Cepstrum Based Method for Moving Shadow Detection in Video

Fuat Cogun and A. Enis Cetin *

Bilkent University, Bilkent 06800, Ankara, Turkey
fuatc@ee.bilkent.edu.tr, cetin@bilkent.edu.tr

Abstract. Moving shadows constitute problems in various applications such as image segmentation and object tracking. Main cause of these problems is the misclassification of the shadow pixels as target pixels. Therefore, the use of an accurate and reliable shadow detection method is essential to realize intelligent video processing applications. In this paper, the cepstrum based method for moving shadow detection is presented. The proposed method is tested on outdoor and indoor video sequences using well-known benchmark test sets. To show the improvements over previous approaches, quantitative metrics are introduced and comparisons based on these metrics are made.

1 Introduction

In many computer vision applications, moving shadows may lead to inaccurate moving object detection results. All moving points of both objects and shadows are detected at the same time in most common video foreground object detection methods requiring inter-frame differentiation or background subtraction. In addition, moving shadow pixels are normally adjacent to moving object pixels. Hence, moving shadow pixels and object pixels merge in a single blob causing distortions of the object shape and model. Thus, object shape is falsified and the geometrical properties of the object are adversely affected by shadows. As a result of this, some applications such as classification and assessment of moving object position (normally given by the shape centroid) give erroneous results. For example, shadow detection is utmost important in forest fire detection applications [1] because shadows are confused with smoke regions as shown in Fig. 1. Another problem arises when shadows of two or more close objects create false adjacency between different moving objects resulting in detection of a single combined moving blob. Shadow regions retain underlying texture, surface pattern, color and edges in images. In [2] it is pointed out that Hue-Saturation-Value (HSV) color space analysis as a shadow cast on a background does not change significantly its hue. There have been some further studies on HSV color space analysis for shadow detection such as [3] and [4]. In the study of Jiang and Ward [5, 6], classification is done on the basis of an approach that shadows

---

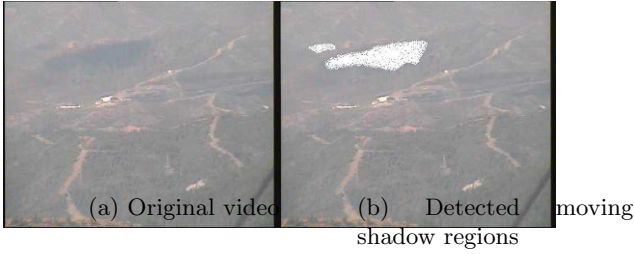(a) Original video   (b)   Detected   moving
shadow regions

Fig. 1.: The source of the shadow regions are the moving clouds

are composed of two parts: self-shadow and cast shadow. The shadow detection approaches are classified as statistical and deterministic type and comparisons of these approaches are made in [9], [10] and [11]. In this paper, a deterministic two-dimensional cepstrum analysis based shadow detection method is proposed. The method is composed of two steps. In the first step, hybrid background subtraction based moving object detection is implemented to determine the candidate regions for further analysis. The second step involves the use of a non-linear method based on cepstrum analysis of the candidate regions for detecting the shadow points inside those regions.

The next section presents the proposed cepstrum based shadow detection method. Results of the proposed method and comparisons with previous approaches are presented in Section 3.

2   Cepstrum Analysis for Moving Shadow Detection

The proposed method for moving shadow detection consists of two parts. In the first part, a method based on hybrid background subtraction [12] is used to determine the moving regions. After determining moving regions, cepstrum analysis is carried out on detected moving regions for yielding the regions with shadows.

The proposed cepstrum analysis method for shadow detection is composed of two parts. The first part includes the separation of the moving regions into 8x8 blocks and the application of the 2D cepstrum to the blocks of interest and their corresponding background blocks to decide whether the texture and color properties are preserved for that moving block or not. If it is decided that the properties are preserved for the block, the algorithm proceeds with the second part. If not, the detection algorithm marks the block as moving object block. In the second part, a more detailed pixel-based approach is considered. 1D cepstrum is applied to each pixel belonging to the block to decide if the pixel is a moving shadow pixel or object pixel. The following subsections present the parts of the proposed cepstrum analysis method.

