2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

Moving Vehicle Detection Based on Fuzzy Background Subtraction

Xiaofeng Lu, Takashi Izumi, Tomoaki Takahashi, and Lei Wang

Abstract—**Background subtraction is a method typically used to segment moving regions in image sequences taken from a static camera by comparing each new frame with a model of the background scene. This paper proposes a novel fuzzy background subtraction algorithm for moving vehicle detection which achieves the high detection rates, and reduces the influence of illumination changes and shadows in the traffic scene. The proposed method adopts the Choquet integral for fusion the similarity measures of three color components of the YCbCr color space and uniform local binary pattern texture. Otherwise, an adaptive selective method for background maintenance is proposed to address the problem of background pollution. The experimental results of several dataset videos show the robustness and effectiveness of the proposed method.**

I. Introduction

VISUAL traffic surveillance system provides most efficient traffic information for traffic control and management, assistance for safe driving in Intelligent Transport Systems (ITS). Moving vehicle detection is an important research area in ITS and is also a major application of computer vision. Background subtraction is often one of the first tasks executed in computer vision applications, making it a critical part of the system. The output from background subtraction is the input to higher level processes, for example, vehicle tracking and vehicle identification. Traditional background subtraction is mainly divided into three phases: background initialization, foreground detection, and background maintenance. Therefore, the performance of background subtraction depends mainly on the techniques used in these three phases. For the complex conditions, such as illumination changes, shadows, and go-and-stop vehicles, background subtraction algorithm should handle these situations where objects are introduced or removed from the scene dynamically. Moreover, the algorithm should minimize the computational complexity and operate in real-time.

Numerous different methods for moving target detection have been proposed. A simple way to model the background is to acquire a background image that does not include any moving object or to calculate an average image [1] or median image [2] of the scene, subtract each new video frame from it, and threshold the result. These background modeling methods have low accuracy and are not adapted to traffic congestion situations. One way to statistically represent the background is to assume that the time history of intensity values of a pixel can be modeled by a Single Gaussian (SG) approach [3][4]. However, this model does not work well in the case of dynamic natural environments. To solve this problem, the Gaussian Mixture Model (GMM) has been used to model dynamic backgrounds [5][6][7][8]. But, for the rapidly changing background, these methods cannot be accurately modeled with just a few Gaussian distributions, and cause problem in achieving sensitive detection [9]. To deal with the limitation of parametric methods, a non-parametric technique using Kernel Density Estimation (KDE) was developed for building a statistical representation of the background scene [10]. The probability density function for pixel intensity is estimated directly from the data without any assumptions about the underlying distributions.

In this paper, we propose a novel method for foreground detection and background maintenance. The proposed method adopts the YCbCr color feature and texture feature using fuzzy approach. This method fuses three components of the YCbCr color space and uniform local binary pattern texture by using the Choquet integral to reduce the influence of shadows and illumination changes. And we propose an adaptive selective method for background maintenance to address the problem of background pollution.

Xiaofeng Lu is with Graduate School of Science and Technology, Nihon University, Chiba, Japan (e-mail: csgy11004@g.nihon-u.ac.jp).

Takashi Izumi is with College of Science and Technology, Nihon University, Chiba, Japan (e-mail: izumi@ecs.cst.nihon-u.ac.jp).

Tomoaki Takahashi is with College of Science and Technology, Nihon University, Chiba, Japan (e-mail: takahashi.tomoaki@nihon-u.ac.jp).

Lei Wang is with Faculty of Computer Science and Engineering, Xi'an University of Technology, Xi'an, China (e-mail: leiwang@xaut.edu.cn)

This work was supported by the National Natural Science Foundation of China (Grant No. 61272283) and the specialized research program of the educational office of Shaanxi province (No. 09JK645).

II. Proposed algorithm

A. General description

Background subtraction commonly includes the three phases, such as background reconstruction, foreground detection, and background maintenance. In the initialization phase, we adopt the color statistical background model in the YCbCr color space to reduce the influence of complex background. Similar method is described in the literature [14]. For the foreground detection, we firstly define a similarity measure between pixels in current and background images. The pixels corresponding to background should be similar in the two images while pixels corresponding to foreground should not be similar. Based on this understanding, we propose a new foreground detection method for fusing the similarity measures by the Choquet integral. Otherwise, we propose an adaptive selective background maintenance method for the complex conditions.

