-time implementation of the proposed approach is evaluated in comparison with the previous approaches.

1 Introduction

Real-time extraction of the moving objects from the video sequence is an important topics for the various applications in computer vision. The application includes the measurement of the number of the traffic cars, the observation of the traffic situation, the automatic detection of a trespasser, the data compression in video images, or the analysis of non-rigid motion.

Among the segmentation approaches of moving object, the "Background subtraction" is the most basic and speedy approach. It works well if and only if the background image has the constant brightness, and it fails under the condition that the brightness of the moving object has the close value as the background.

*Address: Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan. E-mail: tomy@center.nitech.ac.jp

†Address: 1 Hirosawa, Igatani-cho, Kariya-shi, Aichi 448-8542, Japan. E-mail: sfukui@auecc.aichi-edu.ac.jp

‡Address: Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan. E-mail: iwahori@center.nitech.ac.jp

§Address: Gokiso-cho, Showa-ku, Nagoya 466-8555 Japan. E-mail: itoh@ics.nitech.ac.jp

While in [1], an approach using both the background subtraction and the "Region Growing with Spacio-Temporal Watersheds" is proposed to extract the moving object . The purpose of this method is object-based coding of the image compression technologies and its real-time implementation is difficult.

While the "Normalized Distance"[2] has been proposed to give the better result for the effect of illumination change. This is also based on the background subtraction. It is still weak to the noise and fails to the points with the low brightness or the observed points with the similar texture as the background.

"Peripheral Increment Sign (PIS)"[3] has been also proposed for the condition that the expectation for the illumination change is difficult. It is applicable to the real-time implementation because of the simple filtering process, but the segmentation of the moving object with itself is still weak to the noise effect.

On the other hand, the probabilistic approach using EM algorithm is proposed to remove the shadow region of the moving object from the estimated moving region [4]. The problem is that it is difficult to estimate the appropriate initial value or the classification is difficult under the complicated background.

This paper proposes a new and a robust approach to the noise and the illumination change effect of the observed image. The proposed approach is basically the background subtraction with the background image which can be estimated from the global illumination change in the observed image. A mapping table between the present image and the original background taken in advance, is generated to estimate the present background image. After the background subtraction for each pixel, the mean and variance for the difference is calculated for each block not for each pixel to get higher robustness for the segmentation. It is also applicable to the illumination change by estimating the background brightness which is occluded by the moving object itself. The performance for the real-time implementation of the proposed approach is evaluated in comparison with the previous approaches.

Section 2 shows the detail of the proposed approach, and section 3 shows the experimental results. The last section includes the conclusion and the future works.

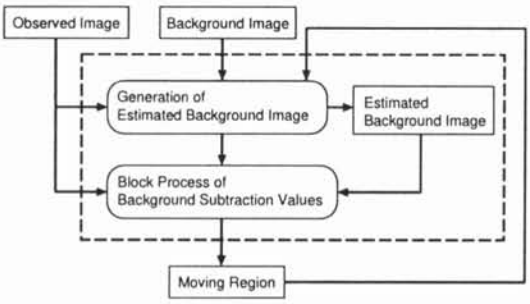


Figure 1: Outline of Proposed Approach

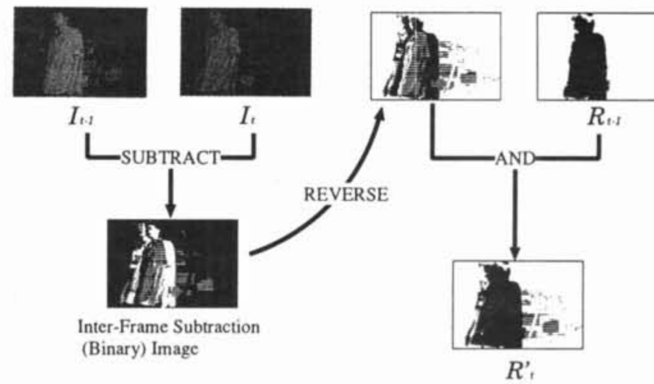


Figure 2: Acquisition of Region R'_t

2 Proposed Approach

The outline of the proposed approach is shown in Figure 1. The first process is introduced to estimate the background image, and the second process is introduced to extract the moving objects according to a block process of the background subtraction values.

The required condition assumes that "background image" is obtained in advance as only the original background image.

2.1 Generation of Estimated Background Image

Let the original background image obtained beforehand be B and let the observed image at time $t(t = 1, 2, \cdots)$ be I_t. The region without the moving objects is defined as the background region R_t.

