Detection and Tracking of Moving Cast Shadows

Jürgen Stauder[1] *and Roland Mech*[2]

[1]IRISA/INRIA, France, Jurgen.Stauder@irisa.fr, http://www.irisa.fr/prive/Jurgen.Stauder
[2]University of Hannover, Germany, mech@tnt.uni-hannover.de, http://www.tnt.uni-hannover.de/~mech
(WIAMIS'99 Workshop on Image Analysis for Multimeadia Interactive Services, 31.5.-1.6.1999, Berlin, Germany)

Abstract

This paper presents a method for detection and tracking of moving cast shadows on a dominating scene background in a monocular video sequence. The method assumes moving shadows on a dominant smooth shaped background. The shadow causing light sources are assumed to be strong enough to cause visible temporal frame differences by moving cast shadows. These differences are detected and classified into regions covered and regions uncovered by moving shadows. Entire moving cast shadows are detected and tracked by temporal integration of the covered background regions while subtracting the uncovered background regions. Results obtained with various test sequences show that those shadows are detected and tracked successfully that are also easily visible to a human observer.

I Introduction

A moving cast shadow in a video sequence is generated by an object occluding a light source with respect to another object. It's generation depends on the scene geometry. This includes position and shape of the objects but also position and kind of light sources. The light sources, the objects and the generated cast shadows in a scene are linked by light rays. If knowledge on the objects is available, these light rays can be used to detect cast shadows in video images [7][1]. If cast shadows can be detected by other methods, the light rays can be used to determine the 3D object shape [3].

Furthermore, cast shadows play a role in object segmentation. For video editing, surveillance, remote data bases and other typical MPEG-4 applications [2], moving objects need to be segmented. As cast shadows move together with the objects, they are likely to be segmented erroneously as objects.

Without knowing any 3D geometry, image regions changed by moving cast shadows can be identified using assumptions on the scene and on the illumination [12][11][10][6][15]. In [13], all assumptions have been recently combined to detect image regions changed by moving cast shadows. By a combination with the Analysis Model of the European project COST 211 [4][9][5], it has been shown that hereby shadow artifacts of object segmentation can be reduced drastically.

In this paper, the detection and tracking of entire cast shadows is addressed. Therefore, a nonlinear filtering approach will be proposed that processes the image regions that are changed by moving cast shadows. First, these image regions will be detected as in [13]. Then, they will be classified into two classes. Regions of the first class are changed, because they have been covered by a moving cast shadow. Regions of the second class are changed, because they have been uncovered by a moving cast shadow. Finally, a temporal integration applied to the classified regions will be proposed to detect and track entire cast shadows. An in-depth description of the proposed algorithm will be published in [14].

The paper is organized as follows. In Section II, the detection of image regions changed by moving cast shadows is reviewed. Section III describes the new shadow detection and tracking algorithm. In Section IV, experimental results are presented and Section V gives a conclusion.

II Detection of Changes Caused by Shadows

In [13], image signal changes caused by moving cast shadows are detected based on four assumptions:

(1) *Shadow causing light source is strong with respect to other light.*

This implies that the frame difference is large in regions changed by cast shadows [12]. Thus, a change detection is applied to a current and to a preceding image from the image sequence [11]. In this paper, camera and background are assumed to be static. Thus, the change detection mask contains only image regions changed by moving objects or moving cast shadows. If camera or background are moving, a global motion compensation may be performed as described in [9]. Three more assumptions are used to retain only those changes that are caused my moving cast shadows.

(2) *Shadows are on a dominant background.*

Then, image regions changed by moving cast shadows should have the same motion as the background. If camera and background are static, this means zero motion. Changed image regions with zero motion are detected by two steps. First, in the current and in the preceding image edges are detected. Second, the edges are classified into moving edges, if the corresponding frame difference has high frequencies. Other edges are classified as static edges. Changed image regions with zero motion are defined to be regions with no edges or regions with static edges. [10][11].

(3) *The background has a smooth shape.*

This implies that the illumination incident on the background is spatially smooth [12]. Then, image regions changed by moving cast shadows should have a smooth frame ratio between the current and the preceding image. The frame ratio is tested for smoothness by evaluating its variance that is calculated for each pel in a local neighborhood.

(4) T*he cast shadows have a penumbra.*

At certain scene geometries, a cast shadow has thick contours, called penumbra, with a soft illumination transition [6][15]. Thus, image regions changed by moving cast shadows should preferably have thick contours.

III Shadow Detection and Tracking

For detection of entire regions of moving cast shadows, a nonlinear filtering approach is proposed in this paper. The filter decomposes into two parts.