B. Color and Texture Similarity Measures

1) Color feature: RGB color is directly available from the sensor or the camera. However, the RGB color space has an obvious drawback: these three color components are reciprocally dependent, which increase the sensitivity to illumination changes. This causes great difficulties in shadow suppression. A number of color space comparisons are presented in the literatures [13][14][15]. Though analyzing the experimental results, we select the YCbCr color space for

978-1-4799-2072-3/14/$31.00 ©2014 IEEE

background model and foreground detection because of its suitability in reducing the influence of shadows and illumination changes.

The color similarity measure $S_k(x,y)$ at pixel (x,y) is defined as

$$S_k(x,y) = \begin{cases} \frac{C_k^I(x,y)}{C_k^B(x,y)} & if \quad C_k^I(x,y) < C_k^B(x,y) \\ 1 & if \quad C_k^I(x,y) = C_k^B(x,y) \\ \frac{C_k^B(x,y)}{C_k^I(x,y)} & if \quad C_k^I(x,y) > C_k^B(x,y) \end{cases} \quad (1)$$

where $k \in \{1,2,3\}$ is one of the color components *Y, Cb* and *Cr*. $C_k^I(x,y)$ and $C_k^B(x,y)$ represent the color values of the current frame and background frame at time *t*, respectively. Note that $S_k(x,y)$ takes a value between 0 and 1. $S_k(x,y)$ is close to 1 if $C_k^I(x,y)$ and $C_k^B(x,y)$ are very similar. Similar method is described in the literature [12].

2) ULBP feature: The presented texture extraction method is based on the uniform local binary pattern (ULBP). The ULBP operator improves the rotation invariance of the local binary pattern (LBP) code and quantifies the occurrence statistics of individual rotation-invariant patterns corresponding to certain micro-features in the image. The LBP operator has 2^k modes, whereas the ULBP operator has only *k*+2 modes, reducing the number of histogram entries, making the bins more discrete and the histogram less susceptible to noise interference [16].

The ULBP texture measure is computed as follows:

$$ULBP_{K,R}^{ri} = \begin{cases} \sum_{i=0}^{K-1} f(p_i - p_c) & if \quad U(LBP_{K,R}^{ri}) \le 2 \\ K+1 & otherwise \end{cases} \quad (2)$$

where

$$LBP_{K,R}^{ri} = min\{ROR(LBR_{K,R}, i) | i = 0,1,\cdots,K-1\} \quad (3)$$

The condition function is defined as:

$$U(LBP_{K,R}^{ri}) = |f(p_{K-1} - p_c)| + \sum_{i=1}^{K-1} |f(p_i - p_c) - f(p_{i-1} - p_c)| \quad (4)$$

where p_c corresponds to the pixel value of the center pixel (x_c, y_c), such as gray level, intensity value, etc., and $p_i$denotes the pixel values of the K neighborhood pixels on a circle of radius *R*. The function *f (x)* is defined as:

$$f(x) = \begin{cases} 1 & if \quad x \ge 0 \\ 0 & if \quad x < 0 \end{cases} \quad (5)$$

The texture similarity measure $S_T(x,y)$ at the pixel (x,y) is defined as follows:

$$S_T(x,y) = \begin{cases} \frac{I_T(x,y)}{B_T(x,y)} & if \quad I_T(x,y) < B_T(x,y) \\ 1 & if \quad I_T(x,y) = B_T(x,y) \\ \frac{B_T(x,y)}{I_T(x,y)} & if \quad I_T(x,y) > B_T(x,y) \end{cases} \quad (6)$$

where $I_T(x,y)$ and $B_T(x,y)$ denote the texture ULBP code of pixel (x,y) in the current and background frames, respectively. Note that $S_T(x,y)$ is close to 1 if $I_T(x,y)$ and $B_T(x,y)$ are very similar.

C. Fuzzy measure and the Choquet integral

In this experiment, we adopt the Choquet integral to fuse the four similarity measures.

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

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

$$C_h = \sum_{i=1}^{n} (\mu(x_{\sigma(i)}) - \mu(x_{\sigma(i-1)}))h(A_{\sigma(i)}) \quad (7)$$

where σ is a permutation of the indices such that $0 \le \mu(x_{\sigma(1)}) \le \mu(x_{\sigma(2)}) \le \cdots \le \mu(x_{\sigma(n)}) \le 1$, $X = \{x_1, x_2, \cdots, x_n\}$ and $A_{\sigma(i)} = \{\sigma(1), \sigma(2), \cdots, \sigma(n)\}$.

