14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

ADDING GEOMETRICAL TERMS TO SHADOW DETECTION PROCESS

László Havasi, Tamás Szirányi and Michael Rudzsky

Hungarian Academy of Sciences
H-1111 Budapest, Kende u. 13-17, Hungary
Péter Pázmány Catholic University
H-1083 Budapest Práter u. 50/a., Hungary
email: {havasi, sziranyi}@sztaki.hu
Faculty of Computer Science, Technion-Israel Institute of Technology
32000, Haifa, Israel
email: rudzsky@cs.technion.ac.il, web: www.cs.technion.ac.il/~rudzsky

ABSTRACT

The elimination of strong shadow in outdoor scenes containing human activity is addressed in the paper. The main contribution of the introduced method is the integration of geometrical information into the shadow detection process. This novel approach takes into account the collinearity of shadow and light direction and completed with a simple colour based pre-filtering. The final classification step is carried out via a Bayesian iteration scheme which is general enough to handle further characteristics of the problem: weak shadow and reflection.

1. INTRODUCTION

Moving object detection is a key issue in most computer vision applications especially for surveillance purposes. Depending on the scene settings the cast shadow usually generates problems while extracting moving objects (e.g. silhouettes). The problem occurs often in outdoor scenes and indoor configurations when the floor is a reflective surface. In most cases shadow can cause merging of objects, shape distortion and object losses. Thus, shadow detection is critical for accurate object detection, which is a relevant step of information extraction for further processing: tracking, event detection [1] or traffic monitoring [2]. Many approaches have been proposed in the literature that deal with shadow. A good survey can be found in [3][4]. Most of the publications are focused on the colour based shadow detection [2][3][5][8]. In order to remove shadow points, these methods have defined conditions in some proper colour space.

Method, called SAKBOT, was introduced in [2]. It was developed for moving object detection and tracking. This complex algorithm contains both colour and motion information. Additionally, the final foreground mask is improved with knowledge-based feedbacks. The method introduced in [6] utilizes some predefined object model present estimation about shadow pixels near to the detected objects. Summarising, the basic features that can be used to distinguish between shadow and object points are: colour, texture, motion. In spite of the notable amount of publications there is no approach which utilizes the general geometrical model of cast shadow and includes the geometrical characteristics into the classification process. Geometrical information is important in describing the creation of shadow. Like other characteristics, the geometrical description is not a unique feature, so without other features it is not sufficient for accomplishing of classification purposes in all cases.

This paper presents a new approach for motion mask classification allowing for the general geometrical model of a point like light source. Our goal is not to present an all-in-one algorithm for shadow detection, but the main idea is to clamp features into a probabilistic framework. During evaluation outdoor sequences will be used, which is impaired by strong shadow and showing pedestrians.

2. PROPOSED METHOD

The goal of shadow detection is to eliminate the shadow points from the extracted foreground mask. The foreground mask can be determined using several approaches. Our implementation based on [7] which is a popular background modelling method for the extraction of foreground mask. The output mask is defined by:

$$m_i = \begin{cases} 1, & \text{where change is detected} \\ 0, & \text{otherwise} \end{cases} \qquad (1)$$
$$M = \{m_i, i \in S\}$$

where S denotes all pixels in the image (index i corresponds to one pixel). In the followings I determines the input image and B denotes the computed background image, respectively. Intensity changes cause changes in object pixels as well as in cast shadow pixels.

Thus, the initial foreground-background mask M contains both object and shadow pixels as foreground.

2.1 Basic Colour Filtering

The usual method to distinguish between moving cast shadow and object points is the investigation of pixels in Hue-Saturation-Value (HSV) colour space [2][5]. This preprocessing step is a simple filtering before higher level processing. In the paper we focus on strong shadow in outdoor environment, thus we have implemented only the condition related to Value (V):

$$c_i = \begin{cases} 1, \text{ if } \alpha \le \dfrac{V(I_i)}{V(B_i)} \le \beta \wedge V(I_i) < 0.6 \\ 0, \text{ otherwise} \end{cases} \quad (2)$$

$$C = \{c_i, i \in S\}$$

Here C denotes a shadow mask with elements c_i equal “1” for shadow pixels and “0” otherwise, according to colour based conditions: α and β are bounds obtained from experiments.

In our experiments this colour based method works reliable only in case of weak shadow. If the shadow is strong, the ratio is around 0.4. Unfortunately, however the ratio changes to 0.9 near to the boundary of shadow. That is why shadow elimination is not possible by using only the colour information. Figure 1 displays the histogram of the ratio inside the manually selected shadow mask.

