

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

An Improved Background Reconstruction Algorithm Based on Basic Sequential Clustering

Mei Xiao

Key Laboratory of Automobile Transportation Safety Technology Ministry of Communication,
School of Automobile, Chang'an University, Xi'an, 710064, Shaanxi, People's Republic of China

Abstract: Based on the assumption that background appears with large appearance frequency, a new background reconstruction algorithm based on basic sequential clustering is proposed in this research. First, pixel intensity in period of time are classified based on mend basic sequential clustering. Second, merging procedure and reassignment procedure are run to classified classes. Finally, pixel intensity classes, whose appearance frequencies are higher than a threshold, are selected as the background pixel intensity value. So the improved algorithm can rebuilt the background images of various scenes. Compared with the background reconstruction method based on basic sequential clustering, the simulation results show that those near classes are avoided at all and the effect of input order of data has been reduced greatly in our method. And the background model can represent the scene well.

Key words: Basic sequential clustering, merging procedure, reassignment procedure, background reconstruction

INTRODUCTION

Identifying moving objects from a video sequence is a fundamental and critical task in video surveillance, traffic monitoring and analysis. Several approaches are known to separate foreground from background. Background subtraction is a common approach to identifying the moving objects, especially, for a video sequence from a stationary camera. Pixels in the current frame that deviate significantly from the background are considered to be moving objects. Background reconstruction and foreground detection are two major steps in a background subtraction algorithm. Background reconstruction uses the new video frame to calculate and update a background model. Foreground detection then identifies pixels in the video frame that cannot be adequately explained by the background model and outputs them as a foreground mask. Present research focuses on background reconstruction.

A large number of background reconstruction methods for fixed cameras have been proposed in recent years. We classify background modeling techniques into approach based on background assumption (Long and Yang, 1990; Gloyer *et al.*, 1995; Cucchiara *et al.*, 2003; Kornprobst *et al.*, 1999; Hou and Han, 2004; Xiao and Han, 2007), statistical model (Wren *et al.*, 1997; Stauffer and Grimson, 1999; Javed *et al.*, 2002; Lee *et al.*, 2003; Zivkovic and Van, 2004; Elgammal *et al.*, 2000) and prediction techniques (Ridder *et al.*, 1995; Toyama *et al.*, 1999).

A adaptive smoothness method (Long and Yang, 1990) based on the assumption that intensity of background is stable for a long period. Their method finds intervals of stable intensity and uses a heuristic which choose the longest, most stable interval as the one most likely to represent the background. Median filter is one of the most commonly-used background modeling techniques (Gloyer *et al.*, 1995). The background estimate is defined to be the median at each pixel location of all the frames in the buffer. The assumption is that the pixel stays in the background for more than half of the frames in the buffer. An improved algorithm about Median filtering is presented in Cucchiara *et al.* (2003) that has been extended to color by replacing the median with the medoid. Assumed that background would be defined as the most often observed part over the sequence, they (Kornprobst *et al.*, 1999) presented a approach to deal with the background reconstruction and motion segmentation based on Partial Differential Equations (PDE), the result is good, but their method is complex and the parameters are difficult to choose. These methods can reconstruct background even though foreground objects are present in the sequence and can avoid blending. Pixel Intensity Classification (PIC) is presented in Hou and Han (2004). The method assume that background pixel intensity appears in image sequence with the maximum probability, then classified pixels intensity based on the calculated pixel intensity difference between inter-frame, finally, the intensity value with the maximum frequency is selected as the background pixel intensity value. In this

study a multi background reconstruction based on online cluster has been proposed. Our algorithm adopted the assumption that background wouldn't be the part appear a short time of the sequence. In this approach, first, pixel intensity is classified based on online cluster, then cluster center and appear probability of each class is calculated, finally, a single or multi pixel intensity, with the appear probability is higher than a threshold, are selected as the background pixel intensity value. Based on the assumption that background appears with large appearance frequency, a background subtraction algorithm based on Basis Sequential Clustering (BSC) is proposed in literature (Xiao and Han, 2007). First, pixel intensity in period of time are classified based on basic sequential clustering and pixel intensity classes, whose appearance frequency is higher than a threshold, are selected as the background pixel intensity value. Simulation results show that the algorithm can hand complex situations contains small motions and motion detection can be performed correctly. It is obvious that the results of BSC are strongly dependent on the order in which the data are presented to the algorithm. And it may happen that two of the formed clusters are very closely located.

