ISSN 2464-4617 (print) WSCG 2016 - 24th Conference on Computer Graphics, Visualization and Computer Vision 2016
ISSN 2464-4625 (CD-ROM)

Background Modeling using Perception-based Local Pattern

K. L. Chan

Department of Electronic Engineering
City University of Hong Kong
83 Tat Chee Avenue
Kowloon, Hong Kong

itklchan@cityu.edu.hk

ABSTRACT

Background modeling is an important issue in video surveillance. A sophisticated and adaptive background model can be used to detect moving objects which are segregated from the scene in each image frame of the video via the background subtraction process. Many background subtraction methods are proposed for video acquired by a stationary camera, assuming that the background exhibits stationary properties. However, it becomes harder under various dynamic circumstances – illumination changes, background motions, shadows, camera jitter, etc. We propose a versatile background modeling method for representing complex background scenes. The background model is learned from a short sequence of spatio-temporal video data. Each pixel of the background scene is represented by samples of color and local pattern. The local pattern is characterized by perception-inspired features. In order to cater for changes in the scene, the background model is updated along the video based on the background subtraction result. In each new video frame, moving objects are considered as foregrounds which are detected by background subtraction. A pixel is labeled as background when it matches with some samples in the background model. Otherwise, the pixel is labeled as foreground. We propose a novel perception-based matching scheme to estimate the similarity between the pixel and the background model. We test our method using common datasets and achieve better performance than various background subtraction algorithms in some image sequences.

Keywords

Background modeling, Moving object detection, Dynamic background, Background subtraction, Local pattern

1. INTRODUCTION

One of the most challenging problems in computer vision is to detect and recognize moving objects such as humans or vehicles in complex environments automatically. Video surveillance [Hsi08a] is obviously one well-known application. For instance, automatic video surveillance systems for human motion monitoring typically consist of the human detection, tracking of targets along the video sequence, and inference of the motion. Besides, other areas such as gait analysis [Cun03a] and video segmentation and retrieval [Lu04a], also benefit from the advance in moving object detection research. The detection of moving objects as foregrounds in the video is the first key problem. To detect moving targets, one common approach is to create a model representing the background scene. The background model is used to detect moving objects by the background subtraction process. At the same time, the background model is updated to cater for the changes in the scene. In each image frame of the video, the background subtraction process is to find out those pixels that are similar to the background. The pixels that are not similar to the background belong to the moving objects (foregrounds). The process involves matching of the pixels with the background model.

Background model can be created and updated from the video. One common assumption is that the video is acquired by a fixed camera and the scene is stationary or changes slowly. However, the scene is not always static. The captured environment can have dynamic elements such as illumination changes, waving trees, water, etc. Strong wind can cause camera jitter. Therefore, sophisticated background modeling methods are proposed for tackling scene variations and background movements. Sobral and Vacavant [Sob14a] presented a recent review and evaluation of 29 background subtraction methods. One approach is to represent the background scene by parametric model. For instance, pixelwise background color can be modeled by Gaussian distribution. Stauffer *et al.* [Sta00a] proposed the modeling of background colors using mixture of Gaussian (MOG) distributions as individual scene pixels may exhibit multiple colors because of background motions or illumination changes. Background model is initialized using an EM algorithm. Pixel values that do not match any of the background distributions are regarded as foreground. Parameters of the MOG model are updated after foreground detection. Since its introduction, MOG has gained widespread popularity and inspired many improvements. For instance, in contrast with a fixed number of Gaussians in the original MOG model, Zivkovic [Ziv04a] proposed an algorithm for selecting the number of Gaussian distributions using the Dirichlet prior. A comprehensive survey on the improvements of MOG model can be found in [Bou08a].

Another approach is to create non-parametric background model. This category of background subtraction methods does not assume the pdf of the scene follow a known parametric form. Elgammal *et al.* [Elg02a] proposed an algorithm for estimating the pdf directly from previous pixels using kernel estimator. Barnich and Van Droogenbroeck [Bar09a, Bar11a] proposed a sample-based background subtraction algorithm called ViBe. Background model is initialized by randomly sampling of pixels on the first image frame. Pixel of the new image frame is classified as background when some samples intersecting the sphere of the pixel. A random policy is also employed for updating the background model at the pixel location and its neighbor. Hofmann *et al.* [Hof12a] proposed a similar non-parametric sample-based background subtraction method. The method can adaptively adjust the foreground decision threshold and model update rate along the video sequence. Haines and Xiang [Hai14a] presented a non-parametric background modeling method based on Dirichlet process Gaussian mixture models. Gao *et al.* [Gao14a] and Liu *et al.* [Liu15a] regarded the observed video frames as a matrix, which can be decomposed into a low-rank matrix of background and a structured sparse matrix of foreground.