The first part of the filter has the character of a derivation. Inputs are a current image $s_k(x,y)$ and the preceding image $s_{k-1}(x,y)$ from a video sequence. Output is a trinary signal $d_k(x,y)$ assigning to each pel of the current image one of the following three values:

- Value 0: No changes caused by cast shadows
- Value -1: Image has been uncovered by a cast shadow since the preceding image
- Value 1: Image has been covered by a cast shadow since the preceding image

The first part of the filter is realized by the detection of image regions that are changed by moving cast shadows (see Section II) and a following classification according to

$$d_k(x,y) = \begin{cases} 0 & \text{no shadow changes} \\ -1 \quad \text{if} & s_k(x,y) \geq s_{k-1}(x,y) \\ 1 & s_k(x,y) < s_{k-1}(x,y) \ . \end{cases} \tag{1}$$

The second part of the filter has the character of an integration. The result of the integration will be a binary shadow mask $m_k(x,y)$ indicating by values 0 and 1 the absence and or the presence of a moving cast shadow. Assuming that for the preceding image such a mask $m_{k-1}(x,y)$ is already present, the integration operation is simply

$$\tilde{m}_k(x,y) = m_{k-1}(x,y) + d_k(x,y) \tag{2}$$

with a following clipping

$$m_k(x,y) = \begin{cases} 1 & \text{if} \quad \tilde{m}_k(x,y) \geq 1 \\ 0 & \quad\quad \tilde{m}_k(x,y) \leq 0 \end{cases} . \tag{3}$$

Finally, the shadow mask is median filtered, simplified and small regions are eliminated. Also, image regions that are covered by a moving object are deleted from the shadow mask to prevent artifacts of non-detected background uncovered by a shadow while being covered by an object. Image regions covered by moving objects are determined by the COST 211 Analysis Model [4][9][5].

A cast shadow will be completely detected as soon as it has completely uncovered its location in the first frame. Due to the concept of derivation/integration, pels that are always covered by a cast shadow can not be detected by this algorithm.

IV Experimental Results

In this section, first results are presented that have been obtained for the test sequences shown in Fig. 1. The MPEG-4 “Table Tennis” scene is illuminated by one strong and another weak spot light generating two cast shadows at the left side of the player at the wall. Further, a reflection of the player can be seen on the table. The “Jürgen” scene is illuminated by three and the “Erik” scene by one spot light generating the main shadows at the left side of the persons. Here, in the image plane the shadows are directly neighbored to the persons.

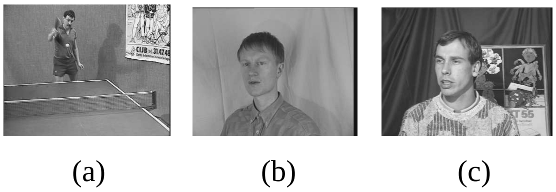

Figure 1: Sample images of natural image sequences: (a) “Table Tennis” (image no. 151, format SIF, 360 × 240 pels, frame rate 30Hz), (b) “Jürgen” and (c) “Erik” (both image no. 2, format CIF, 352 × 288 pels, 10Hz frame rate)

The temporal evolution of shadow tracking for the sequence ”Table Tennis” in Fig. 2 shows that the main shadow at the wall is well tracked. Portions of the shadow that are not detected are either shadowed since the scene cut and can thus not be detected by an integration of temporal changes. Or, they were covered by the moving person and were thus deleted from the mask (see Section III). A successful example is the shadow of the players arm. It is completely visible at all time. Further, its motion is sufficient large to make the shadow leaving quickly its position from the first image after the scene cut. The lower shadow of the person at the wall is not detected by the proposed algorithm. It is more diffuse and even hardly visible by a human observer. There is also a diffuse mirror image of the player on the table. It is mostly darker than the rest of the table, that is illuminated by the brighter wall. Thus, the players mirror image is partially detected as shadow.

Fig. 3 shows sample shadow tracking results for the test sequences “Jürgen” and “Erik”. The sample results show a phase, where the persons had moved most to its left side, such that a maximum area of the cast shadow has been integrated.

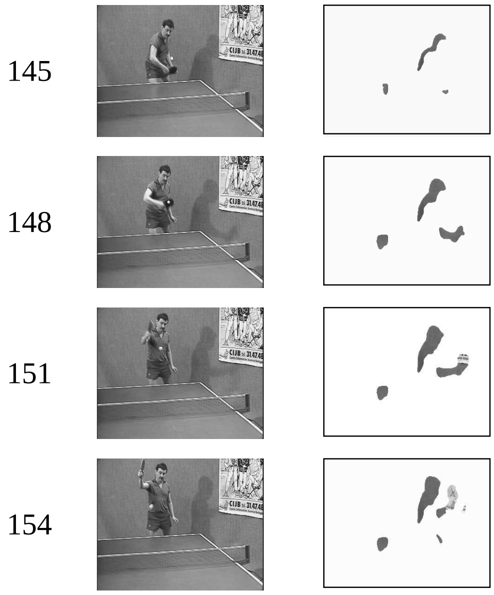

Figure 2: Shadow tracking results for the test sequence ”Table Tennis” for the images 145, 148, 151, 154, 169, 172, 175 and 178. The algorithm started with image $k-1$=133.