Some of the commonly-used statistic model are Time Averaged Background Image (TABI), Gaussian model (Wren *et al.*, 1997; Stauffer and Grimson, 1999) and Non-parametric Model (Elgammal *et al.*, 2000). A simple method of background reconstruction is TABI, a background approximation is obtained by averaging a long time image sequences, the method is effective in empty scene, once situations with many moving objects, especially if they move slowly and the foreground objects always can be blending into the background image. A single Gaussian (Wren *et al.*, 1997) is considered to model the statistical distribution of a background pixel, it does not cope with multimodal backgrounds. Extended background subtraction approach (Stauffer and Grimson, 1999) by using an adaptive multi-modal subtraction method that modeled the pixel color as a Mixture of Gaussians (MOG). This method could deal with slow changes in illumination, repeated motion from background clutter and long term scene changes. But The MOG method is computationally intensive and its parameters require careful tuning. Literature (Javed *et al.*, 2002; Lee *et al.*, 2003; Zivkovic and Van, 2004) put forward kinds of modify, but computationally intensive has not solved yet. Adapted a multi-model background estimate (Elgammal *et al.*, 2000) at each pixel location, Gaussian was chosen to be the kernel estimator and the method uses the median of the absolute differences between successive frames as the width of the kernel. The model can handle situations

where the dynamic background but contains small motion. However, the method is computationally intensive and it is not easy to choose its parameters.

Kalman Filter (KF) (Ridder *et al.*, 1995), Wiener filter (Toyama *et al.*, 1999) for background reconstruction are based on prediction techniques. Ridder *et al.* (1995) 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 are easily to be blended into background. A simpler version of the Kalman filter called Weiner filter (Toyama *et al.*, 1999), it is a linear predictor based on a recent history of values. Any pixel that different significantly from its predicted value are declared foreground.

THE STEPS OF THE IMPROVED BACKGROUND RECONSTRUCTION ALGORITHM

Literature (Hou and Han, 2004; Xiao and Han, 2007) reports the character of the pixel intensity in detail. We adopted the same assumption that was mentioned in literature (Xiao and Han, 2007), that background would not be the parts which appear in the sequence for a short time, a background reconstruction algorithm based on Basic Sequential Clustering (BSC) was proposed in this research. The steps of algorithm are described as follows:

Step 1: Classify the pixel intensity based on BSC.

The basic idea of BSC is the following: Starts with the guess that the first input value is the initial class and the number of the initial class set to 1. As each new data is considered, it is either assigned to an existing cluster or assigned to a newly created cluster, depending on its distance from the already formed ones. Let N be the frames that selected from the sequences and marked as $F = \{f_1, f_2, \dots, f_n\}$. Pixel (x, y) is used to illustrate the BSC. The BSC is stated as Table 1.

Where, $f_t(x, y)$ represents the intensity value of the pixel (x, y) of the $t(t = 1, 2, \dots, N)$ frame. $C_i^t(x, y)$ ($i=1, 2, \dots, n$) and $m_i^t(x, y)$ represent the center and the number of the i th cluster separately. $d(f_t(x, y), C_i^t(x, y)) = |f_t(x, y) - C_i^t(x, y)|$ denote the distance (or dissimilarity) between a input date $f_t(x, y)$ and a cluster $C_i^t(x, y)$, we adopt the same distance in follow step.

Let δ represent a threshold. The following process is run to same pixel (x, y) . The omission of pixel position is necessary for concise writing.

Step 2: Run merging procedure.

