Auto-adaptive threshold and shadow detection approaches for pedestrians detection

Alessandro Moro
University of Trieste(IT)
alessandro.moro@units.it

Kenji Terabayashi
Chuo University(JP)
terabayashi@mech.chuo-u.ac.jp

Kazunori Umeda
Chuo University(JP)
umeda@mech.chuo-u.ac.jp

Enzo Mumolo
University of Trieste(IT)
mumolo@units.it

ABSTRACT

Video surveillance has taken a huge importance in the everyday life, in order to detect, to recognize, to prevent dangerous situations. This work aims to propose two methods able to improve the robustness and speed of the detection of moveable objects or pedestrian using a fixed stereo camera. The first method is an improvement of a real-time adaptive non-parametric threshold with a salt-and-pepper noise removal used to detect changes into images. The second method instead proposes to improve the detection of the shadow points in order to affine the extraction of the pedestrians or objects from the scene. Comparison results about accuracy and performance are shown at the end of the paper.

Author Keywords

Real-time threshold, Shadow detection, Video Surveillance

ACM Classification Keywords

Algorithms, Experimentation

1. INTRODUCTION

The detection of moving objects from a video streams is a basic and fundamental problem of a large amount of vision system including objects/human detection, tracking, traffic control, and semantic annotation. Common approaches to the detection task are faced using fixed cameras based algorithms. In this contest, fixed camera based algorithms generally use a reference image obtaining relatively good results and, often, capable to work in real-time. Normally the largest techniques used to detect the changes are threshold-based, because suited for real-time application. However, many of them are susceptible by light condition. In a complex situation where the light is not equally distributed on the scene, important part cannot be detected making a general system blind. In other situation where the light condition is particularly hard because near to saturation or too dark scene, a general system can detect too many changes introducing a too high level of noise. Moreover, the detection of cast shadow as foreground object is a common problem that could lead to undesirable consequences. For example, shadows could connect different objects or people walking in a group, generating a single object (blob) or create an instance of not existing objects ("ghost" effect).Generally the methods proposed in literature work well in limited contests or require a phase of training which introduce the limitation due to the training model or consider the environment elements (like source lights) like equally distributed. In this paper we propose an extension and improvement of two techniques suited for the following considerations: proposed algorithm must be real-time, easily adaptable to the different light and environment condition, not depend on *a priori* knowledge. The purpose of this paper is to describe a method flexible and robust, able to work well in different conditions of illumination, and able to extract automatically the relevant areas of an image. The remaining parts of this paper are organized as follows. Section 2 contains the description of the proposed methods. Section 3 shows the results and a rapid comparison with other works. Section 4 contains conclusions and future work.

1.1 Related works

In the change detection and shadow detection task, several algorithms have been proposed where change detection is particularly used in video sequence [1, 7, 11]. In recent works some authors treat threshold analysis as a component of the shadow detection problem [3, 5, 8, 10]. Image subtraction based on the use of a threshold is a popular method and a lot of methods based on thresholds have been proposed in literatures. Su *et al.* [8] focused on two types of thresholdings categories (estimation of the scatter of regions of change of a difference image, spatial properties based methods), and proposes a non-parametric algorithm to calculate the global threshold. This method is slower than traditional approaches (Poisson, Euler) but improve the detection. The negative aspect is that a single threshold is calculated for the entire image. In [6] a thresholds value is obtained considering Gaussians factor. It takes the advantages of the Gaussian models but as the previous method only one threshold is considered.

Shadow regions generally are detected based on information of luminance, chrominance and gradient density. As considered in [10], a large number of false alarms or miss detections can be reduced by the assumption of known geometry information. Some authors try to overcome linear hypothesis of the shadow model and create a new representation of the illumination model. Martel-Brisson *et al.* [4] describes a property-based method which describes the cast shadow direction and unsupervised kernel-based approach to estimate it. It can detect the shadows on objects of similar chromacity of the background but the negative aspect is that model detects also the shadow of the foreground elements removing parts of detected objects. A grayscale based algorithm proposed for car detection tasks is described in [9]. Spatial and temporal constraints during the detection process are unified in a conditional random field (CRF) based probabilistic model that permits neighborhood interactions in both labels and observations. This method is particularly suited to detect the shadow of vehicle because based on a model but not flexible for a general environment. A color and gradient-based method to remove shadow is described in [3]. It shows good results in different conditions but it`s cumbersome and use several filters (morphological, connect) which may create artifact or remove necessaries elements.

