SELECTED PAPER

Background Subtraction Using Multiradial Proportion Filter

Shinya Miyamori, Kazunori Saito, Yohei Fukumizu and Hironori Yamauchi

Graduate School of Science and Engineering, Ritsumeikan University
1-1-1 Noji-higashi, Kusatsu-shi, Shiga 525-8577, Japan
Phone/Fax: +81-77-561-2867
E-mail: {ri013063@ed, ksaito@fc, fukumizu@se, yamauchi@sc}.ritsumei.ac.jp

Abstract

In the real-time segmentation of the moving objects in an image sequence, methods based on background subtraction are used widely. Bipolar radial reach correlation (BP-RRC) has realized robust background subtraction through the evaluation of a local texture per pixel to suppress the influence of brightness variations. However, this method has the weak point that it is affected by nonstatic backgrounds. We propose a robust and stabilized background subtraction algorithm that can cope with texture changes caused by various illumination changes and nonstatic backgrounds by improving this texture background model. We verified the proposed method using image sequences that include a loose lighting variation, a rapid lighting variation, and a moving object in the background. We proved that the adaptability of the proposed method is higher than that of the conventional method on a nonstatic background as a result of verification.

1. Introduction

The function of extracting moving objects from a video sequence is a fundamental and crucial problem in manyvision systems such as video surveillance [1-2] and human detection. Typically, the common approach for discriminating a moving object from the background scene is background subtraction, which involves deriving a background image, subtracting each frame from this image and binarizing the result. Here, the result is a binary image that highlights regions of moving objects. The simplest form of the reference image as a background is a time-averaged image. Many adaptive background-modelling methods have been proposed to cope with gradual illumination changes in the scene. Stauffer and Grimson modeled each pixel in a camera scene using an adaptive parametric mixture model of the Gaussian distributions [3]. Lee presented an effective learning algorithm that improved the convergence rate in modeling the background using the Gaussian mixtures [4]. Koller et al. used a Kalman filter to track the changes in background illumination for every pixel [5]. These methods can work well with the gradual illumination changes. However, they cannot handle the problem that illumination changes suddenly in a scene. Although Matsuyama et al. used a correlation to make segmentation robust during illumination change, the outline of moving objects cannot be extracted correctly, because this block-based algorithm forces a low spatial resolution [6]. These algorithms cause the following process, e.g., tracking and recognition to fail, because the accuracy and the efficiency of detection have very great impacts on these tasks.

One of the solutions to the above problems is to use the pixel-based texture background model proposed by Satoh and coworkers [7-8], namely, bipolar radial reach correlation (BP-RRC). This method has realized robust background subtraction through the evaluation of a local texture per pixel to suppress the influence of illumination change. In [7-8], a method that adaptively adjusts the domain according to the local characteristic of a background image was introduced. This makes it possible to cope with the various backgrounds and object appearance within an image sequence. The method can yield regions of moving objects for each pixel, which causes a problem in background subtraction, in which many textures are used.

Since BP-RRC uses texture information, it has the feature that it does not detect offset/gain change. However, it has the problem that it cannot respond to texture change when an object in the background moves. We propose a background subtraction algorithm, multiradial proportion filter (MPRF), which generates and chooses the combination of filters by adjusting texture to background variations. Thereby, we developed a method of separating the background stably, even when texture changes occurred caused by the movement of various objects in the background. The remainder of this paper is organized as follows. Details of the MPRF are described in Section 2. Section 3 shows the experimental results for each method. Conclusions are given in Section 4.

2. Proposed Method

2.1 Basic concept

The fundamental idea of MPRF is focused on the preservability of a brightness difference between two points in an image [8]. Here, we focus on the relationship of the brightness for pixels $p$ vs $p_l$; for example, $p < p_l$ in Figure 1. This relationship does not change until the brightness of either two pixels reaches saturation, even when brightness in the image changes completely. MPRF judges the change in a background using this characteristic. In other words, this method does not detect the offset/gain change of brightness.
2.2 Radial proportion filter

The basic structure of MPRF is shown in Figure 2. From the point of interest \( p \) in the image, we draw virtual lines radially in eight directions. These lines are called proportion lines. First, the number of pixels that have a brightness difference from the pixel of interest is counted. Here, the proportions of the number of pixels with positive and negative brightness to the total on a line are calculated. Next, the line on which the difference in brightness to the point of interest \( p \) exceeds a threshold is calculated. In total, eight sets corresponding to the radial directions from point \( p \) are obtained. After relationships among the brightnesses in each of the sets are evaluated, a majority decision from the total of eight sets is made which enables us to discriminate moving objects robustly.