In basic sequential clustering algorithm, it may happen that two of the formed clusters are very closely

Table 1: The BSC procedure

```

F(x, y) = {f1(x, y), f2(x, y), ..., fn(x, y)};
Begin initialize δ1
  n = 1;
  C1n(x, y) = f1(x, y);           %normalize the center of initial cluster
  m1n(x, y) = 1;                 %normalize the number of initial cluster
  Do input ft(x, y)(∀t ∈ N)
    Find d(ft(x, y), C1k(x, y)) = min1 ≤ j ≤ n d(ft(x, y), C1j(x, y)) %find the nearest cluster
    If d(ft(x, y), C1k(x, y)) ≤ δ1 then
      m1k(x, y) = m1k(x, y) + 1;   %update the number of cluster
      C1k(x, y) =  $\frac{f_t(x, y) + (m_1^k(x, y) - 1) \cdot C_1^k(x, y)}{m_1^k(x, y)}$ ; %update the center of cluster
    Else
      n = n + 1;                   %create new cluster
      C1n(x, y) = ft(x, y);       %normalize the center of new cluster
      m1n(x, y) = 1;             %normalize the number of new cluster
  Until t = N
  Return
End

```

Table 2: The Merging procedure

```

(A): Find C1i, C1j (i < j) such that d(C1i, C1j) = mink,r=1,...,n,k≠r d(C1k, C1r)
  If d(C1i, C1j) ≤ δ2, then
    Merge C1i, C1j to C1i and eliminate C1j;
    C1i =  $\frac{C_1^i \cdot m_1^i + C_1^j \cdot m_1^j}{m_1^i + m_1^j}$ ;           %update the center of ith cluster
    m1i = m1i + m1j;           %update the number of ith cluster
    Rename the clusters C1jth, ..., C1n to C1j, ..., C1n-1, respectively
    n = n - 1;
    Go to (A);
  Else
    Stop;
  End {If}

```

located and it may be desirable to merge them into a single one. Consider the data $F = \{1, 11, 5, 10, 6, 9, 8, 4, 70, 100\}$. If we present the F in the above order to the BSC and we set $\delta_1 = 10$, we obtain $C_1 = \{4, 9.5, 70, 100\}$ and $m_1 = \{4, 4, 1, 1\}$. We can see C_1^1 and C_1^2 are very closely located. Such cases cannot be handled by the BSC. One way out of this problem is to run the simple merging procedure, after the termination of the BSC.

If the number of clusters is n after the termination of the BSC, $C_1^i (i=1,2,\dots,n)$ is the center of ith cluster. Then the merging procedure may be written as Table 2.

δ_2 is a user-defined parameter that quantifies the closeness of two clusters, C_1^i and C_1^j .

Step 3: Run reassignment procedure.

The other drawback of the BSC is their sensitivity to the order of presentation of data. Suppose, for example,

that in using BSC, f_2 is assigned to the first cluster C_1^1 and after the termination of the algorithm four clusters are formed. Then it is possible for f_2 to be closer to a cluster different from C_1^1 . However, there is no way for f_2 to move to its closest cluster once assigned to another one. A simple way to face this problem is to use the reassignment procedure.

It is assumed that there are p clusters after the termination of the merging procedure, C_2^i and $m_2^i (i=1,2,\dots,p)$ are the center and the number of ith cluster that the merging procedure has created.

Then the reassignment procedure may be stated as Table 3.

In this procedure, $b(f_i)$ denotes the closest to cluster.

Step 4: Calculate the appearance frequency of all clusters.

Table 3: The reassignment procedure

```

b(fi) = 0, ∀ fi ∈ F;
m3i = 0, ∀ i ∈ {1, 2, ..., p};
C3i = 0, ∀ i ∈ {1, 2, ..., p};
For t = 1 to N
    Find C2j such that d(fi, C2j) = mink=1,...,p d(fi, C2k);
    set b(fi) = j;
End {For}
For t = 1 to N
    For j = 1 to p
        If b(ft) = j
            m3j = m3j + 1; % update the number of cluster
            C3j =  $\frac{C_3^j \cdot (m_3^j - 1) + f_t}{m_3^j}$ ; % update the center of cluster
        End {If}
    End {For}
End {For}

```

Table 4: The background selection procedure

```

i = 1; j = 1;
For i = 1 to q
    If Wi ≥ δ3
        Bj = C3i; mBj = m3i; j = j + 1;
    End {If}
End {For}

```

Assume that q clusters are formed now. The appearance frequency of ith cluster:

$$W^i = \frac{m_3^i}{\sum_{i=1}^q m_3^i}, (i=1, 2, \dots, q) \quad (1)$$

where, m₃ⁱ is the number of ith cluster that the reassignment procedure has created.

