UDC 004.424.6, DOI: 10.2298/CSIS1001201D

Robust moving object detection under complex background

DING Ying [1,2], LI Wen-hui [1], FAN Jing-tao [2], and YANG Hua-min [2]

[1] School of Computer Science and Technology, Jilin University, Changchun 130012, China
liwenhui2050@163.com
[2] School of Computer Science and Technology, Changchun University of Science and Technology, Changchun 130022, China
{dingying, fanjingtao, yhm}@cust.edu.cn

Abstract. We present a novel method to robustly and efficiently detect moving object, even under the complexity background, such as illumination changes, long shadows etc. This work is distinguished by three key contributions. The first is the integration of the Local Binary Pattern texture measure which extends the moving object detection work for light illumination changing. The second is the introduction of HSI color space measure which removes shadows for the background subtraction. The third contribution is a novel fuzzy way using the Choquet integral which improves detection accuracy. The experiment results using several dataset videos show the robustness and effectiveness of the proposed method.

Keywords: moving object detection, Local Binary Pattern, HSI, Choquet integral.

1. Introduction

Background subtraction is often one of the first tasks in machine vision applications, making it a critical part of the system. The output of background subtraction is an input to a higher-level process that can be, for example, the tracking of an identified object. The performance of background subtraction depends mainly on the background modeling technique used to model the scene background. Especially natural scenes put many challenging demands on background modeling since they are usually dynamic in nature including illumination changes, swaying vegetation, rippling water, flickering monitors etc. A robust background modeling algorithm should also handle situations where new stationary objects are introduced to or old ones removed from the scene. Furthermore the shadows of the moving and scene objects can cause problems. Even in a static scene frame-to-frame changes can occur due to noise and camera jitter.

In this paper, we proposed a novel model for background maintenance and subtraction. The model aggregates color and texture features using fuzzy approach. The goal of the new model was to address all of the above-mentioned difficulties. Our contributions are: (1)Extend the background subtraction work for light illumination changes by integrating Local Binary Pattern (LBP) texture measure, (2)Remove shadows for the background subtraction by using HSI color measure, (3) Improve detection accuracy in a fuzzy way using the Choquet integral.

2. Related work

Different kinds of background model for detecting moving objects have been proposed in the literature, and some of them can be found to be robust to the challenges met in video sequence. These different methods are classified following the model used:

Basic Background Modeling (BBM): In this case, Background Representation is modeled using the average [1] or the median [2] or the histogram analysis over time [3]. Once the model is computed, the foreground detection is made as follows:

$$\left| f(I_t(x,y)) - f(B_t(x,y)) \right| > Th \tag{1}$$

Otherwise, pixels are classified as background. Where This a constant threshold, $I_t(x, y)$ and $B_t(x, y)$ are respectively the current and the background images at time t, $f(x)$ is a feature value of x, such as intensity, gradient, etc.

Statistical Background Modeling (SBM): Background Representation is modeled using a single Gaussian [4-6] or a Mixture of Gaussians [7, 8] or a Kernel Density Estimation [9, 10]. Statistical variables are used in the foreground detection to classify the pixels as foreground or background. Recent SBM use Generalized Gaussian Mixture Modeling [11], Bayesian approaches [12, 13], Support Vector Regression learning approaches [14] or Codebook [15-17].

Background Estimation (BE): Background representation is estimated using a filter. For the foreground detection, any pixel of the current image that deviates significantly from its predicted value is declared foreground. This filter may be a Wiener filter [18], a Kalman filter [19] or a Tchebychev filter [20].

All these methods present the same following steps and issues: background modeling, background initialization, background maintenance, foreground detection, choice of the picture's element (pixel, a block or a cluster), choice of the features which characterize the picture's element (color features, edge features, stereo features, motion features and texture features). Often, these features are used separately and the most used is the color one. The combination of several measuring features can strengthen the pixel's classification as background or foreground. In a general way, the Choquet and Sugeno integrals have been successfully applied widely in classification problems [21], in decision making [22] and also in data modeling

[23] to aggregate different criteria. In the context of foreground detection, these integrals seem to be good model candidates for fusing different measures from different features. Each integral has its particularity. The Choquet integral requires to interpret the scale as a continuum and the Sugeno integral allows to work with an ordinal scale. Recently, Zhang and Xu [24] have used texture feature and color features to compute similarity measures between current and background pixels. Fida EL BAF[25] has fuzzy intensity and texture feature to foreground detection for infrared videos.