![Figure 2: Radial proportion filter](image)

2.3 Determination of proportion values

The direction vectors \( d_k \) \((k = 0, 1, \ldots, 7)\) showing the radial directions along which extend proportion lines are defined as follows:
- \( d_0 = (1,0)^T \), \( d_1 = (1,1)^T \),
- \( d_2 = (0,1)^T \), \( d_3 = (-1,1)^T \),
- \( d_4 = (-1,0)^T \), \( d_5 = (-1,-1)^T \),
- \( d_6 = (0,-1)^T \), and \( d_7 = (1,-1)^T \).

Around the point of interest \( p = (x, y)^T \) in the image, the proportion value \( l_k^+ \) is the proportion of the number of points that have a positive difference in brightness among all pixels on the proportion line and \( l_k^- \) is the proportion having negative difference in brightness. In each proportion line, \( l_k^+ \) and \( l_k^- \) are respectively defined as

\[
l_k^+ (p) = \frac{|\{l f(p + l d_k) > f(p) & l \leq l_{\text{max}}\}|}{l_{\text{max}}} \quad (1)
\]

\[
l_k^- (p) = \frac{|\{l f(p + l d_k) < f(p) & l \leq l_{\text{max}}\}|}{l_{\text{max}}} \quad (2)
\]

where \( f(p) \), \( l_{\text{max}} \), and \( l_{\text{max}} \) are the brightness at \( p \), a threshold parameter, and the maximum length of the proportion line, respectively. The positive proportion value \( l_k^+ (p) \) and the negative proportion value \( l_k^- (p) \) are evaluation values of the brightness around the points of interest.

2.4 Multiradial proportion filter

MRPF is a mixture model of RPF. In MPRF, a positive and a negative two-dimensional mixed Gaussian distribution is defined for each pixel in the eight directions in the current image \( g(p) \), respectively. The probability density distribution of each mixed Gaussian for which \( l_k^+ \) distributions are given is as

\[
l_k = \left( \begin{array}{c} l_k^+ \\ l_k^- \end{array} \right)
\]

\[
P(l_k) = \sum_{k=1}^{K} \pi_k N(l_k | t_{k,m}, \Sigma_{k,m})
\]

\[
= \sum_{m=1}^{M} \pi_m \frac{1}{2 \pi | \Sigma_m |} \exp \left\{ -\frac{1}{2} (l_k - t_{k,m})^T \Sigma_{k,m}^{-1} (l_k - t_{k,m}) \right\}
\]

where \( M \), \( l_{k,m} \), and \( \Sigma_{k,m} \) \((m \in \{1, 2, \ldots, M\})\) are the number of mixtures of MPRF, the average, and the dispersion matrix, respectively.

2.5 Distinction by MPRF

MRPF consists of a distinction step and an updating step. Each step is explained hereinafter.

2.5.1 Distinction step

RPF evaluation values, \( b_k(p) \) of the pixel of interest \( p \) are calculated by evaluating the Mahalanobis distance of proportion values between the current image and the background model that was presumed from observational data.

\[
h_{k,m} = (l_{k,m}(p) - l_k(p))^T
\]

\[
\Sigma_{k,m}(p)^{-1} \left( l_{k,m}(p) - l_k(p) \right)
\]

\[
b(p)_{k,m} = \begin{cases} 0 & \text{if } T_1 \leq h_{k,m} \\
1 & \text{otherwise}
\end{cases}
\]

Here, \( T_1 \) is a threshold value. The MPRF evaluation value \( b_k'(p) \) is calculated using \( b(p)_{k,m} \) determined from eq.(6).

\[
b_k'(p) = \prod_{m=1}^{M} b(p)_{k,m}
\]

The correlation \( B \) is calculated in order to evaluate the similarity between a background image and an input picture:

\[
B(p) = \sum_{k=0}^{K} b_k'(p)
\]

A binary-format image \( C(p) \) is obtained by the thresholding of \( B(p) \).

\[
C(p) = \begin{cases} 1 & B(p) > T_B \quad (0 \leq T_B \leq 1) \\
0 & \text{otherwise}
\end{cases}
\]
the foreground and background.

2.5.2 Updating step

MRPF updates the proportion value and the dispersion matrix of a background model using an on-line EM algorithm for each frame. The updating equation of the average values and the dispersion matrix of each Gaussian mixture model at the $i$-th updating are as

$$l'_{k,m,i}(p) = (1 - \eta_{m,i})l'_{k,m,i-1}(p) + \eta_{m,i}l_k(p) \quad (10)$$

$$\Sigma_{k,m,i}(p) = (1 - \eta_{m,i})\Sigma_{k,m,i-1}(p) + \eta_{m,i}(l'_{k,m,i-1}(p) - l_k(p))(l'_{k,m,i-1}(p) - l_k(p))^T \quad (11)$$

where $\eta_{m,i}$ is the step size that is calculated using the update count and each parameter [4].