Recently, methods for modeling background scene by local pattern are proposed. Heikkilä and Pietikäinen [Hei06a] proposed to model the background of a pixel by local binary pattern (LBP) histograms estimated around that pixel. Liao *et al.* [Lia10a] proposed the scale invariant local ternary pattern (SILTP) which can tackle illumination variations. St-Charles *et al.* [Stc15a] proposed a pixelwise background modeling using local binary similarity pattern (LBSP) estimated in the spatio-temporal domain. Their method outperforms 32 state-of-the-art methods on the ChangeDetection.net dataset [Goy12a, Wan14a].

In this work, we have two contributions. First, we propose a novel perception-based local pattern which can be used effectively to characterize various dynamic circumstances in the scene. Second, we propose a novel scheme to estimate the similarity between new pixel and the background model for classifying the pixel. The background model and the pixel classification are incorporated into the background subtraction method for moving object detection. The background subtraction result is used to update the background model.

2. BACKGROUND MODEL INITIALIZATION

It is common that the background model is created from the video. The modeling method must be versatile since various scene complications may be encountered. We consider that the feature representing the background scene is an important factor. To make the modeling method generic, there should be as few tunable parameters as possible.

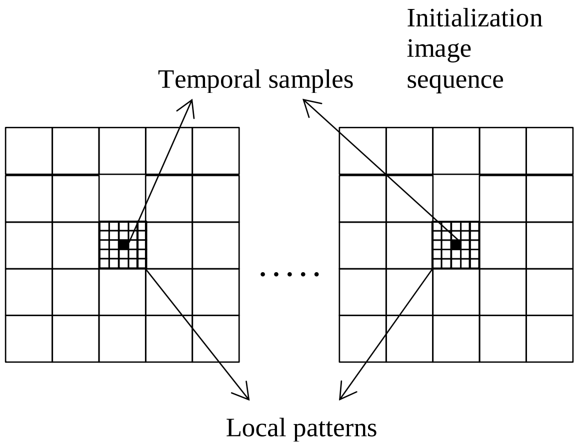

Figure 1. Spatio-temporal sampling of background pixels.

In sample-based background subtraction, the background model is generated by taking previous samples at the same pixel position like [Elg02a], or taking random samples on the first image frame [Bar09a, Bar11a]. We observed various challenges in real scenes. Dynamic background elements such as tree and water produce many false positive errors. Camera jitter also produces false positive errors. It is because the background model does not contain sufficient and representative samples. We propose to take samples from the spatio-temporal domain. As shown in Figure 1, in background model initialization, a number of image frames are used. At a given pixel location (the dark pixels in Figure 1), colors of all the samples (temporal samples) at the same position are entered into the background model for that pixel. In addition, a block is defined centered at that pixel and local pattern feature is extracted from this block of pixels. All spatio-temporal local patterns, sampled from all pixels of a short initialization image sequence, are also entered into the background model. We have performed experimentations and finally fixed the number of initialization image frames as 30 and the block size as 5 x 5 pixels as shown in Figure 1. Static

background can be represented by the temporal samples while dynamic background can be represented by the spatio-temporal local patterns. In case there are moving objects in the initialization image frames, the model still contains background samples as far as the objects are not stationary. The effectiveness of the background model can be seen in the results from camera jitter videos in section 4.