3. New approach

3.1. Approach overview

Moving object detection is based on a comparison between current and background images. In general, a simple subtraction is made between these two images to detect regions corresponding to moving object. So the choice of the features which characterize the pixel element is one of the most important steps. The other one is how to establish the comparison consists in defining a similarity measure between pixels in current and background images.

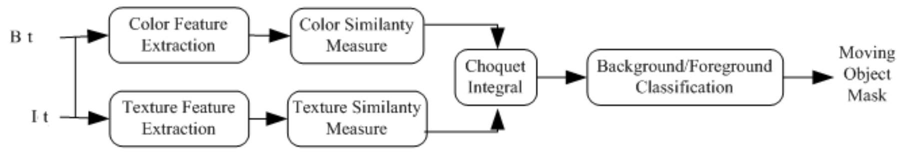

Fig. 1. Our approach overview

In this paper, we define a similarity measure between pixels in current and background images. In this case, pixels corresponding to background should be similar in the two images while pixels corresponding to foreground should not be similar. In Figure 1, the moving object detection process is presented in details. First, the color and texture features are extracted from the background image B_t and the current image I_t. The similarity measures are computed for each feature which is then aggregated by the Choquet integral. The Background/Foreground classification is finally made by threshold the Choquet integral's result. In the following subsections, we describe the rationale for selecting and fusing the set of the adopted features.

3.2. Color feature similarity measure

In order to remove the shadows' disturbance, pixels that could be part of a shadow have to be identified. RGB is the color space commonly acquired directly from a sensor or camera. HSI and YCbCr are closer to human interpretation of colors in the sense that brightness, for intensity, is separated from the base color. The best feature should decrease their sensitive to shadows. We choose HSI color space [26], and define the color features with HSI three components noted C_1, C_2 and C_3. Then, the color similarity measure $S_k(x, y)$ at the pixel (x, y) is computed as:

$$S_k(x, y) = 1 - \frac{|I_k(x,y) - B_k(x,y)|}{255} \tag{2}$$

Where $k \in \{1,2,3\}$ is one of three color features. $B(x, y)$ and $I(x, y)$ respectively represent the background and current frame at time t. Note that $S_k(x, y)$ is between 0 and 1. Furthermore, $S_k(x, y)$ is close to one if $B_k(x, y)$ and $I_k(x, y)$ are very similar.

3.3. Texture feature similarity measure

The proposed texture-based method for background subtraction is based on the Local Binary Pattern (LBP) texture measure. The LBP is a powerful means of texture description [27-29]. The operator labels the pixels of an image block by threshold the neighborhood of each pixel with the center value and considering the result as a binary number (LBP code):

$$LBP(x_c, y_c) = \sum_{i=0}^{K-1} f(p_i - p_c) 2^i \tag{3}$$

Where p_c corresponds to the pixel value of the center pixel (x_c, y_c), such as gray, intensity value etc. and p_i to the pixel values of the K neighborhood pixels. The function $f(x)$ is defined as follows:

$$f(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{4}$$

Then, for each pixel texture in the current image and the background image, the LBP code is less sensitive to illumination changes and is able to derive an accurate local texture difference measure [27].To avoid light illumination changes' affect, here we define a texture similarity measure at pixel (x,y) between the current image and the background image as:

$$S_T(x, y) = 1 - \frac{|I_{LBP}(x,y) - B_{LBP}(x,y)|}{255} \tag{5}$$

Where $I_{LBP}(x,y)$ and $B_{LBP}(x,y)$ are respectively denotes the texture LBP code of pixel (x, y) in the background and current images. Note that $S_T(x,y)$ is between 0 and 1. Furthermore, $S_T(x,y)$ is close to one if $I_{LBP}(x,y)$ and $B_{LBP}(x,y)$

are very similar. In the false positive foreground areas caused by quick lighting changes, there are no texture changes between the current frame and the background. Hence, $S_T(x,y) \approx 1$. The foreground mask will be removed for the areas with $S_T(x,y) \geq T_s$. For this operation, we have chosen the choquet integrals.

3.4. Aggregation of Features by Choquet Integrals

Many fusion techniques can be used to fuse the color and the texture features. We present brief necessary concepts around fuzzy measures and the Choquet integrals [30].