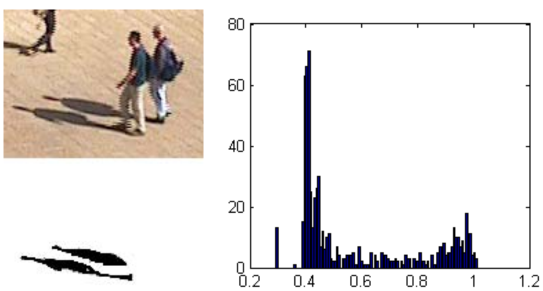

Figure 1 – Difficulty of colour based shadow detection in case of strong shadow: the shadow region has not a histogram with only one peak, and thresholds α and β are not the same for the whole image.

Another problem is that the adaptive determination of the bounds α and β is still a challenging task. Thus, in our implementation these values were adjusted to cover a relatively large region of the full range (α=0.4, β=0.8). Detection results are summarized in the following figure.

Figure 2 – Results of colour based shadow detection: upper left-input image, upper right-motion-detection mask, lower left-foreground mask determined by using colour features [5] and lower right-“worse-case” shadow mask (α=0.4, β=0.8) used for input to classification method. (The binary masks are without morphological post-processing.)

2.2 Geometrical Shadow Modelling

For most outdoor situations, the direction of daylight shadows is controlled by the position of sun. Because the rays of sunlight are essentially parallel, they converge in an infinite vanishing point.

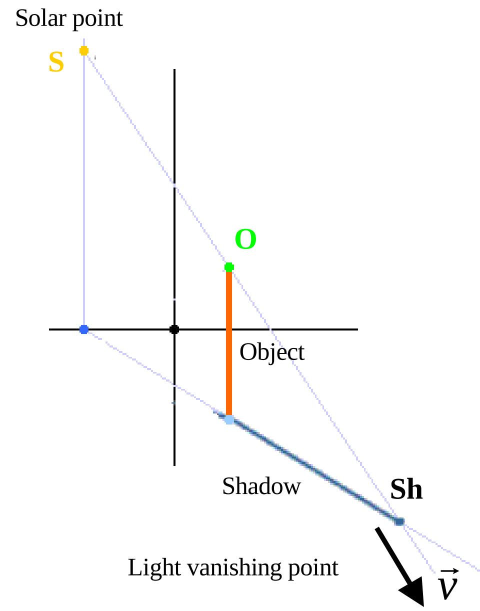

Figure 3 – Geometry of shadow: the three points; object (original), shadow and vanishing point are collinear. Because in outdoor cases the distance of the light source from the object casting the shadow is near infinity, the knowledge of a common direction ($\vec{v}$) is sufficient instead of position of vanishing point.

Because of the far vanishing point the geometrical model may be simplified; the knowledge of a direction (2D vector) is enough for the estimation of shadow region. Details about the geometry can be found in [9][10]. In our previous work we have described a method to compute the vanishing point in case of camera-mirror setting using motion statistics [10]. Shafer in [9] points out that an object and its cast shadow share a similar geometrical relationship to that found in the camera-mirror case. Thus, the determination of this direction is possible both automatically and manually.

The importance of knowledge of the light vanishing point lies in the fact that it simply enables the integration of geometrical constraint into the shadow detection process. Obviously, the shadow point must lie on the line going through the original point with direction $\vec{v}$.

Note that, $\vec{v}$ has unit length: $\|\vec{v}\| = 1$. Unfortunately, this geometrical model is not enough for the exact determination of the corresponding shadow point, because it is a point-to-line transformation instead of a point-to-point analogy. Accordingly, we will use this extra knowledge together with other features (colour, motion etc.) to achieve a better classification results. In the following section an iterative method based on Bayesian formula has been introduced to find the most probable shadow points and foreground mask.

2.3 Classification Process

The aim of classification is to decide about every foreground pixel in the initial foreground-background mask (*M*) whether it is a foreground pixel or a shadow pixel. This two-classes problem is equivalent to find the probable class for an arbitrary pixel in the given scene setting (*M* and *C*). In the literature there are several different approaches to accomplish such classification tasks [11]. Nevertheless, in this section a simple Bayesian iteration will be introduced. We selected this probabilistic framework, because it is rather general and is suitable for further improvements. This method was used successfully for blind deconvolution in [13] and [12].

