LETTER

Robust Motion Detection Based on the Enhanced ViBe

Zhihui FAN[†a)], *Student Member*, Zhaoyang LU[†], Jing LI[†], Chao YAO[†], *and* Wei JIANG[†], *Nonmembers*

SUMMARY To eliminate casting shadows of moving objects, which cause difficulties in vision applications, a novel method is proposed based on Visual background extractor by altering its updating mechanism using relevant spatiotemporal information. An adaptive threshold and a spatial adjustment are also employed. Experiments on typical surveillance scenes validate this scheme.

key words: *motion detection, object detection, shadow removal, background substraction, surveillance*

1. Introduction

Detecting moving objects in video is an important procedure for vision applications, such as video surveillance and activity analysis. Shadows cast by moving objects are mostly connected to objects and significantly differs from the background, and they are frequently misclassified as part of foreground objects, which makes it difficult to detect their exact shapes. Therefore, eliminating shadows from foreground is crucial. Nowadays, it remains a challenging task to develop an efficient and robust motion-detecting algorithm. Recently, Visual background extractor (ViBe) algorithm [1] has attracted more and more attention due to its extremely high speed and better adaption to dynamic environments. Nevertheless, the performance of ViBe is poor in shadow scenes. Pixels and their neighbors usually change simultaneously over time, there exists a highly close relationship between them. Therefore, spatiotemporal information could be extracted to facilitate motion-detecting. A novel method called EViBe (Enhanced ViBe for shadow removing) is proposed based on ViBe by employing a different updating mechanism utilizing relevant spatiotemporal information, which is exploited by two factors: local difference of intensity (LD) and chrominance difference (CD).

2. The Proposed Method EViBe

ViBe models each pixel by storing a set of values collected in the past at the same location or in the neighborhood [1]. It then compares this set to the current pixel value to determine whether the pixel belongs to the background, and updates the model by choosing randomly which values would be substituted. As Fig. 1 (a) illustrated, only common sample pixels are included in the background model by ViBe, thus shadow pixels are likely to be misclassified as foreground according to its evaluation criteria; if we can adopt appropriate method to seek suspected shadow pixels and update them into background model timely, suspected shadow pixels in background would increase the chance that shadows pixels be correctly classified into background, as Fig. 1 (b) does. For EViBe only increases their chances without guaranteeing all shadows to be correctly classified. Additional steps are employed and the details are discussed in Sect. 3.

The main steps of EViBe are given in Algorithm 1.

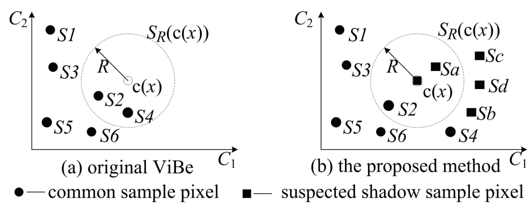

Fig. 1 Comparison between ViBe and the proposed method. In a 2-D color space $(C1, C2)$, S_R is the sphere with radius R centered on a pixel $c(x)$, and black points denote samples.

Algorithm 1 Steps of the proposed method EViBe.

Input: frame sequence(i=1,2,3...,N,...).
Output: foreground(moving objects).

```
for i = 1;i; i++ do
    Model each pixel by choosing sample values from neighborhood randomly.
    Pixel clustering by K-means using positional and chromatic information.
    Convert RGB space to HSV space.
    if i < N then
        Initialize LD and CD in single Gaussian.
        Detect moving objects and update background models by ViBe.
    else
        Seek suspected shadows by LD and CD .
        if suspectedshadowsappear then
            Add suspected shadow pixels into models of pixels belonging to the same cluster.
        else
            Update models by original ViBe.
        end if
        Detect moving objects by ViBe using an adaptive threshold.
    end if
    Spatial adjustment.
    return moving objects.
end for
```

Manuscript received January 15, 2014.
Manuscript revised May 6, 2015.
Manuscript publicized June 10, 2015.
[†]The authors are with State Key Lab. of Integrated Service Networks, Xidian University, Shaanxi, 710071 China.
a) E-mail: fanzhihui0379@126.com
DOI: 10.1587/transinf.2015EDL8013

Suspected shadow pixels would be updated into background models appropriately in algorithm 1. The key is to seek the suspected shadow pixels, which is finished by the following two steps. First *LD* is employed to obtain outliers probably containing real foreground pixels and shadow pixels. Then, *CD* is utilized to seek suspected shadow pixels from above mentioned outliers.