Part I. Cepstral Analysis of Blocks: The cepstrum $\hat{x}[n]$ of a signal x is defined as the inverse Fourier transform of the log-magnitude Fourier spectrum of x. Let $x[n]$ be a discrete signal, its cepstrum $\hat{x}[n]$ is defined as follows:

$$\hat{x}[n] = F^{-1}\{\ln(|F\{x[n]\}|)\}$$

where $F\{.\}$ represents the discrete-time Fourier Transform, $|.|$ is the magnitude, $\ln(.)$ is the natural logarithm and $F^{-1}\{.\}$ represents the inverse discrete-time Fourier Transform operator. In our approach, we use both one-dimensional (1D) and two-dimensional (2D) cepstrums for shadow detection.

Moving regions in video are divided into 8x8 moving blocks as a subset of the whole moving region. Let the i-th moving 8x8 block be defined as $\mathbf{R}_i$. Then, 2D cepstrum of $\mathbf{R}_i$, $\hat{\mathbf{R}}_i$ is defined as follows:

$$\hat{\mathbf{R}}_i = F_{2D}^{-1}\{\ln(|F_{2D}\{\mathbf{R}_i\}|)\}$$

where $F_{2D}\{.\}$ is the 2D discrete-time Fourier Transform and $F_{2D}^{-1}\{.\}$ is the inverse discrete-time Fourier Transform operator.

Similarly, let the i-th corresponding background block for the current image frame be defined as $\mathbf{B}_i$ and its 2D cepstrum as $\hat{\mathbf{B}}_i$. A difference matrix $\mathbf{D}_i$ for the i-th block can be defined as: $\mathbf{D}_i = |\ \hat{\mathbf{R}}_i - \ \hat{\mathbf{B}}_i|$. Theoretically if the block of interest is part of a shadow it should have the following property:

$$\mathbf{R}_i = \alpha\ \mathbf{B}_i$$

where α is a positive real number less than 1. The effect of this on the difference matrix in the 2D cepstral domain is: $\mathbf{D}_i$ having only the (1,1)-indexed value different than zero because of the scaling by constant α. Other entries of $\mathbf{D}_i$ should be equal to zero. So the distance metric is defined as:

$$m_i = \sum_{(a,b)\neq(1,1)} \mathbf{D}_i(a,b)$$

Notice that this operation is done for R, G and B values of the block separately. Therefore, the distance metric M_i is used as follows:

$$M_i = \sqrt{m_{i,r}^2 + m_{i,g}^2 + m_{i,b}^2}$$

where $m_{i,r}$, $m_{i,g}$ and $m_{i,b}$ is the R, G and B component distance metric, respectively. Therefore, the decision algorithm for the first part is:

$$R_i : \begin{cases} \text{moving shadow block,} & \text{if } M_i < \kappa \\ \text{moving object block,} & \text{otherwise} \end{cases}$$

where κ is a determined threshold. After detecting possible candidate 8 by 8 shadow regions, we examine each pixel of such regions one by one to determine the exact boundary of shadow pixels as follows.

Part II. Cepstrum Analysis of Pixels: Red, Green and Blue values and the estimated background values of the pixel positioned at $\mathbf{x} = (x_1, x_2)$ in the nth frame are defined as:

$$\mathbf{v}_{\mathbf{x},n} = (r_{\mathbf{x},n} \ g_{\mathbf{x},n} \ b_{\mathbf{x},n}) \quad \mathbf{b}_{\mathbf{x},n} = (br_{\mathbf{x},n} \ bg_{\mathbf{x},n} \ bb_{\mathbf{x},n})$$

Theoretically, a shadow pixel positioned at $\mathbf{x}$ in n^{th} frame should have the property: $\mathbf{v}_{\mathbf{x},n} = \alpha \mathbf{b}_{\mathbf{x},n}$, where α is a positive real number less than 1. Thus, the shadow pixel frame value is an α scaled version of the same positioned background pixel value in the RGB-space. As a result of this, we obtain the following cepstral relation:

$$\hat{\mathbf{v}}_{\mathbf{x},n}[1] \neq \hat{\mathbf{b}}_{\mathbf{x},n}[1] \tag{1}$$

$$\hat{\mathbf{v}}_{\mathbf{x},n}[i] = \hat{\mathbf{b}}_{\mathbf{x},n}[i], \ i = 2, 3, ... \tag{2}$$