Step 5: Select background of pixel.

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 and Grimson (1999), we choose a single or multi-images as the background images. Suppose there are R(R ≤ q) clusters with a larger appearance frequency in the video sequences. After removal those intensity clusters with a small appearance probability, then the R(R ≤ q) clusters are chosen as the multi-background images. The background selection procedure are showed as Table 4.

Where, δ₃ is a user-defined parameter, m_Bⁱ the number of data of background, B^j (j = 1, 2, ..., R) is the jth background that our algorithm has constructed.

We must readjust appearance frequency of candidate backgrounds:

$$W_B^i = \frac{m_B^i}{\sum_{i=1}^R m_B^i}, i=1, 2, \dots, R \quad (2)$$

SIMULATION RESULTS AND COMPARISONS

Two examples show the results of background reconstruction by using our algorithm. The results calculated by TABI, PIC (Hou and Han, 2004), MOG (Stauffer and Grimson, 1999) and BSC (Xiao and Han, 2007) are given in detail in literature (Xiao and Han, 2007). In the study, we only give the simulation results compare the algorithm of BSC and our method. In the simulation the parameters in BSC are chosen as: N = 100, δ = 13 and δ_c = 0.2. The parameters that are referred in our algorithm are chosen as: N = 100, δ₁ = 13, δ₂ = 13 and δ₃ = 0.2. The video shows the pure detection results without any morphological operations, noise filtering and tracking information of targets.

Figure 1 is result of background reconstruction of synthetic sequence. Suppose that all pixels in 100 frames have values as Table 5. BSC results in two backgrounds have values (158.07, 165.8). From the simulation result, we can see the two backgrounds are very closely and we should merge them into a single one. However, our method construct only one suitable background has values 160, therefore, background reconstruction is performed correctly.

Figure 2 is an indoor sequence. One hundred frames are used to construct the background. The number of background is different in vary pixel, so we use black denote empty background. For example, suppose that pixel (1, 1) has three backgrounds B¹ (1, 1), B² (1, 1) and B³ (1, 1). Pixel (2, 2) has one background B¹ (2, 2), which results in two empty backgrounds B² (2, 2) and B³ (2, 2), so we can see black in B² (2, 2) and B³ (2, 2). From the background image of BSC, we can see part of foreground (see the area that is masked with ellipse in Fig. 2(a1)) is considered as background. It is clear that our method leads to more reasonable results than BSC. Motion detections also show that false detection of our method is less greatly than BSC. The motion detections (Fig. 2(d1) and Fig. 2(d2), Fig. 2(e1) and Fig. 2(e2)) show that our method can eliminate the false motion detection.

Table 5: Intensity of pixel in 100 frames

150	150	150	150	150	150	150	150	150	150	150	170	170	170	170	170	170	170	170	170	170	156	156	156	156	156	
164	164	164	164	164	157	157	157	157	157	157	163	163	163	163	163	158	158	158	158	158	158	162	162	162	162	162
159	159	159	159	159	159	159	159	159	159	159	159	159	159	159	159	159	159	159	159	159	159	159	159	159	159	
161	161	161	161	161	161	161	161	161	161	161	161	161	161	161	161	161	161	161	161	161	161	161	161	161	161	