We define the unknown shadow mask as

$$H = \{h_i, i \in S\} \tag{3}$$

and foreground mask as

$$F = \{f_i, i \in S\} \tag{4}$$

Together with definition of the detected initial foreground-background mask *M* (1), using Bayes conditional probability formula we can get the probability of observing *H* and *F* with given *M* in the following form (the formulas for *F* are similar):

$$P\left(h_i \middle| m_j\right) = \frac{P\left(m_j \middle| h_i\right) P\left(h_i\right)}{\sum_k P\left(m_j \middle| h_k\right) P\left(h_k\right)}, \quad i, j \in S \tag{5}$$

Substituting this equation into the conditional probability formula, we get:

$$P\left(h_i\right) = \sum_j P\left(h_i m_j\right) = \sum_j P\left(h_i \middle| m_j\right) P\left(m_j\right) = \sum_j \frac{P\left(m_j \middle| h_i\right) P\left(h_i\right) P\left(m_j\right)}{\sum_k P\left(m_j \middle| h_k\right) P\left(h_k\right)}, \quad i, j \in S \tag{6}$$

Based on this formula the following iteration scheme can be written [12]:

$$P_{k+1}\left(h_i\right) = P_k\left(h_i\right) \sum_j \frac{P\left(m_j \middle| h_i\right) P\left(m_j\right)}{\sum_k P\left(m_j \middle| h_k\right) P\left(h_k\right)}, \quad i, j, k \in S \tag{7}$$

Where k is the iteration counter. We define the initial probabilities as follows,

$$P\left(m_i\right) = \begin{cases} 1, & \text{where } m_i = 1 \\ 0, & \text{otherwise} \end{cases}, \; i \in S \tag{8}$$

and

$$P\left(h_i\right) = \begin{cases} 1/2, & \text{where } c_i = 1 \\ 0, & \text{otherwise} \end{cases}, \; i \in S \tag{9}$$

$$P\left(f_i\right) = 1 - P\left(h_i\right)$$

During the iteration steps the values of $P\left(f_i\right)$ and $P\left(h_i\right)$ will change. These probabilities converge to stable values which describe the most probable shadow/foreground configuration in a given motion mask. Hereby the classification step is a simple decision to the most probable class:

$$\hat{f}_i = \begin{cases} 1, & \text{if } P\left(h_i\right) < P\left(f_i\right) \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

because,

$$P\left(h_i\right) + P\left(f_i\right) = 1 \quad \forall i \in S \text{ where } m_i = 1 \tag{11}$$

The key issue in the above introduced formula (7) is the determination of the conditional probability term ($P\left(m_j \middle| h_i\right)$). This term enables the completion of probability model with additional knowledge about the problem. First, in case of shadow the indices in the summarizations can be reduced. In the above-defined form the summations are performed over the whole image. According to the geometrical model these 2D summations may be replaced with summations along a straight line (parameterised in 1D), direction of which is equal to $\vec{v}$. This procedure is demonstrated in Figure 4.

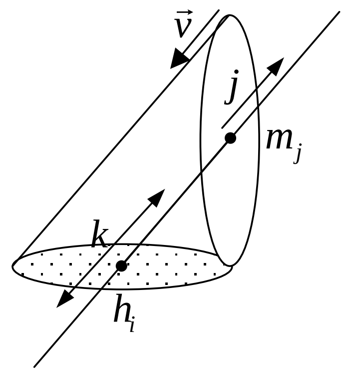

Figure 4 – Simple geometrical constraint: including the collinearity to the conditional probabilities. The notations are introduced in the text. The indices *j* and *k* are related to the cyclical summarizations.

To allow this feature, a modified formula of (7) can be written:

$$P_{k+1}\left(h_i\right) = P_k\left(h_i\right) \sum_j \frac{P\left(m_{r(j,i)} \middle| h_i\right) P\left(m_{r(j,i)}\right)}{\sum_k P\left(m_{r(j,i)} \middle| h_{r(k,i)}\right) P\left(h_{r(k,i)}\right)}, \quad i \in S \text{ and } j, k \in N \tag{12}$$

where the function $r(.)$ returns an image point, which is computed from an initial position (*i*) and a step counter (*j*) along the line:

$$r(j,i) = i + j\vec{v} \tag{13}$$

Based on this notation the expression of conditional probability may be rewritten using two step-counters along the line:

$$d_i(p,l) = P\left(m_{r(p,i)} \middle| h_{r(l,i)}\right), \quad p, l \in N \tag{14}$$

For determination of this value, a simple formula is given:

$$d_i(p,l) = c_{r(p,i)} m_{r(p,i)} m_{r(l,i)} \left(1 - P\left(f_{r(l,i)}\right)\right) \tag{15}$$

This expression validates only the minimal conditions; motion must be present in both points and the shadow point must be in the colour mask (value of $c_{r(p,i)}$ is defined by (2)). The last component relates to the foreground mask, so, the probability that a given point belongs to shadow is equivalent to the probability that the point is not a foreground point. This extension makes connection between *F* and *H* during the iterations.