For each pixel, a similarity measure is computed in different dimensions from the background and current frame. We define the set of criteria $X = \{x_1, x_2, x_3, x_4\}$ with $\{x_1, x_2, x_3\}$ being the three color-component features of the chosen color space, and x_4 being texture features obtained from the ULBP code. For every x_i, let $h(x_i)$ be the importance that the feature x_i takes in the decision of the foreground detection process. The fuzzy functions $\mu(x_i)$ are defined in [0, 1], so that $\mu(x_k) = S_k(x,y)$(where $k \in \{1,2,3\}$) and $\mu(x_4) = S_T(x,y)$. To compute the value of the Choquet integral for each pixel, we first use permutation function σ to rearrange the features x_i in the finite set X with respect to the order: $\mu(x_1) \le \mu(x_2) \le \mu(x_3) \le \mu(x_4)$. The fuzzy measure of different feature is obtained by numerous experimental results.

The pixel (x,y) is considered as the foreground if its Choquet integral value is less than a predetermined threshold TH, as follows:

$$\text{If } C_h(x,y) < TH \text{ then } (x,y) \text{ is foreground} \quad (8)$$

D. Background maintenance

Because of the complex conditions, such as go-and-stop vehicles, illumination changes and shadows, the background frame needs to update dynamically for the accuracy detection. Therefore, the background maintenance process is a critical step in moving target detection. Background maintenance determines how the background adapts itself to take into account the critical situations that can occur. Through analyzing the characteristics of traditional blind and selective background maintenance methods, we propose an adaptive selective background maintenance method based on the literatures [11][21]. Our proposed method solves the traditional problem, such as ghosts, etc., and effectively suppresses the problem of background pollution. Specially, the values of pixels classified as the foreground are taken into account the influence between the current frame and current background.

The adaptive background maintenance algorithm is defined as follows:

$$CB_{n+1} = a_n \times IB_n + (1 - a_n) \times CB_n \tag{9}$$

where CB_n is the current background, and IB_n is an instantaneous background that is computed as follows:

$$IB_n(x,y) = \begin{cases} F_n(x,y) & if\ MP(x,y) = 0 \\ CB_n(x,y) & if\ MP(x,y) = 1 \end{cases} \tag{10}$$

where a_n is a variable learning rate in the interval [0,1]. The value of a_n is based on changes in frame structure and illumination between the current frame and the background frame, and can be computed as follows:

$$a_n = 0.9 \times a_{n-1} + 0.1 \times a\_inst_n \tag{11}$$

In theory, when there are rapid changes in illumination, the weight a_n should be set to a higher value, and when the changes are slow, it should be set to a lower value, because a high weight allows adaptation to rapid changes in illumination, and a low weight reduces the effect of moving targets on background estimation.

Here, a_inst_n is defined as an adaptive weight on illumination normalization between frame F_n and frame F_{n-1}, as follows:

$$a\_inst_n = \frac{sum\_unmov_{n,n-1}}{area\_unmov_{n,n-1}} \tag{12}$$

Because the moving objects' coverage areas in this scene do not reflect illumination changes, the weight on illumination normalization is only calculated in non-moving target areas:

$$area\_unmov_{n,n-1} = \sum_{(x,y)} (1 - MP(x,y)) \tag{13}$$

where $MP(x,y) \in MP_n \cup MP_{n-1}$.

Here, $area\_unmov_{n,n-1}$ represents the number of pixels of the non-moving area in the current frame. In other words, it is the area defined by the current frame minus the union of MP_n and MP_{n-1}, and MP_n and MP_{n-1} are the moving pixels in the frame F_n and the frame F_{n-1}, respectively. They are binary map, where a target pixel value is 1, and a non-target pixel value is 0.

Here, $sum\_unmov_{n,n-1}$ corresponds to a change in color or gray-scale between two consecutive frames F_n and F_{n-1}:

$$sum\_unmov_{n,n-1} = \sum_{(x,y)} \frac{|F_n(x,y) - F_{n-1}(x,y)|}{256} \tag{14}$$

III. Experimental results

In order to verify the effectiveness of the proposed algorithms, we analyze the experimental results obtained from the following two phases, foreground detection, and background maintenance.

Quantitative evaluation was based on the similarity measure derived in the literature [17][18]. Let A be a detected region and B be the corresponding ground truth. The similarity between A and B can be defined as:

$$S(A,B) = \frac{A \cap B}{A \cup B} \tag{15}$$

This quantity approaches 1 when A and B are similar, and 0 otherwise. In the experiment, we calculate the average value of 100 test frames.