Chrominance and luminance information can be effectively separated in HSV color space. Hence value and hue in HSV color space are employed as intensity and chrominance respectively. We designate *LD* as cluster intensity difference. K-means clustering with a chromatic and positional feature is used to divide images into several pixel clusters, and we set the size of cluster (SoC) 40*40. *LD* is defined as the intensity difference between the target pixel and its cluster. The represent intensity of cluster (RoC) for pixel $c(x)$ is expressed in (1).

$$\mathrm{RoC(x)}= \frac{1}{N+M}\left(\sum_{i=1}^{N} V_i + \sum_{q\in C}^{M} V_q\right). \tag{1}$$

where C is a pixel cluster, M is the number of pixels belonging to C, N is the number of past frames,V_q and V_i are the values of pixel q and pixel i in HSV color space respectively. Gaussian model is applied to *LD* and Hue respectively, and two set of pixels: $O1$ and $O2$ are obtained which do not obey. Experiments indicate that single Gaussian model is suitable to *LD* and Hue which sharing the same reliability as 99.7%, $\mathrm{P}(-2.58 < Z < 2.58) = 0.9973$, $\mathrm{Z = N(0,1)}$. Then we get suspected shadow pixels by applying a XOR operator to $O1$ and $O2$ as (2). Where S is the set of suspected shadow pixels, and $c(x)$ is the current pixel.

$$\begin{cases} c(x) \in S, if\ c(x) \in O1\ and\ c(x) \notin O2, \\ \qquad\qquad\qquad\qquad\qquad otherwise. \end{cases} \tag{2}$$

3. Spatial Adjustment and Adaptive Threshold

After the above operations, there might be some small background regions locating in foreground and a few sporadic shadows (misclassified as foreground) locating in background. EViBe eliminates these small background regions by a spatial adjustment, and empirically the small region is defined as a size which is less than 3% of the size of the estimated moving object region. These sporadic shadows are eliminated through a postprocessing (detailed in Sect. 4). Moreover, EViBe adopts an adaptive threshold instead of a fixed threshold in ViBe. The standard deviation σ of the samples in a model is calculated and a matching threshold as $0.5{*}\sigma$ bounded to [30,50] interval is employed.

4. Experiments

To validate EViBe, experiments on three typical surveillance scenes which are publicly available in [2] containing moving cast shadows are carried out, comparing with mixture of Gaussian (MoG) [4], ViBe [1] and PKDE [3]. All tested algorithms run on a PC with 2.8 GHz, 1GB memory and XP. MoG, ViBe and PKDE utilize parameters suggested by their authors. EViBe adopts a set of consistent parameters: number of samples per pixel $\eta = 20$, frequency of random updating $\theta = 16$, number of close samples for being part of background $\lambda = 2$, and adaptive threshold $R = 0.5 * \sigma$. For all algorithms, the only postprocessing is to eliminate small pieces less than 15 pixels. Fig. 2 demonstrates segmentation results. As illustrated, MoG and ViBe fail to remove most casting shadows of objects, PKDE loses the complete contours of objects and cannot absorb shadows completely in the "busStation" scene. In contrast, EViBe could perfectly eliminate the casting shadows and obtain more exact shapes of moving objects.

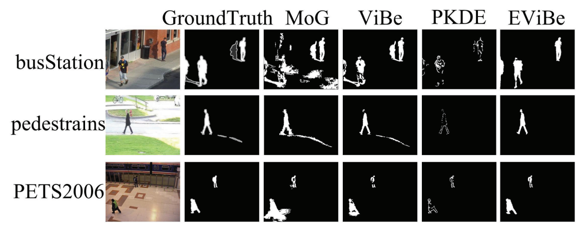

Fig. 2 Performance for comparing methods.

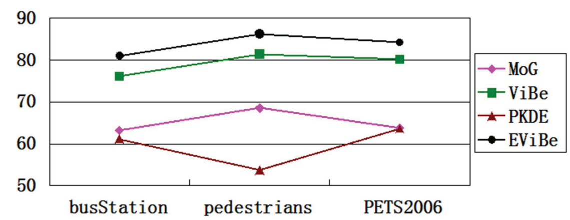

Fig. 3 Performance of F-score(%) on typical surveillance scenes.

We compare the F-score detailed in [3] of all compared algorithms. The result is shown in Fig. 3. It is evident that EViBe has the highest F-score. The running framerate (fps, frames per second) is also tested on the "pedestrians" dataset with the frame size of 360 * 240. The frame rate of 31.8 fps, 65.4 fps, 42.7 fps, and 40.3 fps are achieved for MoG, ViBe, PKDE and EViBe respectively. EViBe can run in real time.

5. Conclusions

A robust motion detection algorithm called EViBe is presented by altering the updating mechanism of ViBe employing relevant spatiotemporal information, which is exploited by two factors *CD* and *LD*. An adaptive threshold and a spatial adjustment step are also adopted. Experimental results validate this approach.

Acknowledgments

This work was supported by the following funding resource: Fundamental Research Funds for the Central Universities, K50510010007.

References

[1] O. Barnich and M. Van Droogenbroeck, "ViBe: A universal background subtraction algorithm for video sequences," IEEE Trans. Image Process., vol.20, no.6, pp.1709–1724, 2011. doi: 10.1109/TIP.2010.2101613

[2] A video database for testing change detection algorithms: http:wordpress-jodoin.dmi.usherb.ca/dataset2014/, accessed Aug. 2014.

[3] S. Liao, G. Zhao, V. Kellokumpu, et al., "Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes," Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on IEEE, San Francisco, California, USA, pp.1301–1306, June 2010. doi:10.1109/CVPR.2010.5539817

[4] J. Stander, R. Mech, and J. Ostermann, "Detection of moving cast shadows for object segmentation," IEEE Trans. Multimedia, vol.1, no.1, pp.65–76, 1999. doi:10.1109/6046.748172