In dynamic scenes, the colors of background elements can vary due to illumination change. The variations of colors must be allowed in matching the new pixel with the background model. Inspired by the perception-inspired confidence interval [Haq13a], we propose a novel local pattern that can cater for color variations. The confidence interval of a sample having a color component value c is defined as ($c - d$, $c + d$). According to Weber's law [Gon10a], d depends on the perceptual characteristics of c. That is, d should be small for darker color and large for brighter color. The perception-based linear relationship is formulated as

$$d = 0.11 * c \tag{1}$$

Each pixel of the block (except the center pixel) is compared with the center pixel. If its color is outside the confidence interval of the center pixel, its feature value f is set equal to

$$f = b_{half} - d_{city} + 1 \tag{2}$$

where d_{city} is the city-block distance between a given pixel of the block and the center pixel, b_{half} is the half size of the block. If its color is within the confidence interval of the center pixel, its feature value is 0. Therefore, neighbor closer to the center pixel will contribute a larger feature value if they are perceptually different. Different neighbor farther from the center pixel will contribute a smaller feature value. Finally all feature values of the block are summed to form the pattern value for the center pixel. Figure 2a illustrates the formation of a local pattern for a block of 3 x 3 pixels. The first row shows the formation of LBP for a noise-free image. The second row indicates that LBP is not robust to random noise in the image. The third row also shows that LBP cannot keep its invariance against scale transform. Figure 2b illustrates the formation of perception-based local pattern under the same circumstances. The confidence interval for the patterns in the first and second row is (56.96, 71.04). The confidence interval for the pattern in the third row is (113.92, 142.08). It can be seen that perception-based local pattern is robust against random noise and scale transform. Its pattern value is equal to 4.

63	68	42		0	1	0
64	64	27	→	1		0
61	95	83		0	1	1

65	69	42		1	1	0
63	64	27	→	0		0
60	95	83		0	1	1

130	138	84		1	1	0
126	128	54	→	0		0
120	190	166		0	1	1

a

63	68	42		0	0	1
64	64	27	→	0		1
61	95	83		0	1	1

65	69	42		0	0	1
63	64	27	→	0		1
60	95	83		0	1	1

130	138	84		0	0	1
126	128	54	→	0		1
120	190	166		0	1	1

b

Figure 2. Formation of local pattern: (a) LBP, (b) perception-based local pattern.

We observed that the choice of color model can have significant impact on the accuracy of moving object detection. We used invariant color feature to represent the color of the pixel. In our method, we adopted the $c_1c_2c_3$ normalized color model [Gev99a].

$$c_1 = \arctan\frac{R}{\max(G, B)} \tag{3}$$

$$c_2 = \arctan\frac{G}{\max(R, B)} \tag{4}$$

$$c_3 = \arctan\frac{B}{\max(R, G)} \tag{5}$$

3. MOVING OBJECT DETECTION AND BACKGROUND MODEL UPDATING

Figure 3 illustrates the framework of our moving object detection method. The background model is initialized using 30 initial image frames of the video as mentioned in the previous section. In the background/foreground segmentation, all pixels of the current image frame are classified as background or foreground. Since we have generated a strong background model that characterizes the spatial and temporal variations of background colors, we adopt a conservative policy in the joint background/foreground segmentation. If all color component values of the pixel match with some temporal color samples or spatio-temporal local patterns of the background model, the pixel is labeled as background. Otherwise, it is labeled as foreground.

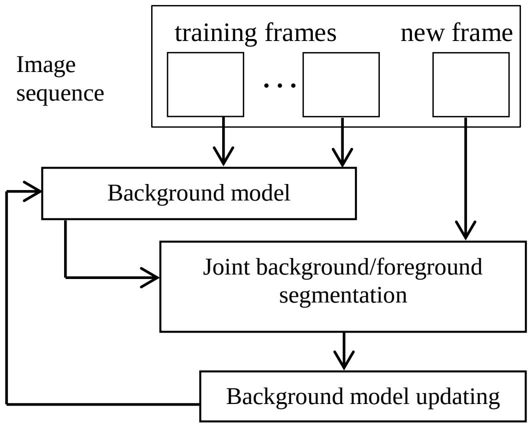

Figure 3. Overview of our moving objects detection method.

We propose a novel scheme to estimate the similarity between the pixel and the background model which strikes for balance between efficiency and perceptual accuracy. First, the pixel is compared with the temporal color samples of the background model. The perception-based confidence interval of the pixel is defined. Once two temporal color samples in the background model are found fall within the confidence interval, the pixel is labeled as background. In static scene, the background subtraction can be accomplished quickly by this process. In dynamic scene, it may not be possible to find similar color samples at the same spatial location along the temporal domain. Then, the pixel is compared with the spatio-temporal local patterns in the background model. A block with this pixel at the center is defined. Pattern values for this pixel are calculated using the same method as mentioned in the previous section. Local pattern of the pixel is compared with the patterns stored in the background model. We define a spatio-temporal search space of 11 x 11 pixels x 30 frames. Two patterns are considered similar if the absolute difference of their pattern values is ≤ a tolerance value. We fixed the tolerance value to 3. If two patterns in the background model match with the local pattern of the pixel, the pixel is labeled as background. Otherwise, the pixel is labeled as foreground. The algorithm of background subtraction is shown below.