2.6 Threshold values

Threshold values $T_f$ and $T_b$ discriminate between the foreground and background. Equation (6) evaluates the similarity in each line; it equals 1 when the model matches and 0 otherwise. Equation (8) evaluates the similarity of a background model and the observational data from the result of Eq (6). It equals 8 when everything matches and 0 when there is no match. Therefore, since $B(p)$ can be considered to be 8 times the Bernoulli trials, we can model this frequency function with a binomial distribution:

$$P(B(p) = u) = \binom{8}{u}q^u(1 - q)^{8-u} \quad (12)$$

where $q$ is an expected value that takes 1 for each line. The probability distributions $p_b$ and $p_f$ of $B$ for a background and a foreground region are shown in Figure 3. False positives and false negatives are minimized by setting the threshold values $T_f$ and $T_b$ to minimize the overlap area of the two probability distributions. Although the background probability distribution can be derived from $T_b$, the probability distribution of the foreground requires the covariance matrix in addition to $T_f$. Then, a threshold value is determined from the experimental results.

3. Experiment

In this section, we demonstrate the performances of the proposed algorithm, the Gaussian mixed model (GMM) [4], and BP-RRC [8] to the manual segmentation. As experimental data, we used the Wallflower dataset [9] with resizing to the resolution of $320 \times 240$. Light Switch in the Wallflower dataset has frame sequences of changing brightness suddenly owing to the light on and off, while Time of Day has gradual changes. Also, Waving Trees has a swaying tree in the background.

3.1 Experimental results

Example frames of the experimental results are shown in Figure 4. We evaluated performance in terms of the number of FNs (false negatives, i.e., the number of foreground pixels that are marked as background) and FPs (false positives, i.e., the number of background pixels that are marked as foreground). In order to compare performance among image sequences, we show the results also as a ratio of FNs to the foreground pixels and that of FPs to the background pixels. We calculated the average processing time per frame for these algorithms. These performance results are summarized in Table 1. We used a computer with Intel® Core™2 Quad 2.83 GHz of CPU in the experiment.

Although the GMM was very effective for gradual illumination variations and nonstatic backgrounds, it generated FPs in the case of sudden illumination change such as the Light Switch result shown in Figure 4. In BP-RRC, FN was generated in part of the image where the texture changes due to swaying trees and/or multiple illuminations. The proposed method can also enable us to detect the nonstatic backgrounds robustly. Furthermore, the proposed method has succeeded in restraining FP in many scenes because of the introduction of proportion values. As shown in Table 1, the proposed method requires more processing time compared with the others. However, if the updating interval of MRPF can be extended, the proposed method can shorten the processing time.

4. Conclusions

We verified that BP-RRC was a more effective background model than the GMM with regarding to illumination variations, which are not based on observational data, and pointed out a problem that cannot be completely solved within the waving background. Thus, in this paper, a new filter (MRPF) using positive and negative proportions on a radial line was proposed for detecting moving objects from image sequences. MRPF can respond to various background scenes by updating several background models adaptively. Performance comparison of MRPF with the GMM and BP-RRC indicate that MRPF gives higher robustness in nonstatic backgrounds and in illumination variations than the conventional methods. Our method successfully enables moving object detection continuously under sudden illumination change, without the need for an extra process that redefines the background model.
Table 1: Experimental results

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Error Type</th>
<th>Time of Day</th>
<th>Problem Type</th>
<th>Waving Trees</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False neg.</td>
<td>763[pixel]</td>
<td>1695[pixel]</td>
<td>5115[pixel]</td>
</tr>
<tr>
<td>MPRF [this paper]</td>
<td>False pos.</td>
<td>3238[pixel]</td>
<td>4768[pixel]</td>
<td>4846[pixel]</td>
</tr>
<tr>
<td></td>
<td>Processing time</td>
<td>179.6[ms/frame]</td>
<td>158.7[ms/frame]</td>
<td>153.7[ms/frame]</td>
</tr>
<tr>
<td></td>
<td>False pos.</td>
<td>1071[pixel]</td>
<td>13438[pixel]</td>
<td>5315[pixel]</td>
</tr>
<tr>
<td></td>
<td>Processing time</td>
<td>15.7[ms/frame]</td>
<td>16.2[ms/frame]</td>
<td>16.1[ms/frame]</td>
</tr>
<tr>
<td>GMM</td>
<td>False neg.</td>
<td>2084[pixel]</td>
<td>2079[pixel]</td>
<td>6435[pixel]</td>
</tr>
<tr>
<td></td>
<td>False pos.</td>
<td>1394[pixel]</td>
<td>63243[pixel]</td>
<td>2063[pixel]</td>
</tr>
<tr>
<td></td>
<td>Processing time</td>
<td>15.7[ms/frame]</td>
<td>15.4[ms/frame]</td>
<td>25.4[ms/frame]</td>
</tr>
</tbody>
</table>

References