2. PROPOSED METHOD

Detection of moving objects is an important topic for analyzing the contents of a sequence of images. The first step of a monitoring system is the generation of a correct model of background in order to extract moving objects from the video stream. In this approach, two methods are proposed to improve algorithms described in [8, 10] which face the problems of determines the threshold values and shadow detection. The first method increases the threshold performance using tracking data and combining neighbor windows. The second method improves the shadow detection introducing light condition, considering shadow direction and using a different color model.

2.1 Auto thresholding

In order to extend increase the detection rate, the algorithm described in [8] has been extended. Two basic components have been included. Use tracking information to refine the thresholds values. Extend the use of the algorithm in hard light conditions by composition of neighbor windows. Given an image frame and a rectangular research window *w*, thresholds values into the entire possible window *W* are calculated as defined in [8]. Moving the window over the image of the halved of its size in vertical and horizontal direction, it will consider each pixel such that $\forall_{x,y} \exists A \in W, \forall_{i>0, i \in N} \exists w^i \in A$ such that the threshold will be defined as

$$T_{(t+1,x,y)} = \frac{f(x,y,w,|A|) + p_{(t,x,y)} \cdot T^{\vartheta}_{(t,x,y)}}{1 + p_{(t,x,y)}} \quad (1)$$

T^{ϑ} is the threshold calculated into a rectangular area which size and position depends on the size of a detected object ϑ on the scene in instant *t* and *p* is a function which assume value 1 if an object is detected, 0 otherwise. *f* a recursive function so defined

$$f(x,y,w,i) = \begin{cases} \frac{T^{w^i}_{(t,x,y)} + f(x,y,w,i-1)}{2}, \text{if } \frac{T^{w^i}_{(t,x,y)} + f(x,y,w,i-1)}{2} > \varepsilon \\ \varepsilon \text{ , otherwise.} \end{cases} \quad (2)$$

In the situation of empty areas the small variations of the foreground images respecting the background generate a salt and pepper effect which is cause of detection fails and reduction of time performances. To overcome this problem, based on the idea in [2], a simple non-parametric function is used. The new threshold value $T'_{(t,x,y)}$ will be

$$T'_{(t,x,y)} = \begin{cases} \infty, \text{if } n_{(x,y)} > \frac{H \cdot W}{2} \\ T_{(t,x,y)}, \text{otherwise.} \end{cases} \quad (3)$$

Where H and W are height and width of half of research window, I the intensity of the image, and $n_{(x,y)}$ is a noise function defined as

$$\sum_{y=1}^{H-1} \sum_{x=1}^{W-1} g\left(I_{(x,y)}, T_{(x,y)}\right) \sum_{y'=y-1}^{y+1} \sum_{x'=x-1}^{x+1} h(I_{(x',y')}, T_{(x',y')}) \quad (4)$$

Where *g(a,b)* is a function that return 1 if a > b and 0 otherwise, and *h(a,b)* is a function that return 1 if a ≤ b and 0 otherwise. The coordinate are relative of the window frame analyzed.

Fig. 1 Example of noise remove. On the left changes detected using the dinamic threshold. On the right remove of the noise

2.2 Shadow detection

As described in the previous paragraphs, shadow detection has an important role in the detection. The shadow detection described in [10] has been improved applying three main changes that will be focused in the next equations. It has been introduced a light condition (5) which consider the shadow as a reduction of light.

$$\theta_{(t+1,x,y)} = \begin{cases} \alpha\Psi_{(x,y)} + \beta\Lambda_{(x,y)} + (1 + \alpha + \beta)\theta_{(t,x,y)}, \text{if } \frac{I_{(x,y)}}{\eta} < I'_{(x,y)} \\ \infty, \text{otherwise} \end{cases} \quad (5)$$

Where θ corresponds to a shadow value. At that value will be applied a threshold to determine if a pixel is a shadow. To a small value corresponds a shadow point. α, β, and η are constant values which defines the weight of textures, colors and intensity. It combines the relationship between pixel Ψ (8) and within pixel Λ (10) using temporal information.

Shadow generally has a direction. As defined in [10], suppose to be I(x, y) the intensity of the pixel located in (x, y) and E(x, y) the direct light and ρ(x, y) the reflected light from a 3D point and projected into the point (x, y). Simplifying the models it is possible to consider that an occlusion as shadow casting reduces the irradiance but not the diffuse reflectance.

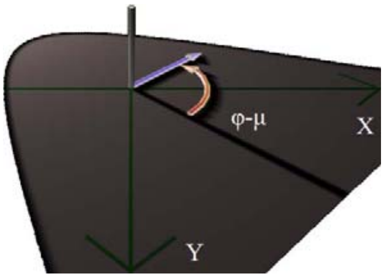

Fig. 2 Representation of the shadow model. The source light create a shadow cone but the presumed shadow direction is miscalculated (blue arrow).