Algorithm – background subtraction

For each new pixel
  Define perception-based confidence interval
  Search temporal color samples
  If number of matches = 2
    Label pixel as background
    Step over to the next pixel
  Else
    Calculate perception-based local pattern
    Search spatio-temporal local patterns
    If number of matches = 2
      Label pixel as background
      Step over to the next pixel
    Else
      Label pixel as foreground

In the background model updating, the total number of color samples and local patterns will remain the same. If the new pixel matches with the temporal color samples, one temporal color sample will be updated by the following equation

$$c_b^{new} = (1\text{-}\alpha)c_b^{old} + \alpha c_p \qquad (6)$$

where c_p is the color of the new pixel, c_b is the matched temporal color. We set α equal to 0.05. If the local pattern of the new pixel matches with the patterns of the background model, one local pattern will be updated by the following equation

$$l_b^{new} = (1\text{-}\alpha)l_b^{old} + \alpha l_p \qquad (7)$$

where l_p is the local pattern value of the new pixel, l_b is the matched local pattern value in the background model.

The use of chromaticity in matching the pixel with background model means the background/foreground segmentation is robust to gradual illumination change. We also observe that cast shadow is more likely to be classified as background rather than foreground by using chromaticity. We use the same set up in the experimentation. There are no tunable parameters.

4. RESULT

We implement our method using MATLAB and run on a 2.1 GHz PC with 1 Gbyte memory. For a low-resolution image frame of 320 x 240 pixels, the computation time per image frame is about 5 seconds. In the first experimentation, we evaluate our method quantitatively in terms of Recall (Re), Precision (Pr), F-Measure (F1), False Positive Rate (FPR), and False Negative Rate (FNR) using the Change Detection dataset [Goy12a]. Recall gives the ratio of detected true positive pixels (TP) to total number of foreground pixels present in the ground truth which is the sum of true positive and false negative pixels (FN). Precision gives the ratio of detected true positive pixels to total number of foreground pixels detected by the method which is the sum of true positive and false positive pixels (FP). F-Measure is the weighted harmonic mean of Precision and Recall. It can be used to rank different methods. The higher the value of Re, Pr, and F1, the better is the accuracy. The lower the value of FPR and FNR, the better is the accuracy.

Table 1 shows the average F1 of our method and some well-known parametric and non-parametric background subtraction algorithms obtained from 5 categories of video (baseline - B, dynamic background - DB, camera jitter - CJ, intermittent object motion - IOM, shadow - S), containing 26 image sequences of 47,040 image frames. The best result in a given column is highlighted. No method can achieve the best result in all categories. GMM [Sta00a], KDE [Elg02a] and ViBe [Bar11a] can achieve the best F1 in one category. Our method can achieve the best F1 in two categories of dynamic background and camera jitter, and the results in other categories are close to the best F1.

	B	DB	CJ	IOM	S
GMM	0.825	0.633	0.597	**0.520**	0.716
KDE	**0.909**	0.596	0.572	0.409	0.766
ViBe	0.866	0.459	0.569	0.488	**0.798**
Our method	0.884	**0.635**	**0.671**	0.475	0.712

Table 1. Average F1 of various methods on the Change Detection dataset

We then present a detail comparison of our method with ViBe. We select ViBe because it was showed that ViBe performs better than many state-of-the-art parametric and non-parametric algorithms such as [Ziv04a]. Tables 2 and 3 show the results of our method and ViBe on the dynamic background category respectively. In the tables, the best average results are highlighted. There are six image sequences (boats, canoe, fall, fountain01, fountain02, overpass). The videos contain strong background motions such as moving water and tree shaken by the wind. Our method can achieve higher F1 than ViBe in all image sequences. Our method can achieve better result than ViBe in 3 out of 5 average quantitative measures. Tables 4 and 5 show the results of our method and ViBe on the camera jitter category respectively. There are four image sequences (sidewalk, boulevard, traffic, badminton). The videos were captured by vibrating cameras. All videos are very challenging. Our method can achieve higher F1 than ViBe in 3 out of 4 image sequences. Our method can achieve better result than ViBe in 3 out of 5 average quantitative measures.

