A Background Reconstruction for Dynamic Scenes

Mei Xiao  
Electronic & Information Engr.  
Xi’an Jiaotong University  
Xi’an Shaanxi P.R.China  
xiaomeiazhanglei@hotmail.com

Chongzhao Han  
Electronic & Information Engr.  
Xi’an Jiaotong University  
Xi’an Shaanxi P.R.China  
czhan@mail.xjtu.edu.cn

Xin Kang  
Electronic & Information Engr.  
Xi’an Jiaotong University  
Xi’an Shaanxi P.R.China  
xkang@mail.xjtu.edu.cn

Abstract - Based on assumption that background would not be the parts which appear in the sequence for a short time, a background reconstruction algorithm based on online clustering was proposed in this paper. Firstly, pixels intensities are classified based on online clustering. Secondly, cluster centers and appearance probabilities of each cluster are calculated. Finally, a single or multi intensities clusters with the appearance probability greater than threshold are selected as the background pixel intensity value. Simulation results show that the algorithm can represent situation where the background contains bi-model or multi-model distribution, and motion segmentation can be performed correctly. The algorithm with inexpensive computation and low memory can accommodate the real-time need.

Keywords: Background reconstruction, online clustering, motion segmentation, video sequence analysis

1 Introduction

Segmentating moving objects from a video sequence is a fundamental and critical task in video surveillance, traffic monitoring and analysis. Background subtraction is a common approach to identifying the moving objects, especially for a video sequence from a stationary camera. Background reconstruction is the heart of background subtraction algorithm. Though many background reconstruction algorithms have been proposed in the literatures, the problem of background reconstruction for complex dynamic scenes is still far from being completely solved. We classify background reconstruction techniques into statistical model [1-3] prediction technique [4-5] and background assumption [6-10]. They are described as below:

A simple method of background reconstruction was time-averaged background image (TABI) which a background approximation was obtained by averaging a long time image sequences. The method was effective on situations with no moving objects. However, once scenes had many moving objects, especially when they moved slowly, the foreground objects were always blended into the background image. Wren et al [1] used a single Gaussian to model the statistical distribution of a background pixel, nevertheless it does not cope with multimodal backgrounds. Stauffer et al [2] extended background reconstruction approach by modeling the pixel color as a mixture of Gaussians (MOG). The method could deal with slow changes in illumination, repeated motion from background clutter and long term scene changes. But the MOG was computationally intensive and its parameters require careful tuning. Elgammal et al [3] adapted a multi-model background estimate at each pixel location. Gaussian was chosen to be the kernel estimator and the median of the absolute differences between successive frames was used as the width of the kernel. The model could handle dynamic situations which contained small motions. However, the foreground detection is more complex. Ridder et al [4] and Makhoul [5] for background reconstruction were based on prediction technique. Ridder et al [4] modeled each pixel with Kalman Filter which made their system more robust to lighting changes in the scene, but the method recovers slowly and the foreground objects were easily blended into background. Makhoul [5] modeled each pixel with Weiner Filter which was a simpler version of the Kalman filter. Any pixel different significantly from its predicted value were declared foreground. Long et al [6] presented an adaptive smoothness method based on the assumption that background were stable for a long period. Their method found intervals of stable intensity and used a heuristic which chose the longest, most stable interval as the one most likely to represent the background. Median filter was one of the most commonly-used background modeling techniques [7-8]. The background estimation was defined to be the median at each pixel location of all the frames in the buffer. The assumption was that the pixel stayed in the background for more than half of the frames in the buffer. Kornprobst et al [9] assumed that background would be defined as the most often observed part over the sequence and presented a approach to deal with the background reconstruction based on Partial Differential Equations (PDE). The result was good, but their method was complex and the parameters were difficult to choose. Hou et al [10] proposed a pixel intensity classification (PIC) method. The method assumed that background pixel intensity appeared in image sequence with the maximum probability and the intensity value with the maximum frequency was selected as the background pixel intensity value.

In this paper, we propose a background reconstruction method which can effectively construct the background for dynamic scene. The proposed method works well for both static and dynamic scenes. The paper is organized as follow. In section 2 the background reconstruction algorithm based on online clustering is presented. The
Simulation results and comparisons are then analyzed and shown in section 3. Finally, section 4 concludes the paper.