Considering that the irradiance influence the cast shadow not linearly than the intensity ratios between neighboring shadow pixels depends on the source direction. So if has considered all the pixel of the image, $\forall_{x,y} \exists w, \forall_{i,j} | i \in [x-1, x+1], j \in [y-1, y+1], i \neq x \wedge j \neq x$ the intensity ratio will be

$$\frac{I_{(x,y)}}{I_{(i,j)}} = \frac{E_{(x,y)}}{E_{(i,j)}} \cdot \sin\left(\frac{(\varphi-\mu)}{2}\right) + \frac{\rho_{(x,y)}}{\rho_{(i,j)}} \quad (6)$$

Where the angle of the shadow is φ respect the Cartesian axes, and μ is the angle of research of the shadow (Fig.2).

Unknowing the source light *a priori*, to calculate the shadow effect to the intensity ratio it has been minimized the logarithmic variation of intensities instead to force a direction. As before, if consider the pixel of coordinate *x, y* $\forall_{x,y} \exists w, \forall_{i,j} | i \in [x-1, x+1], j \in [y-1, y+1], i \neq x \wedge j \neq x$ than the intensity ratio *d* will be

$$d_{(x,y)} = \min\left(\left|\ln I_{(x,y)} - \ln I_{(i,j)}\right|\right) \quad (7)$$

The error score for discriminating the pixel (x, y) as shadow can be calculated as difference of foreground $d(i,j)$ and background intensity ratio $d'(i,j)$ in a small window ω centered.

$$\Psi(x,y) = \sum_{c \in R,G,B} \sum_{(i,j) \in \omega(x,y)} |d(i,j) - d'(i,j)| \quad (8)$$

A different color model is used to the color constancy within pixel in order to speed up the elaboration time.

As the shadow change the brightness of the background, the colors tends remains the same. Comparing the color information between background and foreground can help in shadow detection. Several color models with different properties has been developed but in order to have a faster and at the same time accurate model, it is used a 2D color space from the RGB and a normalized color space.

$$\begin{cases} C_1(x,y) = \tan^{-1}\left(\frac{I_R(x,y)}{I_B(x,y)}\right) \\ C_2(x,y) = \tan^{-1}\left(\frac{I_G(x,y)}{I_B(x,y)}\right) \end{cases} \quad (9)$$

And then, the color score error can be defined like a variation between background and foreground.

$$\Lambda(x,y) = |C_1(x,y) - C_1'(x,y) + C_2(x,y) - C_2'(x,y)| \quad (10)$$

3. RESULTS

In order to evaluate the performance of the proposed algorithm four videos has been taken from indoor and outdoor environments in different moment of the day using Point Gray Research Bumblebee camera with 640 x 480 pixel resolution. The results have been obtained using an Intel Core2 Quad CPU, 2.83 GHz with 4 GB ram. The videos are obtained with frame rate at 30 frames/sec. For indoor environment movies are taken in static background where the light source is controlled, instead the outdoor environment suffers of light change (Fig.3). The movies have been taken in a sunny day in three different time of the day, from the 11am to 5pm.

Fig. 3 Example of environment studied

In order to evaluate the real-time thresholding, and the detection rate and false alarm rate of shadows, it has been create a collection of ground truth labeled manually first (Fig.4). Frames in the videos are selected in approximately equal time interval, discarding the situation of no object showing up or just partial appeared.

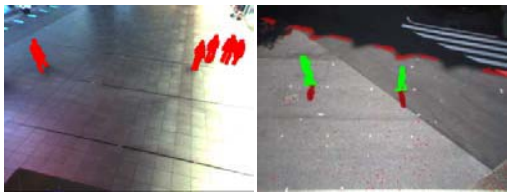

Fig. 4 The labeled ground truth in the testing videos. On the left in red the pedestrian areas. On the right in green the pixel corresponding to pedestrian areas, in red the pixel corresponding to shadow

Let "CD" be change detected, and "CN" be change not detected. Let "FN" be false negative the case that a change is detected where there are no changes. Then the accuracy of the real-time thresholding is evaluate calculating the detection rate (DR = CD / (CD+CN)) and the accuracy rate (AR = CD / (CD+FN)). For the evaluation purpose and due to the difficulty to define the shadow areas, only the real pedestrian or objects have been considered. As a result, it has been compared the performances of the proposed algorithm to the base real-time adaptive threshold [5] and two fixed thresholds determined by visual inspection. The experimental results are shown in Table 1. For the shadow detection, let "TP" be true positive, "FP" be false positive, and "FN" be false negative. The accuracy of the shadow detection is then evaluated by calculating detection rate (DR = TP/(TP+FN)) and false alarm rate (FAR = FP/(TP+FP)). The parameters have been detected probing all the possible values in a reasonable range and selecting the configuration gave the best result.