Since, the iteration formula contains the sum of these values (and because it is a 2D probabilistic distribution function), the normalization is necessary before use:

$$\sum_{p,l} d_i(p,l) = 1 \tag{16}$$

Till now, we used indices along the line without any upper and lower bounds. There is no need to compute the sums along the full line, because we can define the probable utmost distance between the original point and its corresponding shadow point. This distance is symbolized by parameter γ. Thus the ranges of the indices are

$$\begin{aligned} p &= -\gamma \ldots \gamma \\ l &= 0 \ldots \gamma \end{aligned} \tag{17}$$

The values of $d_i(.)$ form a 2D pdf function. Its layout is visualized in matrix form in figure 5.

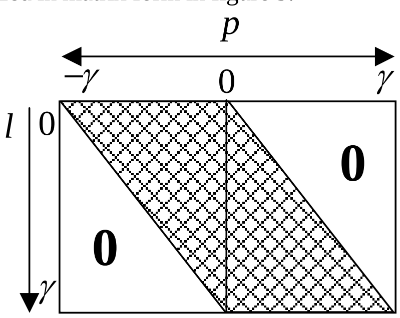

Figure 5 – Layout of matrix $d_i(.)$. The filled region indicates the probable non-zero elements. This zone-structure is because the place of shadow is always relative to the original point.

After substituting (14) into (12) we get the final formula of the iteration step:

$$P_{k+1}(h_i) = P_k(h_i) \sum_j \frac{d_i(0,j) P\left(m_{r(j,i)}\right)}{\sum_k d_i(k-j,j) P\left(h_{r(k,i)}\right)}, \tag{18}$$

$$i \in S \text{ and } j,k = 0 \ldots \gamma$$

The formulas for the foreground probabilities ($P(f_i)$) can be derived in the same way.

3. ALGORITHM EVALUATIONS

In the following, we present some samples of outdoor sequence. The first row contains the relevant part of input images. The second row displays the detected motion mask, see (1). The further rows demonstrate the foreground and shadow probabilities during the iteration steps.

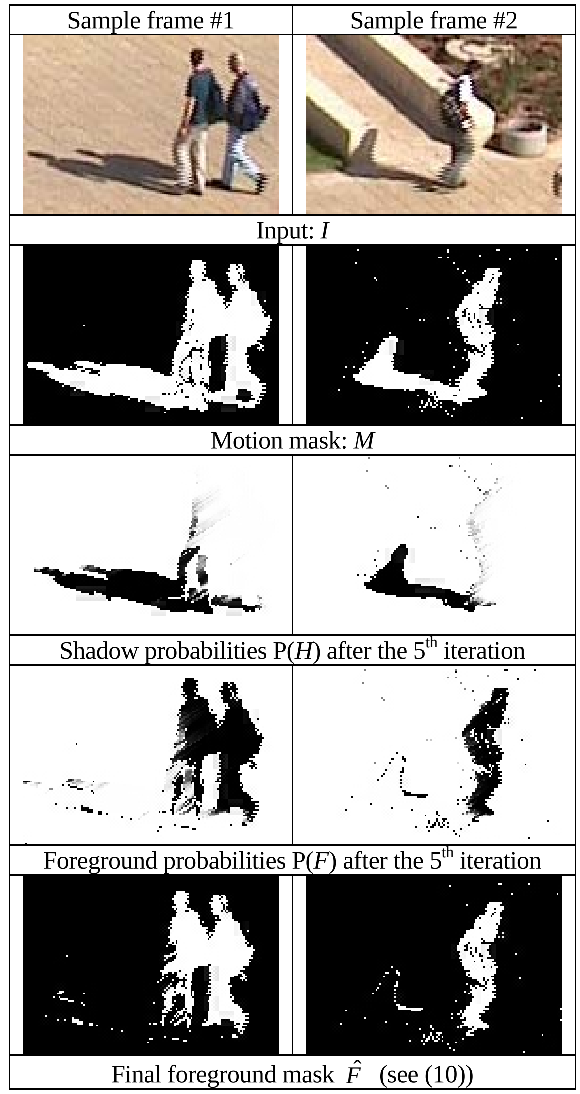

Figure 6 – Experimental results on strong shadow. The final foreground mask is the output of the classifier, for details see text.

In our experiments the method converges within 5 iterations.