Let λ be a fuzzy measure on a finite set X, and non-additive measure on a subset of X is any function $\mu: X \rightarrow [0, 1]$.

Definition 1 The Choquet integral of μ with respect to λ is defined by:

$$C_\lambda = \sum_{i=1}^{n} (\mu(x_{\sigma(i)}) - \mu(x_{\sigma(i-1)}))\lambda(A_{\sigma(i)}) \qquad \textbf{(6)}$$

Where finite set X = {$x_1,\ldots,x_n$}, and σ is a permutation of the indices such that $\mu_{\sigma(1)} \leq \ldots \leq \mu_{\sigma(n)}$ and $A_{\sigma(i)} = \{\sigma(i),\ldots,\sigma(n)\}$.

As defined above, the computed measures are obtained by dividing the feature values in background and current image with endpoints denoted by 0 and 1. For each pixel, color and texture similarity measures are computed as formula (2) (5) from the background and the current frame. We define the set of criteria $X = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ with $(\alpha_1, \alpha_2, \alpha_3)$ = three components color features of the chosen HSI color space and α_4 = texture feature $LBP(x, y)$.

With each criterion, we associate a fuzzy measure, for each α_i, let $\lambda(\alpha_i)$ be the degree of importance of the feature x_i in the decision whether pixel corresponds to background or foreground. Define $\lambda(\alpha_1) = \lambda(\{\alpha_1\})$, $\lambda(\alpha_2) = \lambda(\{\alpha_2\})$, $\lambda(\alpha_3) = \lambda(\{\alpha_3\})$, and $\lambda(\alpha_4) = \lambda(\{\alpha_4\})$ such that the higher the $\lambda(\alpha_i)$, the more important the corresponding criterion in the decision. To compute the fuzzy measure of the union of any two disjoint sets whose fuzzy measures are given, we use an operational version proposed by Sugeno [30] which called λ-fuzzy measure. To avoid excessive notation, let denote this measure by λ', where λ' is a parameter of the fuzzy measure used to describe an interaction between the criteria that are combined. Its value can be determined through the boundary condition, i.e. $\lambda(X) = \lambda(\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}) = 1$. The fuzzy density values over a given set $K \subseteq X$ are computed as:

$$\lambda(K) = \frac{1}{\lambda'}[\prod_{x_i \in K}(1 + \lambda' \cdot \lambda(x_i)) - 1] \qquad \textbf{(7)}$$

The fuzzy function $\mu(\alpha_i)$ are defined in $[0,1]$ so that, $\mu(\alpha_1) = S_1(x, y)$, $\mu(\alpha_2) = S_2(x, y)$, $\mu(\alpha_3) = S_3(x, y)$ and $\mu(\alpha_4) = S_T(x, y)$. To compute the value of Choquet integral for each pixel, we need firstly to rearrange the features α_i in the set X with respect to the order: $\mu(\alpha_1) \leq \mu(\alpha_2) \leq \mu(\alpha_3) \leq \mu(\alpha_4)$.

The pixel at position *(x, y)* is considered as foreground if its Choquet integral value is less than a certain threshold $T_{c,t}$, which denote the threshold at time instant *t*, as follows:

If $C_{\mu,t}(x, y) < T_{c,t}(x, y)$ then pixel *(x, y)* is foreground or moving object, else background.

4. Experimental results

The performance of the proposed method was evaluated using several video sequences. Both indoor and outdoor scenes were included. We have compared our method with the improved GMM modeling. Algorithms were implemented under Microsoft Visual C++ using the OpenCV library. The experimental results demonstrate the robustness of our algorithm in complex environments.

4.1. Experiments on indoor dataset

(a)background model (b)current frame (c)GMM detect result (d)our method detect result

Fig. 2. Moving object analysis of a indoor test sequence with a person come in

Fig. 2 compares a moving object detection result on the indoor test sequence from Wallflower [31], where a person is walking in a room, by GMM algorithm and our approach. Fig. 2a is the background model, and Fig. 2b is frame 650 (random choose) which after a person come in the office. The absolute color components change greatly with the illumination, even when no foreground object is present for the light changing. In Fig. 2c, large areas of false positive foreground were detected by the GMM method for light illumination change. As mentioned above, LBP is invariant to monotonic changes in gray scale. This makes it robust against illumination changes; Fig. 2d shows that our method successfully handles the light illumination changes by integrating texture information.