When intensities in the whole image change under the conditions of illumination or the gamma control, the intensity $I_t(x, y)$ in R_t differs from that of the original background image $B(x, y)$. In such a case, the moving regions cannot be extracted by the simple background subtraction. So, we consider to estimate the background image B_t and extract the moving regions using the subtraction between B_t and I_t. The generation of the estimated background image B_t is performed by the following procedure.

1. Acquisition of region used for generation of intensity conversion table
2. Creation of intensity conversion table
3. Creation of B_t by intensity conversion table

2.1.1 Acquisition of Region used for Generation of Intensity Conversion Table

Let (x, y) and (x', y') be the different points in the image and suppose that $I(x, y)$ is equal to $I(x', y')$ if $B(x, y)$ is equal to $B(x', y')$ when the intensities change in the whole image. This assumption enables to make a mapping table of $B(x, y)$ to $I(x, y)$ based on the relations between $B(x, y)$ and $I(x, y)$ at the background region.

To make the table, the intensity $I_t(x, y)$ should be obtained from the region without the moving regions. Let the candidate of the background region be R'_t. The R'_t is taken as shown in Figure 2. First, the difference values between I_t and I_{t-1} is obtained and binarized by thresholding. Secondly, the binarized values is reversed. And finally, the R'_t is taken as the logic operation AND between the regions with the reversed values and the last estimated background region R_{t-1}. R_0 is necessary to estimate B_1, here, R_0 is treated as the whole region of image.

2.1.2 Intensity Conversion Table

The mapping table of $B(x, y)$ to $I(x, y)$ is made by using the obtained region R'_t. This table converts the intensities from the background image B to the estimated background image B_t. This table is named as "Intensity Conversion Table". A histogram shown in Figure 3 is produced from each set $(I_t(x, y),\ B(x, y))$ in the region R'_t to make the conversion table. The histogram is named as "Intensity Mapping Histogram".

The intensity conversion table is created from this histogram as shown in Figure 4. For example, look at the part where the intensity at B is 26, from the histogram. As shown in the histogram at the 'intensity at B = 26' in Figure 4, some pixels which have the same intensity at B may not have the same value at I_t under the noise effect. So, the conversion table is generated using the mean value of the intensities of a set of pixels at I_t. Since the mean value at I_t becomes 11 in the case of this example, the set $(26, 11)$ of intensities at B and I_t is obtained and added to the conversion table. The set consists of all pixels which have the same intensity in $B(x, y)$ as the candidate of the background region. While, the linear interpolation is applied for the points of range that do not appear in the conversion table.

2.1.3 Creation of B_t by Intensity Conversion Table

The created table converts the intensities at the moving region in I_t from the intensities at B. As a result, the background B_t without the moving objects is estimated. Figure 5 shows a example of B_t.

2.2 Block Process of Background Subtraction Values

The difference value between the observed I_t and the background B_t is calculated for each pixel. With the difference value alone, the moving regions of the brightness close to B_t cannot be detected. I_t is divided

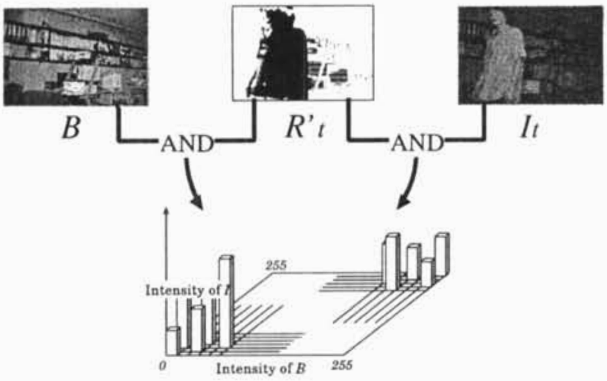


Figure 3: Intensity Mapping Histogram

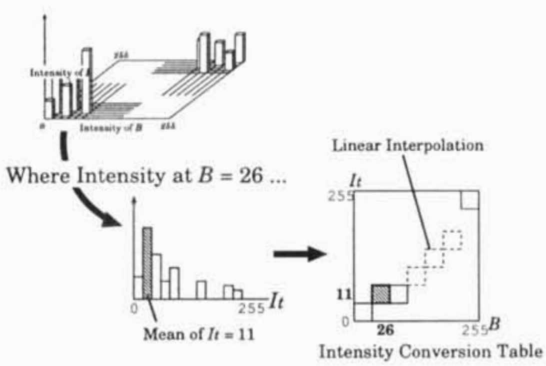


Figure 4: Creating Intensity Conversion Table from Histogram

into some blocks, then calculate the mean m and the variance V of the difference values for each block. If the difference values between I_t and B_t are small for the block, m and V will also become small. Consequently, The block that has small m and V can be regarded as the background region.