The proposed algorithm improves the robustness respect the original work. In particular the shadow detection improves the detection rate reducing the case of false alarm. Moreover, as shown in table 1 and table 2 on the bottom, the elaboration time allow an important aspect that the algorithms are feasible for the real-time application.

	Proposed Meth.	Base Meth.[8]	Threshold 1	Threshold 2
DR	0.918	0.927	0.874	0.815
AR	0.303	0.103	0.267	0.587
Time (ms)	18	6	3	3

Table 1 Detection rate. This table compare the detection rate and accuracy using the proposed real-time method, the original method and two fixed threashold values

	Proposed Method	Multiple Cues
DR	0.958	0.919
FAR	0.045	0.057
Time Par Pixel (s)	$3.6 \cdot 10^{-7}$	$3.9 \cdot 10^{-7}$

Table 2 Comparison of DR and FAR with using the proposed method and the Multiple Cues method

Fig. 5 Two examples of elaboration. From left to right. Original image, detection of changes (red), shadow detected (green) and pedestrian detected (blue)

4. CONCLUSION

This study presents a real-time adaptive threshold algorithm with shadow detection for the applications of video surveillance. The proposed method can automatically detect the changes on the images without manually setting any threshold values and is adaptable for indoor and outdoor applications. The contribution of this study is to propose a combination of methods to detect moveable object or pedestrian in the scene which not need any training and is easily adaptable of different environment.

So far, the proposed algorithm works well in several environments under light and dark condition. When the objects are too similar in intensity to the original scene but different in color are difficult to be captured and if the color of a pixel belong to a possible shadow range it will be consider as shadow. In the next will be focused the possibility to calculate the shadow parameters at real-time and dynamically change the regions size and form to evaluate the threshold values. An improvement of the real-time adaptive threshold will be given considering the color information and detection context to improve the shadow labeling. More environments will be analyzed and the method will be compared with more methods. Test will be extended also in time to evaluate the night time performances.

REFERENCES

[1] A. Amer, E. Dubois, "Memory-based spatiotemporal real-time object segmentation," *Proc. SPIE Int. Symposium on Electronic Imaging, Conf. on Real- Time Imaging*, vol. 5012, pp. 10–21, 2003.

[2] R.H. Chan, Ho Chung-Wa, M. Nikolova, "Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization", IEEE Transactions on Image Processing, vol. 14, pp. 1479-1485, 2005

[3] M. Izadi, P. Saeedi , "Robust region-based background subtraction and shadow removing using color and gradient information", ICPR 2008, pp. 1-5, 2008

[4] N. Martel-Brisson and A. Zaccarin, "Kernel-based learning of cast shadows from a physical model of light sources and surfaces for low-level segmentation", 2008 IEEE Conference on Computer Vision and Pattern Recognition , pp. 1-8, 2008

[5] N. Martel-Brisson and A. Zaccarin, "Learning and Removing Cast Shadows through a Multidistribution Approach", *Pattern Analysis and Machine Intelligence*, vol 29, pp. 1133-1146, 2007

[6] Bo Quin, Chuangde Zhang, Zhengua Fang, Wei Li, "A Quick Self-Adaptive Background Updating Algorithm Based On Moving Region", Proceedings of the 9th Jointh Conference on Information System, 2006

[7] R. J. Radke, S. Andra, O. A-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *IEEE Transactions On Image Processing*, vol. 14, no. 3, pp. 294–307, 2005.

[8] C. Su, and A. Amer, "A Real Time Adaptive Thresholding for Video Change Detection, "*Proc. IEEE Int. Conf. in Image Processing 2006* ,pp. 157-160, 2006

[9] Yang Wang , "Real-Time Moving Vehicle Detection With Cast Shadow Removal in Video Based on Conditional Random Field", Circuits and Systems for Video Technology, vol. 19, pp. 437-441, 2009

[10] M.-T. Yang, K.-H. Lo, C.-C. Chiang, W.-K. Tai, "Moving cast shadow detection by exploiting multiple cues", Image Processing, IET Vol 2, pp. 95-104, 2008

[11] D. Zhang, G. Lu, "Segmentation of moving objects in image sequence: a review," *Circuits Systems Signal Processing*, vol. 20, no. 2, pp. 143–183, 2001.