In the foreground detection phase, we test our proposed method by public datasets [19][20] and real surveillance sequences. The challenging factors include shadow, rotation, scale changes and illumination changes (see Fig.1). In the experiment, we set $TH = 0.80$. Otherwise, in order to highlight the difference of tracking results with different detection algorithms, we implement some state-of-the-art methods including the conventional background subtraction and GMM method (see Fig. 2). The quantitative evaluation of different detection methods is listed in TABLE I.

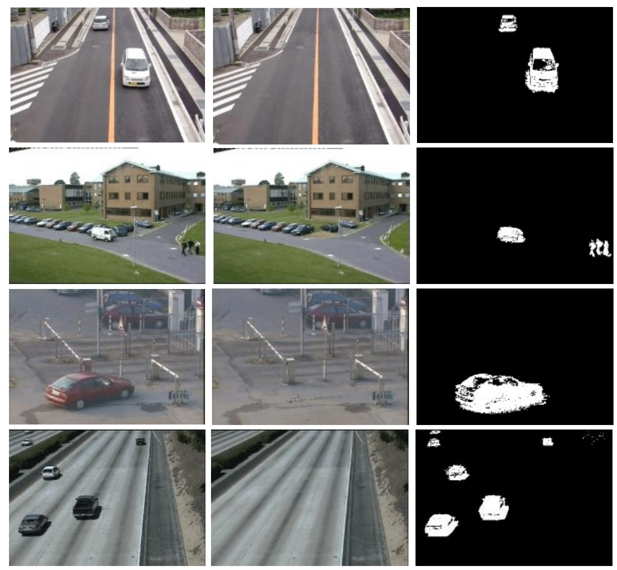


Fig. 1. Results of foreground detection (left: current frames; middle: extracted background frame; right: extracted foreground frame).

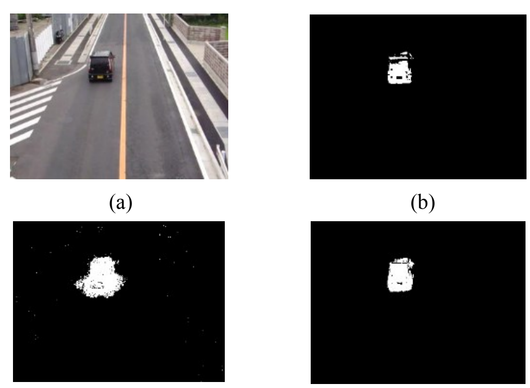


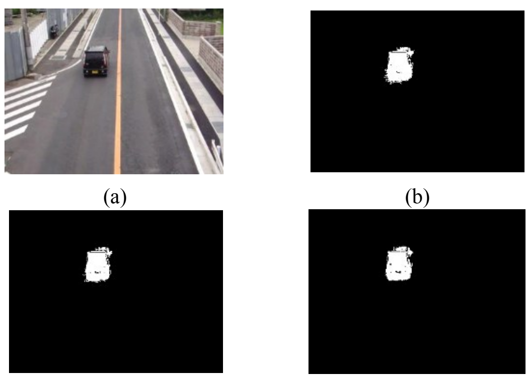
Fig.2. Results of comparing experiment, showing (a) current Frame, (b) conventional background subtraction,(c) GMM,(d) proposed method.

In the background maintenance phase, because different background maintenance methods will produce different background model frames which are segmented by the same segmentation methods, the moving object is not the same.

Therefore, we take the average ratio of the moving object image and the ground truth, and aim to reflect the effect of different background maintenance methods. Experimental results for different background maintenance methods are shown in Fig. 3. The quantitative evaluation is shown in TABLE Ⅱ.

TABLE I
QUANTITATIVE EVALUATION OF OBJECT DETECTION

Method	*S(A,B)* %
Classical background subtraction	66.00
GMM	50.72
Proposed method	81.20



(a) (b) (c) (d)

Fig. 3. Comparison of maintenance methods, showing (a) Current frame, (b) blind maintenance (c) selective maintenance, (d) proposed method.

TABLE Ⅱ
EVALUATION OF BACKGROUND MAINTENANCE METHODS

Maintenance method	*S(A,B)* %
Blind Maintenance	76.65
Selective Maintenance	77.74
Adaptive Maintenance	81.20

The experimental results demonstrate the effectiveness of our proposed method for moving vehicle detection

IV. CONCLUSION

In this paper, we propose a novel fuzzy background subtraction for moving vehicle detection. The proposed method adopts the Choquet integral for fusing color features and texture feature. The YCbCr color space is adopted to model the background frame, and to segment the foreground, which reduces the influence of the complex conditions and achieves the high detection rates. ULBP feature allows more effective distinction of textures and adapts to illumination changes. For background maintenance, we propose an adaptive selective background maintenance method to address the problem of background pollution. Comparing the other methods, the experimental results show that the proposed method was more robust and efficient.

