Segmentation of moving objects in presence of moving shadows

Jürgen Stauder
Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung
Universität Hannover, Appelstraße 9A, D–30167 Hannover 1, FRG
Phone: +49–511–762–5311
Email: stauder@tnt.uni–hannover.de *http://www.tnt.uni–hannover.de/~stauder*

Abstract

A new method for segmentation of moving objects in an image sequence in presence of moving shadows on static background is presented. The method distinguished between moving objects and moving shadows. Therefore, shadows are identified by a) a large frame difference, b) non–moving background texture, c) spatially homogeneous temporal change of shading and d) soft contours. The results show, that the proposed method is superiour to simple change detection or a method from literature evaluating only a) to c).

Background

For object–based coding [9] or content–based functionalities of the forthcoming ISO/MPEG–4 standard [7], moving objects need to be segmentated in image sequences. In the case of static camera and static background, the image area changed by moving objects, i.e. moving objects and uncovered background, can be detected by change detection [1][4][8]. Therefore, the frame difference between a current image and its predecessing image of an image sequence is evaluated by a threshold operation. Changes are assumed to be due to object motion or camera noise. The resulting change detection mask is then postprocessed [4][8]: It is cleared from uncovered background using motion information resulting to an object mask valid for the current image. Before this step, the object mask from the predecessing image, is or–connected with the current change detection mask as a temporal object memory.

Problem

This paper addresses the problem of moving object segmentation in presence of moving cast shadows on static background objects. In this case, a change detector cannot distinguish between moving objects and moving shadows (see Fig. 2 c–d). To solve this problem, image regions changed by moving shadows have to be identified and discriminated against those changed by moving objects.

State of the art

In [6][2], image regions corresponding to shadows are identified in a single image using known 3D object geometry, known illumination direction and known 3D geometry of a planar background for constraint environments as a traffic scene [6] or simple buildings [2].

In [5][13], shadows are identified in a single image by their contour, based on the prior knowledge, that the light emitting surfaces of light sources in typical scenes have a certain extent. Then, cast shadows do not have sharp contours but inherit a soft transition from shadowed to not shadowed areas called penumbra. The shadow contours are modelled by a Gaussian step function [13].

In [12], shadows are identified in a single image by comparison with a reference image showing the same scene without shadows. Avoiding such a reference image, in [11] the frame difference between a current image and its predecessing image from an image sequence is evaluated. The frame difference is pel–wise classified by three criteria:

1. First, the pels are classified into CHANGED und UNCHANGED by change detection as in [4]. The label CHANGED indicates a moving object or a moving shadow.
2. Secondly, all pels are tested for membership to edges classifying them into NO_EDGE, MOVING_EDGE or STATIC_EDGE. The label MOVING_EDGE indicates a moving object, whereas STATIC_EDGE indicates static background (with or without shadow).
3. Thirdly, those pels are labeled to be SHADED or NOT_SHADED, where the frame ratio in their neighbourhood can be or cannot be described, respectively, by a factor that is constant for the neighbourhood. The label SHADED indicates a moving shadow on static background with non–moving texture.

The decisions of the three criteria are evaluated by heuristic rules resulting to a binary segmentation mask of image regions changed by moving objects excluding moving shadows [11].

The segmentation method in [11] was found to give wrong results in the following critical cases:

- Moving shadows on the background with spatially inhomogeneous shading, e.g. due to a object being close to the background, may be classified as MOVING_EDGE and NOT_SHADED, leading to an inclusion of shadows into the object mask, see Fig. 2 e.
- Flat textured and shaded object regions may be classified erroneously as SHADED and thus excluded from object mask, see the person's right shoulder in Fig. 2 f.

Approach

In this paper, the approach from [11] is extended by a fourth criterion called shadow contour criterion. The new criterion is based the idea from [5][13] to identify shadows by their soft contours. To exploit this feature of shadows, the shadow contour detector first locates the contours of moving objects and moving shadows and then classifies them into object and shadow contours [3]. To locate the contours, the rim of the change detection mask is taken as contour. It is refined by spatially adapting the contour to high image luminance gradients. For contour classification, a contour model is used. On the contrary to [13], the transition of luminance in direction perpendicular to a contour is assumed to follow the linear model shown in Fig. 1.

The norm g of image luminance gradient and the height h of the luminance step for each pel of a contour is evaluated. By dividing height h by the norm g of the gradient for each pel of a contour, the width w of the luminance transition is calculated and then thresholded. Pels with a width w exceeding the threshold are labeled as SHADOW_CONTOUR. The other pels lying on contours are labeled as OBJECT_CONTOUR.