The detection for each block is done according to the above process. Subsequently, more detailed outline of the moving object is extracted for each pixel. For the block region that surrounds the outline of the detected moving region, each pixel is classified into either the moving region or the background region, as in the case for the block process.

The moving region is obtained through the above processes, and the detected regions R_t except the moving region are used to generate a background image at next frame.

3 Results

A general purpose digital video camera is used as a video sequence input device. The target image consists of 720 × 480 pixels, and the corresponding block size is taken as 10 × 10. The image is processed as a gray scale obtained from the original RGB images. With a fixed camera, the input image is taken under the general illuminating condition that the intensity of the light source is changed in taking image sequence.

Figure 5: Example of Estimated Background Image

(a) Background Image Obtained in Advance

(b) Observed Image

Figure 6: Experimental Data

Figure 6 shows an example of experimental data. In this example, a scene in the room is taken as a background and a person appear from the right. Figure 6-(a) shows a background image taken in advance and Figure 6-(b) shows an observed image that a moving object (a man) appears. Here, the observed image has the different brightness from the background image.

Figure 7 shows the experimental result. The white region in the result corresponds to the detected moving object, while the black region corresponds to the detected background region.

This approach can extract the moving object with high performance under the illumination change.

Figure 8 shows the comparison between the proposed approach and the others. For the comparison under the same condition, each approach is demonstrated for the block level. The proposed approach is demonstrated without the background subtraction on each pixel around the outline of the moving object. While, for both the standard background subtraction and the PIS, the target image used is converted to a mosaic image beforehand. It is shown that the proposed approach works fine in comparison with the others. From the observation, Figure 8-(a) fails to the case that the moving object has the brightness close to the background and that the background region has

Figure 7: Results

the brightness change. From the PIS, Figure 8-(b) has many noise for the background region. Also, Figure 8-(c) cannot get the exact segmentation for the case that the moving object has the similar texture inside the block as that of the background image and that the block has low intensities.

4 Conclusion

This paper proposed a new approach to extract the moving region using a background image. The approach uses the estimation of the present background image and it has the clear advantage to extract the moving object in real-time with robustness for the illumination change in the background. This is shown through the experimental evaluation in comparison with other previous approaches. The future works include the improvement of the incorrect detection for a background after a moving object passes through, also that for the shadow region produced with the moving objects.

References

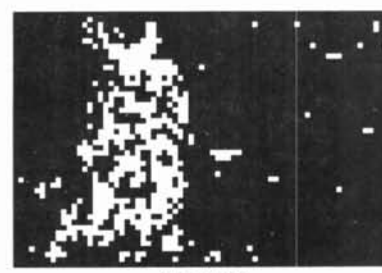
[1] Shinichi SAKAIDA, Masahide NAEMURA, and Yasuaki KANATSUGU, "Moving Object Extraction Using Background Difference and Region Growing with Spacio-Temporal Watersheds," In Transactions of the Institute of Electronics, Information and Communication Engineers D-II, Vol. J84-D-II No. 12, pp. 2541-2555 (2001).

[2] Shigaki NAGAYA, Takafumi MIYATAKE, Takehiro FUJITA, Wataru ITO and Hirotada UEDA, "Moving Object Detection by Time-Correlation-Based Background Judgment Method," In Transactions of the Institute of Electronics, Information and Communication Engineers D-II, Vol. J79-D-II, No. 4, pp. 568-576 (1996).

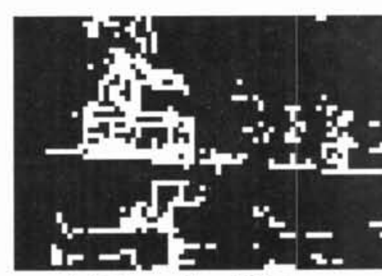
[3] Yutaka SATO, Shun'ichi KANEKO, and Satoru IGARASHI, "Robust Object Detection and Segmentation by Peripheral Increment Sign Correlation Image," In Transactions of the Institute of Electronics, Information and Communication Engineers D-II, Vol. J84-D-II, No. 12, pp. 2585-2594 (2001)

[4] Nir Friedman and Stuart Russell, "Image Segmentation in Video Sequences," In Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence, Providence, Rhode Island: Morgan Kaufmann, 1997.

(a) Standard Background Subtraction



(b) PIS



(c) Normalized Distance

(d) Proposed Approach

Figure 8: Comparing Proposed Approach with Others