Sequence	Re	Pr	F1	FPR	FNR
boats	0.682	0.842	0.754	0.001	0.318
canoe	0.856	0.915	0.885	0.003	0.144
fall	0.713	0.546	0.618	0.011	0.287
fountain01	0.339	0.133	0.191	0.002	0.661
fountain02	0.733	0.470	0.573	0.002	0.267
overpass	0.805	0.780	0.792	0.003	0.195
Average	0.688	**0.614**	**0.635**	**0.004**	0.312

Table 2. Results of our method – dynamic background

Sequence	Re	Pr	F1	FPR	FNR
boats	0.528	0.107	0.178	0.020	0.472
canoe	0.897	0.694	0.783	0.014	0.103
fall	0.833	0.342	0.484	0.036	0.168
fountain01	0.580	0.032	0.061	0.008	0.420
fountain02	0.822	0.428	0.563	0.002	0.179
overpass	0.798	0.600	0.685	0.005	0.202
Average	**0.743**	0.367	0.459	0.014	**0.257**

Table 3. Results of ViBe – dynamic background

Sequence	Re	Pr	F1	FPR	FNR
sidewalk	0.405	0.837	0.546	0.002	0.595
boulevard	0.684	0.867	0.765	0.005	0.316
traffic	0.589	0.736	0.654	0.014	0.411
badminton	0.587	0.927	0.719	0.002	0.413
Average	0.566	**0.842**	**0.671**	**0.006**	0.434

Table 4. Results of our method – camera jitter

Sequence	Re	Pr	F1	FPR	FNR
sidewalk	0.518	0.279	0.363	0.027	0.482
boulevard	0.782	0.444	0.566	0.037	0.219
traffic	0.851	0.559	0.675	0.039	0.149
badminton	0.835	0.562	0.672	0.017	0.166
Average	**0.746**	0.461	0.569	0.030	**0.254**

Table 5. Results of ViBe – camera jitter

Figure 4 shows some visual results from the dynamic background category. The first column shows the original image frames and the results obtained by ViBe. The second column shows the results obtained by our method. The third column shows the corresponding ground truths. The ground truth images contain 5 labels (static, hard shadow, outside region of interest, unknown motion, motion). It can be seen that ViBe produces more false positive errors than our method in all image sequences. ViBe may also produce many false negative errors (see results of boats and overpass). From the figure, it can be seen that our method produces balanced Recall and Precision. That is why our method can achieve higher F1 in all image sequences. Figure 5 shows the visual results from the camera jitter category. Again, ViBe produces more false positive errors than our method in all image sequences. In sidewalk, the stationary human and crossing are erroneously detected as foreground by ViBe. Our method only produces minimal scattered false positive errors in the stationary human, while the crossing is correctly identified as background. In the badminton, the players appear at the beginning of the image sequence. Unfortunately, ViBe erroneously detects those players when they already moved to different places along the image sequence. As shown in the figure, our method can detect the correct number of players.