2 Background reconstruction algorithm based on online clustering

Three curves that represent a pixel intensity change in 100 frames are shown in Fig. 1 and image gray levels is 256. Data 1 curve typically occurs at rainy daytimes, the illumination of scene changes suddenly at the 31st frame. Obviously, from 1st to 30th frames scene corresponds to one background and from 31st to 100th frames scene corresponds to another background. The curve of intensity changed is also similar to the status of objects changed from moving to motionless or form stationary to moving. Data 2 curve typically occurs in traffic monitor system. We can see that pixel intensity changes greatly in several short intervals, for example the interval from 13th to 20th frames, which means moving objects passing through. For data 3 curve, it shows two scatter plots distribution of a pixel value resulting from waving tree. This environments can not be characterized by a single background, so multi-backgrounds are necessary.

2.1 Algorithm Steps

There are many problematic phenomena such as repetitive motion of the background, sudden illumination change and sensor noise in real scene. We aim to construct the background for complex situation where there are parts of problematic phenomena which have been described in [11]. The steps of algorithm are described as follows:

Step 1, classify the pixel intensity based on online clustering [12].

To accommodate the real time needs of many applications, background reconstruction algorithm must be computationally inexpensive and have low memory requirements, so clustering must be performed on-line in our system. The online clustering algorithm is as follows: $N$ frames are selected from the sequences and marked as $(I_1, I_2, ..., I_N)$, $I_i(p)$ represents the intensity value of the pixel $p$ of the $i^{th}$ frame where $i = 1, 2, ..., N$. $C^i(p)$ and $m^i(p)$ represent the center and the number of the $i^{th}$ cluster separately. Let $\delta$ represent a threshold. Starts with the guess that the first input value is the initial class, and the number of the initial class set to 1. The initial guess value for these cluster centers is most likely incorrect. We alter only the cluster center most similar to a new pattern being presented, and the cluster center is changed. The algorithm is illustrated in table 1.

Step 2, remove those clusters with a smaller appearance probability.

Assume $L(p)$ clusters are obtained after clustering operation, $m^1(p), m^2(p), ..., m^{L(p)}(p)$ represent the pixel’s number of each cluster, Appearance probability $W^i(p)$ of the $i^{th}$ class is denoted as follow:

$$W^i(p) = \frac{m^i(p)}{\sum_{i=1}^{L(p)} m^i(p)}, \quad (i = 1, 2, \cdots, L(p)) \quad (1)$$

Those clusters, which appear in the sequence for a long time, are the candidates for the background. $W^i(p)$ is taken as the criterion to describe the probability of the $i^{th}$ class staying in the sequence. We cast away those clusters which appearance probabilities are lower than threshold $\xi$ and preserve the clusters that appearance probabilities are higher than $\xi$ as multiply candidate backgrounds.

Step 3, choose background intensity value.

In real scene, multi-surfaces often appear in the view frustum of a particular pixel and the lighting conditions are often changed, so choosing a single image as background always results in large errors in moving object detection. Motivated by the work of Stauffer et al [2], we choose a single or multi-images as the background images. Suppose there are $n(p), (n(p) \leq L(p))$ clusters with a larger appearance probability in the video sequences. The average intensity means the cluster center of each class, can be marked as $C^i(p)$. After removal those intensity clusters with a small appearance probability, then the $n(p)$ clusters are chosen as the multi-background images:

$$B^i(p) = C^i(p), \quad (i = 1, 2, \cdots, n(p)) \quad (2)$$

We must readjust appearance probability of candidate backgrounds

$$W^i(p) = \frac{m^i(p)}{\sum_{i=1}^{n(p)} m^i(p)}, (i = 1, 2, \cdots n(p)) \quad (3)$$