We use a DFT of size 4 in our implementation and check the second, third and fourth cepstral coefficients, $\hat{\mathbf{v}}_{\mathbf{x},n}[2]$, $\hat{\mathbf{v}}_{\mathbf{x},n}[3]$, $\hat{\mathbf{v}}_{\mathbf{x},n}[4]$ and their counterpart cepstral coefficients of background location, $\hat{\mathbf{b}}_{\mathbf{x},n}[2]$, $\hat{\mathbf{b}}_{\mathbf{x},n}[3]$, $\hat{\mathbf{b}}_{\mathbf{x},n}[4]$. They should be equal if the pixel of interest is a shadow pixel. First cepstral coefficients, $\hat{\mathbf{v}}_{\mathbf{x},n}[1]$ and $\hat{\mathbf{b}}_{\mathbf{x},n}[1]$ should be different due to the effect of the natural logarithm of coefficient α. Using this fact, we define a difference vector: $\mathbf{d}_{\mathbf{x},n} = |\hat{\mathbf{v}}_{\hat{\mathbf{x}},n} - \hat{\mathbf{b}}_{\hat{\mathbf{x}},n}|$. Shadow detection method for moving pixels inside the block is given as follows:

$$\mathbf{x} : \begin{cases} \text{moving shadow pixel,} & \text{if } \mathbf{d}_{\mathbf{x},n}[2] \ \& \ \mathbf{d}_{\mathbf{x},n}[3] \ \& \ \mathbf{d}_{\mathbf{x},n}[4] < \tau \\ \text{moving object pixel,} & \text{otherwise} \end{cases} \tag{3}$$

where τ is an adaptive threshold changing its value as a function of the background pixel value for the current image frame.

3   Experimental Results and Conclusions

In this section, the outcomes of the proposed algorithm are presented and comparisons with some of the previous approaches are made. The benchmark test set available in [14] is used in this paper as it is widely referenced by most of the researchers working in the field. Each video sequence in the benchmark test set has different sequence type, shadow strength, shadow size, object class, object size, object speed and noise level.

The video sequences of campus raw (Fig. 2a) have very low shadow strength as well as high noise level. In Fig. 2b, it is clearly seen that two moving objects are detected perfectly and most of the moving shadow points on the ground are marked with success. In order to compare the performance of the proposed method with the others, quantitative measures are used. In this study, shadow detection accuracy η and shadow discrimination accuracy ξ metrics introduced in [10] are used as the quantitative measures for comparison purposes. The reason for selecting [10] for comparison is due to the existence of detailed classification schemes and utilization of different approaches available in the literature

(a) Original video frame (b
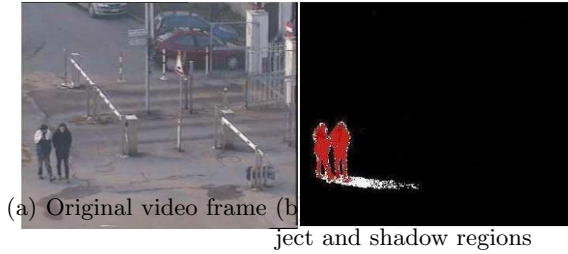
ject and shadow regions

Fig. 2.: "Campus" video sequence

for shadow detection in its content. Table 2 and Table 3 summarizes the performance of the proposed method and the other methods using the same benchmark test set. In the tables, the abbreviations SNP, SP, DNM1, DNM2 and CB stands for the statistical non-parametric approach, statistical parametric approach, deterministic non-model based approach using color exploitation, deterministic non-model based approach using spatial redundancy exploitation and the proposed cepstrum based approach, respectively. The ξ and η values in percentage for the proposed approach are commonly better than the SNP, SP, DNM1 and DNM2 approaches used by the other researchers in the literature.