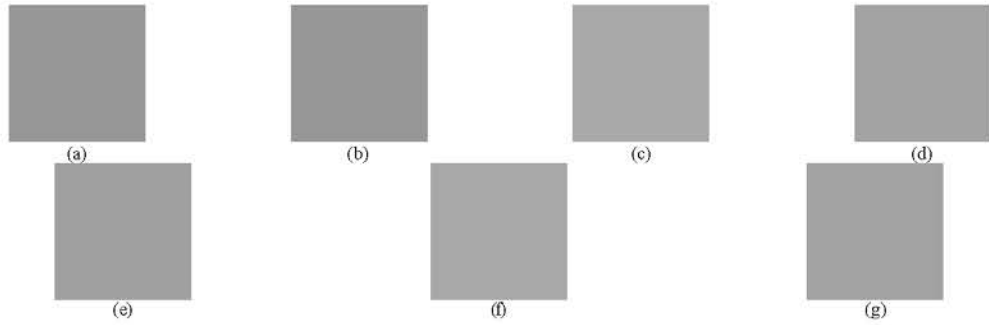


Fig. 1: Result of background reconstruction of synthetic sequence. (a)-(d) are frame #5, #15, #50 and #80 separately; (e) and (f) are the backgrounds B^1 and B^2 of BSC; (g) is the background of out method

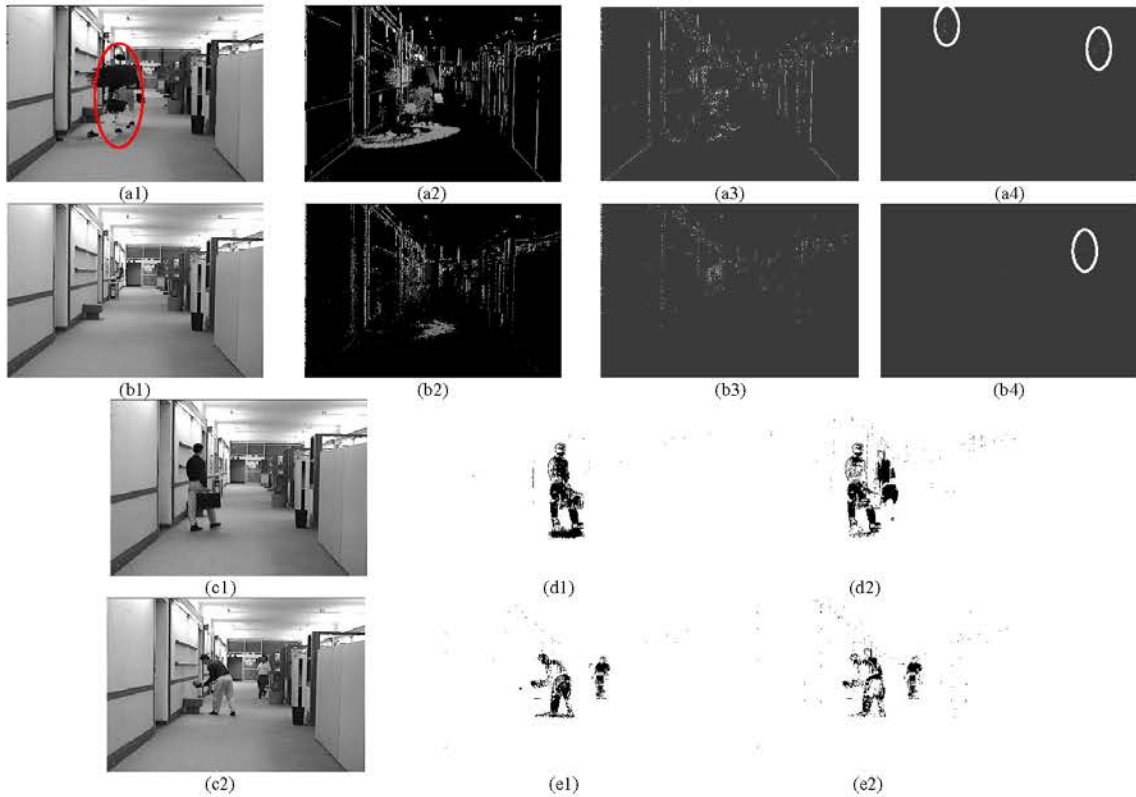


Fig. 2: Result of background reconstruction and motion detection of *Hall* sequence. (a1)-(a4) are the backgrounds B^1 - B^4 of BSC; (b1)-(b4) are the backgrounds B^1 - B^4 of our method; (c1) and (c2) are frame #39 and #100; (d1) motion detection of 39th frame of our method; (d2) motion detection of 39th frame of BSC; (e1) motion detection of 100th frame of our method; (e2) motion detection of 100th frame of BSC