![Fig.1. Example intensity history plot of a pixel in 100 frames](image)
Table 1 Online Clustering

\[
\begin{align*}
\text{begin} & \quad \text{initialize} \quad \delta \\
C^i(p) & \leftarrow I_i(p) \\
\text{do} & \quad \text{input new} \quad I_i(p) \\
& \quad \quad i \leftarrow \arg\min_{i} \left| I_i(p) - C^i(p) \right| \\
& \quad \quad \text{if} \quad \left| I_i(p) - C^i(p) \right| < \delta \\
& \quad \quad \quad m' = m' + 1 \\
& \quad \quad \quad C^i(p) \leftarrow \frac{I_i(p) + (m'(p) - 1) \cdot C^i(p)}{m'(p)} \\
& \quad \quad \text{else} \quad \text{add new class} \quad C^{i+1}(p) \\
& \quad \quad \quad m^{i+1}(p) \leftarrow 1 \\
\text{until} & \quad \text{no more patterns} \\
\text{return} & \quad C^i(p), C^{i+1}(p), \ldots; m^i(p), m^{i+1}(p), \ldots \\
\text{end}
\end{align*}
\]

2.2 Background update

The maintenance and update of background are very important to moving detection. An ideal background maintenance system will be able to avoid several problems in realistic environments. The problems have been discussed in detail by Toyama et al.[11]. In this paper, we adapted the background update strategy proposed in [13] to deal with those problems. The background update schemes have been verified, both in [13] and in our experiments, to have good performances.

3 Simulation results and comparisons

There examples show the results of background reconstruction by using our algorithm. To compare with our approach, the results calculated by TABI(time-averaged background image), PIC[5] and MOG[5] are also given here. In the simulation the parameters are choosen as: \(N = 100, \delta = 10, \xi = 0.18 \) and \(\sigma = 25\). The video shows the pure detection results without any morphological operations, noise filtering and tracking information of targets. Fig. 2 (a), (b) and (c) are the 1st frame, 6th frame and 26th frame; (d) is the background image using TABI; (e) and (f) are the results of motion detection corresponding to (b) and (c) by subtracting background image (d) separately; (g), (h) and (i) are multi-background images using our algorithm; (j) and (k) are the results of motion detection corresponding to (b) and (c) by subtracting multi-background images separately. It can be seen that foreground objects are blended into background which is obtained by TABI, therefore, there are much errors in moving object detection. In highway sequence a large of pixel is single distribution, however, double or more backgrounds have been rebuilt in some pixels because of the noise. From Fig 2(g), (h) and (i), we can see the number of background is different in vary pixel. And there are few pixels with nonzero intensity in image (i) which masked with circle. Our method can removal the effect of moving object and construct the correct background image. Therefore, motion segmentation can be performed correctly.

Fig. 4 (a), (b) and (c) are the 1st frame, 12th frame and 30th frame; (d) is the background image using TABI; (e) and (f) are the results of motion detection corresponding to (b) and (c) by subtracting background image (d) separately; (g), (h) and (i) are multi-background images using our algorithm; (j) and (k) are the results of motion detection corresponding to (b) and (c) by subtracting multi-background images separately. It can be seen that foreground objects are blended into background which is obtained by TABI, therefore, there are much errors in moving object detection. In Sweedeny sequence a large of pixel is single distribution, however, double or more backgrounds have been rebuilt in some pixels because of the noise. From Fig 3(g), (h) and (i), we can see the number of background is different in vary pixel. And there are few pixels with nonzero intensity in image (i) which masked with circle. Our method can removal the effect of moving object and construct the correct background image.
image using PIC; (f-h) are background images using MOG; (i-l) are background images using our method.

Fig. 5 shows the moving detection corresponding to Fig. 4. Fig. 5 (a1) and (a2) are the 1st frame and 24th frame; Fig. 5 (b1) and (b2) are moving detection of PIC corresponding to (a1) and (a2); Fig. 5 (c1) and (c2) are moving detection of MOG corresponding to (a1) and (a2); Fig. 5 (d1) and (d2) are moving detection of our method corresponding to (a1) and (a2). From Fig 5, it can be found there are a lot of false in background image using PIC because of wind and rain, which results in false detection almost in whole image. The detection results of our method are as good as MOG. Only a few parts of vehicle such as the windows are classified into background. Simulations show that our methods can handle the situation where the scene contains small motions such as tree branch motion.
4 Conclusions

In this paper, a robust background construction algorithm was introduced. The method is to class the pixel based on online clustering. Online clustering can economize computation time and save space. A single or multi images are chosen as the background images according to scene characteristic. Simulation results show that the algorithm can handle situation where the scene contains small motions such as tree branch motion.

References


**Appendix**

Fig. 2 (i), Fig. 3 (i) and Fig. 4(l) were zoomed as follows:

![Zoomed Image](image-url)