Table 1.: Shadow detection accuracy (η) values in percentage

	Campus	Highway I	Highway II	Intelligent Room	Laboratory
SNP	80.58	81.59	51.20	78.63	84.03
SP	72.43	59.59	46.93	78.50	64.85
DNM1	82.87	69.72	54.07	76.52	76.26
DNM2	69.10	75.49	60.24	71.68	60.34
CB	**84.21**	**77.38**	**62.73**	**80.67**	**83.26**

Table 2.: Shadow discrimination accuracy (ξ) values in percentage

	Campus	Highway I	Highway II	Intelligent Room	Laboratory
SNP	69.37	63.76	78.92	89.92	92.35
SP	74.08	84.70	91.49	91.99	95.39
DNM1	86.65	76.93	78.93	92.32	89.87
DNM2	62.96	62.38	72.50	86.02	81.57
CB	**81.35**	**85.34**	**86.88**	**93.56**	**94.90**

The proposed cepstral domain method can determine shadow regions retaining the underlying color and texture of the background region. In benchmark data test sets, it is observed that proposed method gives successful results. The shadow pixels and object pixels are segmented accurately in all video sequences. Finally, quantitative measures are defined for comparison with previous

approaches. The detection and discrimination rate comparisons show that the proposed method gives better results than other approaches available in the literature.

References

1. Toreyin, B. U., Cetin, A. E.: Shadow Detection Using 2D Cepstrum. In: Proceedings of SPIE, the International Society for Optical Engineering, vol. 7338, pp 733809-733809-7 (2009)
2. Cucchiara, R., Grana, C., Piccardi, M., Prati, A.: Detecting Objects, Shadows and Ghosts in Video Streams by Exploiting Color and Motion Information. In: Proceedings of the 11th International Conference on Image Analysis and Processing, pp. 360-365 (2001).
3. Cucchiara, R., Grana, C., Piccardi, M., Prati, A., Sirotti, S.: Improving Shadow Suppression in Moving Object Detection with HSV Color Information. In: Proceedings of IEEE International Conference on Intelligent Transportation Systems, pp. 334-339 (2001)
4. Chen, B.: Indoor and Outdoor People Detection and Shadow Suppression by Exploiting HSV Color Information. In: Proceedings of the The Fourth International Conference on Computer and Information Technology, pp. 137-142 (2004)
5. Jiang, C., Ward, M. O.: Shadow Identification. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, pp. 606-612 (1992)
6. Stauder, J., Mech, R., Ostermann, J.: Detection of Moving Cast Shadows for Object Segmentation. In: IEEE Transactions on Multimedia, vol. 1, no. 1, pp. 65-76 (1999)
7. Onoguchi, K.: Shadow Elimination Method for Moving Object Detection. In: Proceedings of Fourteenth International Conference on Pattern Recognition. vol. 1. pp. 583-587 (1998)
8. Sonoda, Y., Ogata, T.: Seperation of Moving Objects and Their Shadows, and Their Application to Tracking on the Loci in the Monitoring Images. Proceedings of Fourth International Conference on Signal Processing, vol. 2, pp. 1261-1264 (1998)
9. Prati, A., Mikic, I., Grana, C., Trivedi M. M.: Shadow Detection Algorithms for Traffic Flow Analysis: A Comparative Study. In: Proceedings of IEEE International Conference on Intelligent Transportation Systems, pp. 340-345 (2001)
10. Prati, A., Mikic, C., Trivedi, M. M., Cucchiara, R.: Detecting Moving Shadows: Algorithms and Evaluation. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 7, pp. 918-923 (2003)
11. Prati, A., Cucchiara, R., Mikic, C., Trivedi, M. M.: Analysis and Detection of Shadows in Video Streams: a Comparitive Evaluation. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 571-576 (2001)
12. Colins, R., Lipton, A., Kanade, T.: A System for Video Surveillance and Monitoring. 8-th Int Topical Meeting on robotics and remote systems, American Nuclear Society, (1999)
13. Joshi, A., Atev, S., Masoud, O., Papanikopoulos, N.: Moving Shadow Detection with Low- and Mid-Level Reasoning. In: Proceedings of IEEE International Conference on Robotics and Automation, (2007)
14. Video Surveillance Online Repository, http://www.openvisor.org