CONCLUSIONS

Background reconstruction is core of background subtraction. An improved background reconstruction based basic sequential clustering is proposed in the paper. The merging procedure and assigned procedure are used after the termination of the BSC algorithm. Our method algorithm has two traits: 1) background reconstruction does not need priori knowledge of scene. 2) the algorithm can consume a lot of time of computation and save memory requirements. In addition, those near classes are avoided at all and the effect of input order of data has been reduced greatly.

REFERENCES

- Cucchiara, R., M. Piccardi and A. Prati, 2003. Detecting moving objects, ghosts and shadows in video streams. *IEEE Trans. Pattern Anal. Mach. Intel.*, 25 (10): 1337-1342.
- Elgammal, A., D. Harwood and L. Davis, 2000. Non-parametric model for background subtraction. *Proceeding of 6th European Conference on Computer Vision*. Dublin Ireland, pp: 246-252.
- Gloyer, B., H.K. Aghajan, K.Y. Siu and T. Kailath, 1995. Video-based freeway monitoring system using recursive vehicle tracking. *Proceeding of IS and T-SPIE Symposium on Electronic Imaging: Image and Video Processing*, pp: 173-180.
- Hou, Z.Q. and C.Z. Han, 2004. A background reconstruction algorithm based on pixel intensity classification in remote video surveillance system. *Proceeding of the 7th Int. Conf. Inf. Fusion*. Fairborn, United States, pp: 754-759.
- Javed, O., K. Shafique and M. Shah, 2002. A hierarchical approach to robust background subtraction using color and gradient information. *Proceeding of the Workshop on Motion and Video Computing*. Orlando, FL, pp: 22-27.
- Kornprobst, P., R. Deriche and G. Aubert, 1999. Image sequence analysis via partial difference equations. *J. Math. Imaging Vision*, 11 (1): 5-26.
- Lee, D.S., J.J. Hull and B. Erol, 2003. A Bayesian framework for Gaussian mixture background modeling. *Proceeding of International Conference on Image Processing*. Barcelona, Spain, pp: 973-979.
- Long, W. and Y. Yang, 1990. Stationary background generation: An alternative to the difference of two images. *Pattern Recognition*, 23 (12): 1351-1359.
- Ridder, C., O. Munkelt and H. Kirchner, 1995. Adaptive background estimation and foreground detection using Kalman-filter. *Proceeding of International Conference on Recent Advances in Mechatronics*, pp: 193-199.
- Stauffer, C. and W.E.L. Grimson, 1999. Adaptive background mixture models for real-time tracking. *Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Los Alamitos, CA, USA., pp: 246-252.
- Toyama, K., J. Krumm and B. Brumitt *et al.*, 1999. Wallflower: Principles and practice of background maintenance. *Proceedings of the 7th IEEE International Conference on Computer Vision*. Piscataway, NJ, USA., pp: 255-261.
- Wren, C.R., A. Azarbayejani, T. Darrell and A.P. Pentland, 1997. Pfister: Real-time tracking of the human body. *IEEE Trans. Pattern Anal. Mach. Intel.*, 19 (7): 780-785.
- Xiao, M. and C.Z. Han, 2007. Background subtraction algorithm based on online clustering. *Moshi Shibie yu Rengong Zhineng*, 20 (1): 35-41.
- Zivkovic, Z. and D.H.F. Van, 2004. Recursive unsupervised learning of finite mixture models. *IEEE Trans. Pattern Anal. Mach. Intel.*, 26 (5): 651-656.