4. FUTURE WORK

The described method is general but in the current state it contains only information for strong shadow removal. In summary, the following main future research directions can be formulated:

1. *Estimation of light direction:* for real-life applications it is necessary to track the changes of vanishing point during the day. The most comfortable way is to use the known relations from astronomy; the knowledge of location and current time is utilizable for the determination of polar angle. Furthermore, for the transformation of the direction onto the camera plane the camera parameters (camera matrix, P) are also required.
2. *Extended colour based pre-filtering:* the basic colour model is feasible to detect weak shadow [15] where the

fluctuation of value is small enough. Nevertheless, one of our goals is the detection of reflection in indoor videos. The applied colour based approach is not well adapted for the modelling of colour mixture in such environment. For this purpose the geometrical model is a good basis, because the reflection is generated from both the background and object textures [14]. Furthermore, the position of vanishing point in the case of reflection is nearly stable.

3. *Determination of corresponding object-shadow point pairs:* useful information for the more precise classification is the computation of a distance map which contains estimations about the distances between original (object) points and their shadow points. This computation needs the extension of the iteration formula, similar to [12]. In [12] the authors used the same formula to determine the localized point spread function (PSF). This distance map is applicable not only to expand the classification, but for 3D information extraction about the objects.

5. DISCUSSION

The paper introduced a new approach for cast shadow detection in outdoor video sequences. We have presented an iterative Bayesian framework to determine the shadow and foreground masks taking into account both colour and geometrical information. The geometrical model of cast shadow is reduced to a simple direction (toward light vanishing point), which assists to implement an efficient localized variant of the iteration scheme. We also have named the main research directions to achieve a general shadow elimination method.

6. ACKNOWLEDGEMENTS

The authors would like to thank to Ehud Rivlin for his remarks. Our work was supported by the EU in the NoE MUSCLE project.

REFERENCES

[1] Z. Zalevsky, E. Rivlin, and M. Rudzsky, "Motion characterization from co-occurrence vector descriptor," *Pattern Recognition Letters*, vol. 26, pp. 533-543, 2005

[2] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Statistical and knowledge-based moving object detection in traffic scene," in *Proc. of IEEE Int'l Conf. On Intelligent Transportation Systems*, 2000, pp. 27–32.

[3] A. Prati, I. Mikic, M. M. Trivedi, and R. Cucchiara, "Detecting Moving Shadows: Algorithms and Evaluation," *IEEE Trans. PAMI*, vol. 25, pp. 918-923, 2003.

[4] R. Cucchiara**,** C. Grana, M. Piccardi, A. Prati, "Detecting Moving Objects, Ghosts and Shadows in Video Streams" in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1337-1342, 2003

[5] Cs. Benedek, and T. Sziranyi, "Markovian Framework for Foreground-Background-Shadow Separation of Real World Video Scenes," *Lecture Notes on Computer Science*, ACCV, 2006

[6] A. Yoneyama, C. H. Yeh, and C. C. Jay Kuo, "Moving Cast Shadow Elimination for Robust Vehicle Extraction based on 2D Joint Vehicle/Shadow Models," in *Proc. of IEEE Int'l Conf. on Advanced Video and Signal Based Surveillance*, 2003.

[7] C. Stauffer, W. Eric L. Grimson, Learning patterns of activity using real-time tracking, *IEEE Trans. PAMI*, vol. 22, pp. 747-757, 2000.

[8] S. Nadimi, and B. Bhanu, "Physical Models for Moving Shadow and Object Detection in Video," *IEEE Trans. PAMI*, vol. 26, pp. 1079-1087, 2004.

[9] S. A. Shafer, "Shadows and Silhouettes in Computer Vision," *Kluwer Academic Publisher*, 1985.

[10] L. Havasi, and T. Sziranyi, "Estimation of the vanishing point in camera-mirror scenes using video," O*ptics Letters*, vol. 31, No. 10, 2006, in press.

[11] A. Webb, "Statistical Pattern Recognition," *John Wiley & Sons*, 2004.

[12] L. Kovács, and T. Szirányi, "Relative Focus Map Estimation Using Blind Deconvolution," *Optics Letters*, vol. 30, pp. 3021-3023, 2005.

[13] W. H. Richardson, "Bayesian-Based Iterative Method of Image Restoration," *JOSA* vol. 62, pp. 55-59, 1972.

[14] A. Levin, A. Zomet, and Y. Weiss, "Separating reflections from a single image using local features," in *Proc. of IEEE Int'l Conf. on Computer Vision and Pattern Recognition*, pp. 306-313, 2004.

[15] G. D. Finlayson, M. S. Drew, and C. Lu, "Intrinsic Images by Entropy Minimization," in *Proc. of European Conference on Computer Vision*, *LNCS*, pp. 582-595, 2004.