In Fig. 3(a), the weak and diffuse shadows at the wall at both sides of the speaker have not been detected. Due to small object motion, the temporal image signal changes are so weak that they are not detected by the change detection. Also, assumption (1) from Section II is violated, because the scene is illuminated by several strong spot lights. On the other hand, the shadow directly neighboring the left side of the speaker has been detected. It is less diffuse, because the person and the shadowed background are closer. In Fig. 3(b), the cast shadow is still not completely detected. A cast shadow is detected completely as soon as he has totally moved away from its position from the first image.

V Conclusions

This paper presents a method for detection and tracking of moving cast shadows on a dominating scene background in a monocular video sequence. First, frame differences caused by moving cast shadows are detected and classified into regions covered and regions uncovered by a moving cast shadow. A moving cast shadow is detected

Figure 3: Shadow tracking results for test sequences (a) "Jürgen" (image 23) and (b) "Erik" (image 38)

and tracked by temporal integration of the covered background regions while subtracting the uncovered background regions.

The presented results show that visible cast shadows that are strong enough to be easily visible by a human are well detected and tracked. The algorithm does not detect weak and diffuse shadows, because they do not cause temporal signal changes that are strong enough to be detected. Further, portions of a cast shadow that are shadowed since the beginning of the sequence are not detected due to the derivation/integration principle of the algorithm. Here, a combination with single frame shadow detection methods would be helpful.

References

[1] M. Bejanin, A. Huertas, G. Mediani, R. Nevatia: "Model validation for change detection", 2nd Int'l IEEE workshop on Applications of Computer Vision, Los Alamitos, USA, 5.-7. December 1994, pp. 160-167.

[2] L. Chiariglione: "MPEG and Multimedia Communications", IEEE Trans. CSVT, Vol.7, No.1, Feb.1997.

[3] A.M. Bruckstein, R.J. Holt, Y.D. Jean, A.N. Netravali: "On the use of shadows in stance recovery", Bell Laboratories, Lucent Technologies, Technical Memorandum, BL0113380-971222-04TM, December 1997.

[4] A. Alatan, L. Onural, M. Wollborn, R. Mech, E. Tuncel, T. Sikora: "Image sequence analysis for emerging interactive multimedia services - The European COST 211 framework", IEEE Trans. CSVT, Vol. 8, No. 7, Nov. 1998, pp. 802-813.

[5] COST 211^{quat} Project Web Pages: http://www.teltec.dcu.ie/cost211

[6] C. Jiang, M.O. Ward: "Shadow segmentation and classification in a constrained environment", CVGIP: Image Understanding, Vol. 59, No. 2, March 1994, pp. 213-225.

[7] D. Koller, K. Danilidis, H.-H. Nagel: "Model-Based object tracking in monocular image sequences of road traffic scenes", Int'l J. of Computer Vision, Vol. 10, No. 3, 1993, pp. 257-281.

[8] B.-T. Phong: "Illumination for computer generated pictures", Communications of the ACM, Vol. 18, No. 6, June 1975, pp. 311-317.

[9] R. Mech, M. Wollborn: "A noise robust method for 2D shape estimation of moving objects in video sequences considering a moving camera", Signal Processing, Vol. 66, No. 2, April 1998, pp. 203-217.

[10] J. Ostermann: "Modelling of 3D moving objects for an analysis-synthesis coder", Conference on Sensing and Reconstruction of 3D-Objects and Scenes, Atlanta, 15.-16.2.1990, SPIE Vol. 1260, pp. 240-249.

[11] J. Ostermann: "Segmentation of image areas changed due to object motion considering shadows", in "Multimedia Communications and video coding", Plenum Press, NY, 1996.

[12] G.G. Sexton, X. Zhang: "Suppression of shadows for improved object discrimination", IEEE Colloquium on Image Processing for Transport Applications, London, UK, Dec. 1993.

[13] J. Stauder: "Segmentation of moving objects in presence of moving shadows", Int'l Workshop on Coding Techniques for Very Low Bit-rate Video, Linköping, Sweden, 28.-30. July 1997, pp. 41-44.

[14] J. Stauder, R. Mech, J. Ostermann: "Detection of Moving Cast Shadows for Object Segmentation", IEEE Trans. on Multi Media, Vol. 1, No.1, March 1999, pp. 65-76.

[15] W. Zhang, F. Bergholm: "An extension of Marr's signature based edge classification and other methods determining diffuseness and height of edges, and bar edge weight", Int'l Conf. on Computer Vision, Berlin, 11.-14. May 1993, pp. 183-191.