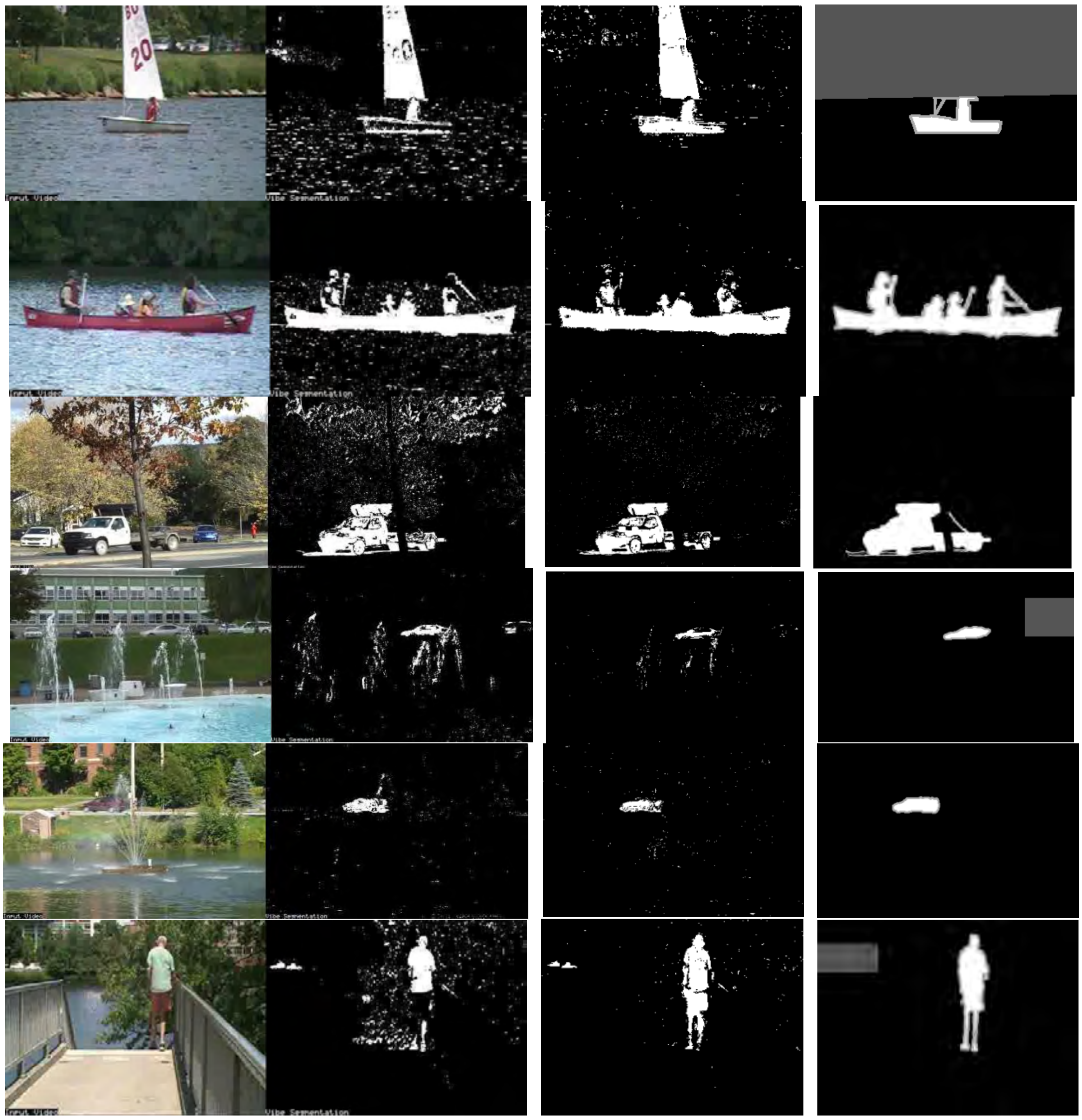

Figure 4. Background subtraction results from the Change Detection dataset dynamic background category – original image frames and results obtained by ViBe (first column), results obtained by our method (second column), ground truths (last column).