REFERENCES

[1] B. Lee, and M. Hedley, “Background Estimation for Video Surveillance,” *Image and Vision Computing New Zealand, IVCNZ 2002*, pp.315-320 , 2002.

[2] N. McFarlane, and C. Schofield, “Segmentation and Tracking of Piglets in Images,” *British Machine Vision and Applications*, pp.187-193, 1995.

[3] C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland, “Pfinder: Real-Time Tracking of the Human Body,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780-785, 1997.

[4] M. Zhao, N. Li, and C. Chen, “Robust Automatic Video Object Segmentation Technique,” *IEEE Conf. on Computer Vision and Pattern Recognition,* pp.246-252, 1999.

[5] C. Stauffer, and W. E. L. Grimson, “Adaptive Background Mixture Models for Real-time Tracking,” in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition, CVPR 1999*, pp.246-252, 1999.

[6] J. Zhang, and C. Chen, “Moving Objects Detection and Segmentation in Dynamic Video Backgrounds,” in *IEEE Conf. on Technologies for Homeland Security*, pp. 64-69, 2007.

[7] R. Tan, H. Huo, J. Qian, and T. Fang, “Traffic Video Segmentation using Adaptive-K Gaussian Mixture Model,” *The Int. Workshop on Intelligent Computing, IWICPAS 2006*, pp.125-134, 2006.

[8] W. E. L. Grimson, C. Stauffer, R. Romano, and L. Lee, “Using adaptive tracking to classify and monitor activities in a site,” in *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp.22-29, 1998.

[9] F. Elbaf, T. Bouwmans, and B. Vachon, “Fuzzy Statistical Modeling of Dynamic Backgrounds for Moving Object Detection in Infrared Videos,” in *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshop*, pp. 60-65, 2009.

[10] A. Elgammal, R. Duraiswami, D. Harwood, and L. S. Davis, “Background and foreground modeling using nonparametric kernel density estimation for visual surveillance,” in *Proc. of the IEEE*, vol. 90, pp. 1151-1163, 2002.

[11] W. Zhang, X.Z. Fang, and X.K. Yang, “Moving vehicles segmentation based on Bayesian framework for Gaussian motion model,” Pattern Recognition Letters, vol. 27, no. 9, pp. 956-967, 2006.

[12] H. Zhang, and D. Xu, “Fusing Color and Texture Features for Background Model,” in *3rd Int. Conf. on Fuzzy Systems and Knowledge Discovery*, vol. 4223, no. 7, pp. 887－893, 2006.

[13] F. E. Baf, T. Bouwmans, and B. Vachon, “A fuzzy approach for background subtraction,” in *15th IEEE Int. Conf. on Image Processing*, pp.2648-2651, 2008.

[14] N.A. Mandellos, I. Keramitsoglou, and C.T. Kiranoudis, “A background subtraction algorithm for detecting and tracking vehicles,” *Expert Systems with Applications*, vol. 38, pp.1619-1631, 2011.

[15] Y. Ding, W. Li, J. Fan, and H. Yang, “A fuzzy background model for moving object detection,” in *11th IEEE Int. Conf. on Computer-Aided Design and Computer Graphics*, pp. 610-613, 2009.

[16] T. Ojala, M. Pietikäinen, and T. Mäenpää, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002.

[17] R. E. Haskell and C. T. Case, “Transient signal propagation in lossless isotropic plasmas (Report style),” *USAF Cambridge Res. Lab., Cambridge, MA Rep. ARCRL-66-234 (II)*, vol. 2. 1994,

[18] L. Li, W. Huang, I. Y. Gu and Q. Tian, “Statistical modeling of complex backgrounds for foreground object detection,” *IEEE Transactions on Image Processing*, vol.13,no.11, pp.1459-1472, 2004.

[19] PETS2001 Dataset: ftp://ftp.cs.rdg.ac.uk/pub/PETS2001/DATASET1/.

[20] ATON Project Tested Dataset: http://cvrr.ucsd.edu/aton/shadow/.

[21] S. Gupte, O. Masoud, R. F. K. Martin, and N. P. Papanikolopoulos, “Detection and Classification of Vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, pp.37-47, 2002.