4.2. Experiments on outdoor dataset

Figure 3 shows the results of our algorithm for the outdoor test sequence, which contains changing environment and shadow. The original sequence has been taken from the PETS database [32] where several persons are walking in a subway station. The proposed algorithm successfully handles this situation. In HSI color space, the feature value changes of pixels in shadow region are very small, so most of the shadows can be removed by integrating HSI color information

(a)background frame (b)current frame (c)GMM detect result (d)our method detect result

Fig. 3. Moving object analysis of an outdoor test sequence

4.3. Experiments on detection accuracy

To see the progression of the performance of each algorithm, we compute the true positive rate (TPR) and the false positive rate (FPR) as follows:

Let *A* be the ground truth point set and *B* be a detected region, the TPR and FPR can be defined as equation 8.

$$TPR = \frac{\sum_{(x,y)\in(A\cap B)} 1}{\sum_{(x,y)\in A} 1}, \quad FPR = \frac{\sum_{(x,y)\in(\bar{A}\cap B)} 1}{\sum_{(x,y)\in \bar{A}} 1} \qquad (\mathbf{8})$$

Table 1 Detection results comparison

experiments	GMM Method		Our Method	
	TPR	FPR	TPR	FPR
1	1.00	0.056	1.00	0.00
2	0.958	0.012	0.968	0.003
3	0.945	0.011	0.947	0.003
4	0.945	0.011	0.951	0.004
5	0.957	0.012	0.960	0.004
6	0.952	0.012	0.968	0.004
7	0.940	0.012	0.949	0.004
8	0.948	0.021	0.950	0.001
9	0.943	0.020	0.959	0.008
10	0.989	0.237	0.985	0.002

Where *(x, y)* is a point in image. The TPR is the proportion of moving object pixels that were correctly classified among all positive samples. And FPR is the proportion of background pixels that were erroneously reported as being moving object pixels. In several experiments, the TPR and FPR result of different methods are compared as shown in Table 1. We have noticed that the TPR is very similar between these two methods, but the FPR of our method is quite lower than GMM's.

5. Conclusion

In this paper, we have presented a novel fuzzy background model for detecting moving objects from video frames. This method using Choquet integral for fuse color features and texture features. It chooses HSI color space instead of RGB, which remove most of the shadow, and aggregates LBP texture feature, which compute easily, to adapt the light illumination change. The proposed algorithm was tested against several standard benchmarks including both indoor and outdoor scenes. Further, the experiments results show that the proposed method is more robust and accurate.

6. References

1. B. Lee and M. Hedley: Background estimation for video surveillance. Image and Vision Computing New Zealand. (2002)
2. N. McFarlane and C. Schofield: Segmentation and tracking of piglets in images. British Machine Vision and Applications, Vol.8, No. 3, 187-193. (2005)
3. J. Zheng and Y. Wang: Extracting roadway background image: A mode based approach. Transportation Research Board, TRB, 82-88. (2006)
4. A. Francois and G. Medioni: Adaptative color background modeling for real-time segmentation of video stream. Proceedings of the International Conference on Imaging Science, Systems, and Technology, ICISST. (1999)
5. C. Wren and A. Azarbayejani. Pfinder: Real-time tracking of the human body. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.19, No. 7, 780 –785. (1997)
6. M. Zhao, N. Li, and C. Chen: Robust automatic video object segmentation technique. IEEE International Conference on Image Processing ICIP. (2002)
7. C. Stauffer: Adaptive background mixture models for realtime tracking. Proceedings IEEE Conference on Computer Vision and Pattern Recognition, CVPR. (1999)
8. R. Tan, H. Huo, J. Qian, and T. Fang: Traffic video segmentation using adaptive-k gaussian mixture model. The International Workshop on Intelligent Computing, IWICPAS. (2006)
9. A. Elgammal and L. Davis: Non-parametric model for background subtraction. 6th European Conference on Computer Vision, ECCV. (2000)
10. C. Ianasi, V. Gui, C. Toma, and D. Pescaru: A fast algorithm for background tracking in video surveillance, using nonparametric kernel density estimation.