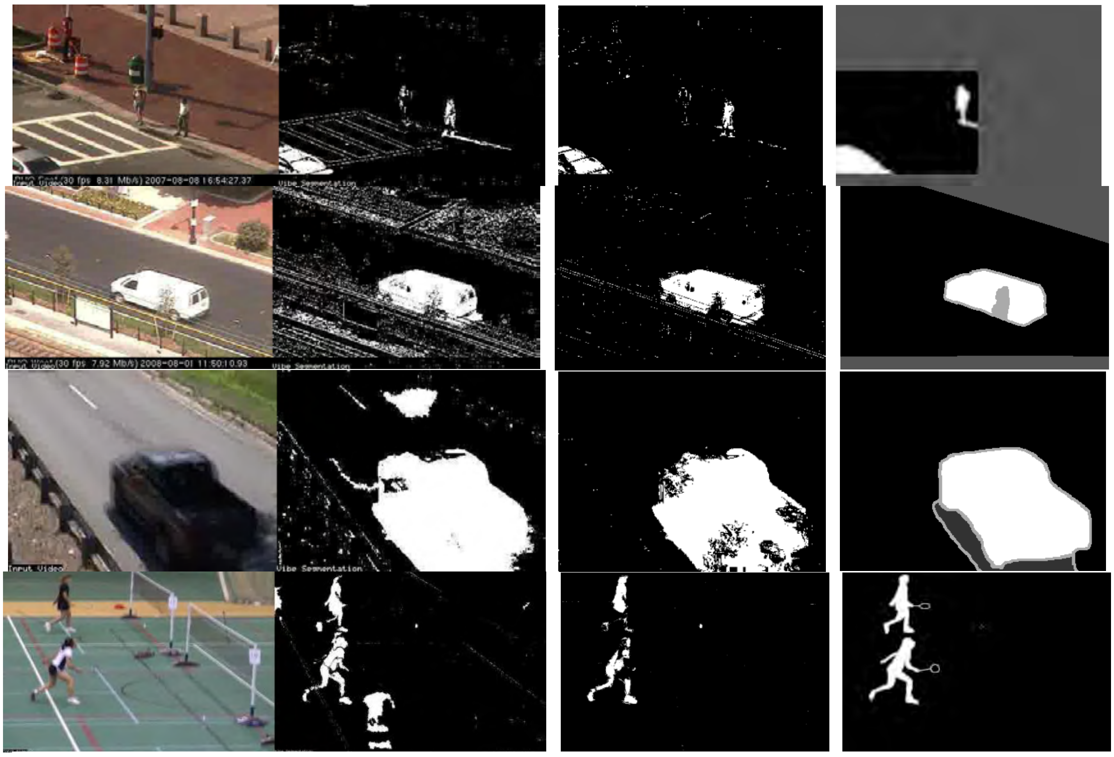

Figure 5. Background subtraction results from the Change Detection dataset camera jitter category – original image frames and results obtained by ViBe (first column), results obtained by our method (second column), ground truths (last column).

In the second experimentation, we compare our method with ViBe and some local pattern based background subtraction algorithms (blockwise LBP – LBP-B [Hei04a], pixelwise LBP – LBP-P [Hei06a]) using the STAR dataset [Li03a]. Table 6 shows the F1 of 9 video sequences. The superiority of local pattern based background model over sampled-based background model can be seen. Our method can achieve the best F1 in 3 video sequences, and the average F1 is second to LBP-P.

Sequence	**LBP-B**	**LBP-P**	**ViBe**	**Our method**
Airport Hall	0.477	**0.503**	0.496	0.429
Bootstrap	0.528	0.520	0.514	**0.569**
Curtain	0.661	0.714	0.775	**0.800**
Escalator	**0.591**	0.539	0.445	0.380
Fountain	0.705	**0.753**	0.425	0.484
Shopping Mall	0.547	**0.629**	0.522	0.548
Lobby	0.503	**0.523**	0.029	0.448
Trees	**0.629**	0.606	0.345	0.600
Water Surface	0.768	0.822	0.801	**0.878**
Average	0.587	**0.635**	0.444	0.600

Table 6. F1 of various methods on the STAR dataset

5. CONCLUSION

We propose a method for the detection of moving objects in video. The background model is represented by samples of color and perception-based local patterns. In moving object detection, each pixel of the current image frame is classified as background if it matches with the background model. Otherwise, the pixel is classified as foreground. This is achieved by our proposed perception-based matching scheme to estimate the similarity between the pixel and the background model. We test and compare our method with various well-known background subtraction algorithms using challenging video datasets. The quantitative measures and visual results show that our method can achieve better performance in some image sequences.

6. REFERENCES

[Bar09a] Barnich, O., and Van Droogenbroeck, M. ViBe: a powerful random technique to estimate the background in video sequences. Proc. of Int. Conf. on Acoustics, Speech and Signal Processing, pp. 945-948, 2009.

[Bar11a] Barnich, O., and Van Droogenbroeck, M. ViBe: A universal background subtraction algorithm for video sequences. IEEE Trans. on Image Processing, Vol. 20, No. 6, pp. 1709-1724, 2011.

[Bou08a] Bouwmans, T., El Baf, F., and Vachon, B. Background modeling using mixture of Gaussians for foreground detection – a survey. Recent Patents on Computer Science, Vol. 1, pp. 219-237, 2008.

[Cun03a] Cunado, D., Nixon, M.S., and Carter, J.N. Automatic extraction and description of human gait models for recognition purposes. Computer Vision and Image Understanding, Vol. 90, pp. 1-41, 2003.

[Elg02a] Elgammal, A., Duraiswami, R., Harwood, D., and Davis, L.S. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. Proc. of IEEE, Vol. 90, No. 7, pp. 1151-1163, 2002.

[Gao14a] Gao, Z., Cheong, L.-F., Wang, Y.-X. Block-sparse RPCA for salient motion detection. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 36, No. 10, pp. 1975-1987, 2014.

[Gev99a] Gevers, T., and Smeulders, A.W.M. Color based object recognition. Pattern Recognition, Vol. 32, pp. 453-464, 1999.

[Gon10a] Gonzalez, R.C., and Woods, R.E. Digital Image Processing. Pearson/Prentice Hall, 2010.

[Goy12a] Goyette, N., Jodoin, P.-M., Porikli, F., Konrad, J., and Ishwar, P. Changedetection.net: a new change detection benchmark dataset. Proc. of IEEE Workshop on Change Detection at IEEE Conf. on Computer Vision and Pattern Recognition, pp. 16-21, 2012.

[Hai14a] Haines, T.S.F., and Xiang, T. Background subtraction with Dirichlet process mixture models. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 36, No. 4, pp. 670-683, 2014.

[Haq13a] Haque, M., and Murshed, M. Perception-inspired background subtraction. IEEE Trans. on Circuits and Systems for Video Technology, Vol. 23, No. 12, pp. 2127-2140, 2013.

[Hei04a] Heikkilä, M., Pietikäinen, M., and Heikkilä, J. A texture-based method for detecting moving objects. Proc. of British Machine Vision Conf., pp. 187-196, 2004.

[Hei06a] Heikkilä, M., and Pietikäinen, M. A texture-based method for modeling the background and detecting moving objects. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 28, No. 4, pp. 657-662, 2006.

[Hof12a] Hofmann, M., Tiefenbacher, P., and Rigoll, G. Background segmentation with feedback: the Pixel-Based Adaptive Segmenter. Proc. of IEEE Workshop on Change Detection at IEEE Conf. on Computer Vision and Pattern Recognition, pp. 38-43, 2012.

[Hsi08a] Hsieh, J.–W., Hsu, Y.–T., Liao, H.–Y.M., and Chen, C.–C. Video-based human movement analysis and its application to surveillance systems. IEEE Trans. on Multimedia, Vol. 10, No. 3, pp. 372-384, 2008.

[Li03a] Li, L., Huang, W., Gu, I., and Tian, Q. Foreground object detection from videos containing complex background. Proc. of ACM Int. Conf. on Multimedia, pp. 2-10, 2003.

[Lia10a] Liao, S., Zhao, G., Kellokumpu, V., Pietikäinen, M., and Li, S.Z. Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes. Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1301-1306, 2010.

[Liu15a] Liu, X., Zhao, G., Yao, J., Qi, C. Background subtraction based on low-rank and structured sparse decomposition. IEEE Trans. on Image Processing, Vol. 24, No. 8, pp. 2502-2514, 2015.

[Lu04a] Lu, C.M., and Ferrier, N.J. Repetitive motion analysis: segmentation and event classification. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 26, No. 2, pp. 258-263, 2004.

[Sob14a] Sobral, A., and Vacavant, A. A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos. Computer Vision and Image Understanding, Vol. 122, pp. 4-21, 2014.

[Sta00a] Stauffer, C., and Grimson, W.E.L. Learning patterns of activity using real-time tracking. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 22, No. 8, pp. 747-757, 2000.

[Stc15a] St-Charles, P.-L., Bilodeau, G.-A., and Bergevin, R. SuBSENSE: a universal change detection method with local adaptive sensitivity. IEEE Trans. on Image Processing, Vol. 24, No. 1, pp. 359-373, 2015.

[Wan14a] Wang, Y., Jodoin, P.-M., Porikli, F., Konrad, J., Benezeth, Y., and Ishwar, P. CDnet 2014: an expanded change detection benchmark dataset. Proc. of IEEE Workshop on Change Detection at IEEE Conf. on Computer Vision and Pattern Recognition, pp. 387-394, 2014.

[Ziv04a] Zivkovic, Z. Improved adaptive Gaussian mixture model for background subtraction. Proc. of Int. Conf. on Pattern Recognition, pp. 28-31, 2004.