Facta Universitatis, Series : Electronics and Energetics, Vol.18, No. 1, 127–144.(2005)
11. M. Allili, N. Bouguila, and D. Ziou: A robust video foreground segmentation by using generalized gaussian mixture modeling. Fourth Canadian Conference on Computer and Robot Vision, CRV. (2007)
12. F. Porikli and O. Tuzel: Bayesian background modeling for foreground detection. ACM InternationalWorkshop on Video Surveillance and Sensor Networks, VSSN. (2005)
13. Y. Sheikh and M. Shah: Bayesian modeling of dynamic scenes for object detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI. Vol.27, No. 11, 1778–1792. (2005)
14. J. Wang, G. Bebis, M. Nicolescu, M. Nicolescu, and R. Miller: Improving target detection by coupling it with tracking. Machine Vision and Application, MVA. (2008)
15. A. Doshi and M. Trivedi: Hybrid cone-cylinder codebook model for foreground detection with shadow and highlight suppression. AVSS. (2006)
16. K. Kim, T. H. Chalidabhongse, D. Harwood, and L. Davis: Real-time foreground-background segmentation using codebook model. Real-Time Imaging. Vol.11, No. 3, 172-185. (2005)
17. A. Leykin and R. Hammoud: Robust multi-pedestrian tracking in thermal-visible surveillance videos. OTCBVS. (2006)
18. K. Toyama and J. Krumm: Wallflower: Principles and practice of background maintenance. International Conference on Computer Vision ICCV. (1999)
19. S. Messelodi and C. Modena: A kalman filter based background updating algorithm robust to sharp illumination changes. International Conference on Image Analysis and Processing, ICIAP. (2005)
20. R. Chang, T. Ghandi, and M. Trivedi: Vision modules for a multi sensory bridge monitoring approach. IEEE Conference on Intelligent Transportation Systems, ITS. (2004)
21. H. Tahani and J. Keller: Information fusion in computer vision using the fuzzy integral. Transaction on Systems, Man, and Cybernetics, Vol.20, No. 3, 733–741. (1990)
22. Y. Narukawa and T. Murofushi: Decision modelling using the choquet integral. Modeling Decisions for Artificial Intelligence, Vol.31, No. 31, 183–193. (2004)
23. M. Sugeno and S. Kwon: A new approach to time series modeling with fuzzy measures and the choquet integral. 4th IEEE International Conference on Fuzzy Systems, ICFS. (1995)
24. H. Zhang and D. Xu: Fusing color and texture features for background model. Third International Conference on Fuzzy Systems and Knowledge Discovery, Vol. 4223, No. 7, 887–893. (2006)
25. Fida EL BAF: Thierry BOUWMANS and Bertrand VACHON: Fuzzy Foreground Detection for Infrared Videos. Computer Vision and Pattern Recognition Workshops, Anchorage, AK, 1-6. (2008)
26. Ribeiro H, Gonzaga A: Hand image segmentation in video sequence by GMM: A comparative analysis. 19th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI 2006), Brazil, 357-364. (2006)
27. M. Heikkil¨a, M. Pietik¨ainen and J. Heikkil: A Texture-based Method for Detecting Moving Objects. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.28, No. 4, 657-666. (2006)
28. T. Ojala, M. Pietik¨ainen, and D. Harwood: A comparative study of texture measures with classification based on feature distributions. Pattern Recognition, Vol. 29, No. 1, 51–59. (1996)

29. T. Ojala etc.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 7, 971–987. (2002)
30. M. Grabisch, T. Murofushi, and M. Sugeno: Fuzzy Measures and Integrals. Theory and Applications (edited volume). Studies in Fuzziness. Physica Verlag. (2000)
31. Wallflower Dataset: http://research.microsoft.com/users/jckrumm/WallFlower/TestImages.htm
32. PETS Dataset. http://pets2006.net.

Ding Ying is studying at Jilin University, she got Master Degree in 2007, at Changchun University of Science & Technology. Her main directions are image processing and computer simulation.

Li Wen-hui got his Ph.D. in 1994 from BREMEN University in Germany. He is currently a Professor of Computer Science at Jilin University.

Fan Jing-tao is studying at Changchun University of Science & Technology, he got Master Degree in 2007, at Changchun University of Science & Technology. His main directions are image processing and computer simulation.

Yang Hua-min got his Ph.D. in 2007 from Changchun University of Science & Technology in China. He is currently a Professor of Computer Science at Jilin University.

Received: April 13, 2009; Accepted: August 27, 2009.