To integrate the new shadow contour criterion into the segmentation method from [11], the following simple additional heuristic rule was found to be effective: Pels that belong to an area changed by moving object according to [11] are reclassified to belong to a moving shadow if and only if there are more SHADOW_CONTOUR pels in a certain neighbourhood than OBJECT_CONTOUR pels.

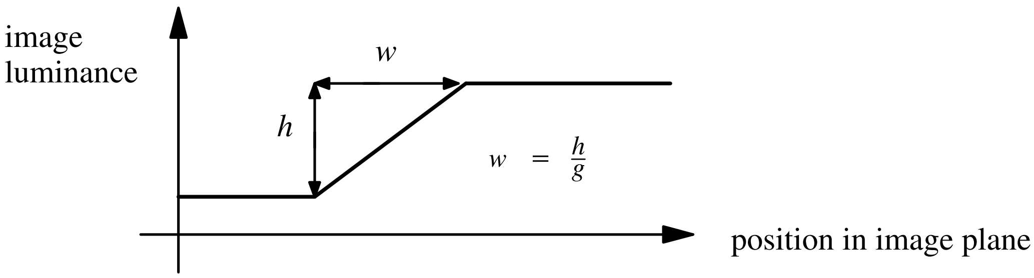

Fig 1 Model of an image luminance transition in direction perpendicular to a contour: The luminance transition is defined by the step height h and the norm g of the gradient. From h and g the width w can be calculated.

Results

Experimental results has been obtained for the change detection method from [4], the three–criteria–based segmentation method from [11] and the proposed four–criteria–based method. To present the results, all three methods are applied to succeeding image pairs of an image sequence. For all methods, the resulting segmentation masks are postprocessed [4][8] as described above. The motion information for postprocessing was automatically derived by 3D object based motion estimation [10]. The threshold for the edge width w was set to 2.4 pel.

Segmentation results have been obtained for test sequences ”Erik” and ”Jürgen” (CIF, 10Hz frame rate). As can be seen in Fig. 2 c–d the change detection detects moving shadows as moving objects. The three–criteria–based segmentation method from [11] behaves better, but some parts of moving shadows are still contained in the object mask, see Fig. 2 e, or some parts of the objects are missing in the object mask, see right shoulder in Fig. 2 f. Using the proposed four–criteria–based segmentation method, see Fig. 2 g–h, moving shadows are much better excluded from the object mask. In case of no moving shadows, as in test sequence ”Claire” (CIF, 10Hz frame rate), the proposed segmentation method shows the same results as reported for the method from [11].

References

[1] T. Aach, R. Mester: ”Statistical model–based change detection in moving video”, Signal Processing, Vol. 31, No. 2, March 1993, pp. 165–180.

[2] 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.

[3] N. Elouardi: ”Verfahren zur Detektion von Objektsilhouetten unter Ausnutzung von Vorwissen über Schlagschatten”, student thesis, University of Hannover, July 1997.

[4] M. Hötter, R. Thoma: ”Image segmentation based on object oriented mapping parameter estimation”, Signal Processing, Vol. 15, No. 3, October 1988, pp. 315–334.

[5] C. Jiang, M.O. Ward: ”Shadow segmentation and classification in a constrained environment”, CVGIP: Image Understanding, Band 59, Nr. 2, März 1994, S. 213–225.

[6] D. Koller, K. Danilidis, H.–H. Nagel: ”Model–Based object tracking in monocular image sequences of road traffic scenes”, Int’l Journal of Computer Vision, Vol. 10, No. 3, 1993, pp. 257–281.

[7] R. Mech, P. Gerken: ”Automatic segmentation of moving objects (partial results of core experiment N2)”, Doc. ISO/IEC JTC1/SC29/WG11 MPEG 96/1549, November 1996.

[8] R. Mech, M. Wollborn: ”A noise robust method for segmentation of movin objects in video sequences”, accepted for ICASSP ’97, Munich, April 21–24, 1997.

[9] Musmann, H. G.; Hötter, M.; Ostermann, J.: ”Object–oriented analysis–synthesis coding of moving images”, Signal Processing: Image Communication, Vol. 1, No. 2, November 1989 pp 117–137.

[10] J. Ostermann: ”Object–based analysis–synthesis coding (OBASC) based on the source model of moving rigid 3D objects”, Signal Processing: Image Communication, Vol. 6, No. 2, May 1994, pp. 143–161.

[11] J. Ostermann: ”Segmentation of image areas changed due to object motion considering shadows”, in Y. Wang u.a. (Ed.): ”Multimedia Communications and video coding”, Plenum Press, New York, 1996.

[12] G.G. Sexton, X. Zhang: ”Supression of shadows for improved object discrimination”, IEEE Colloquium on Image Processing for Transport Applications, London, UK, Dezember 1993, S. 9/1–9/6.

[13] 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. Mai 1993, S. 183–191.

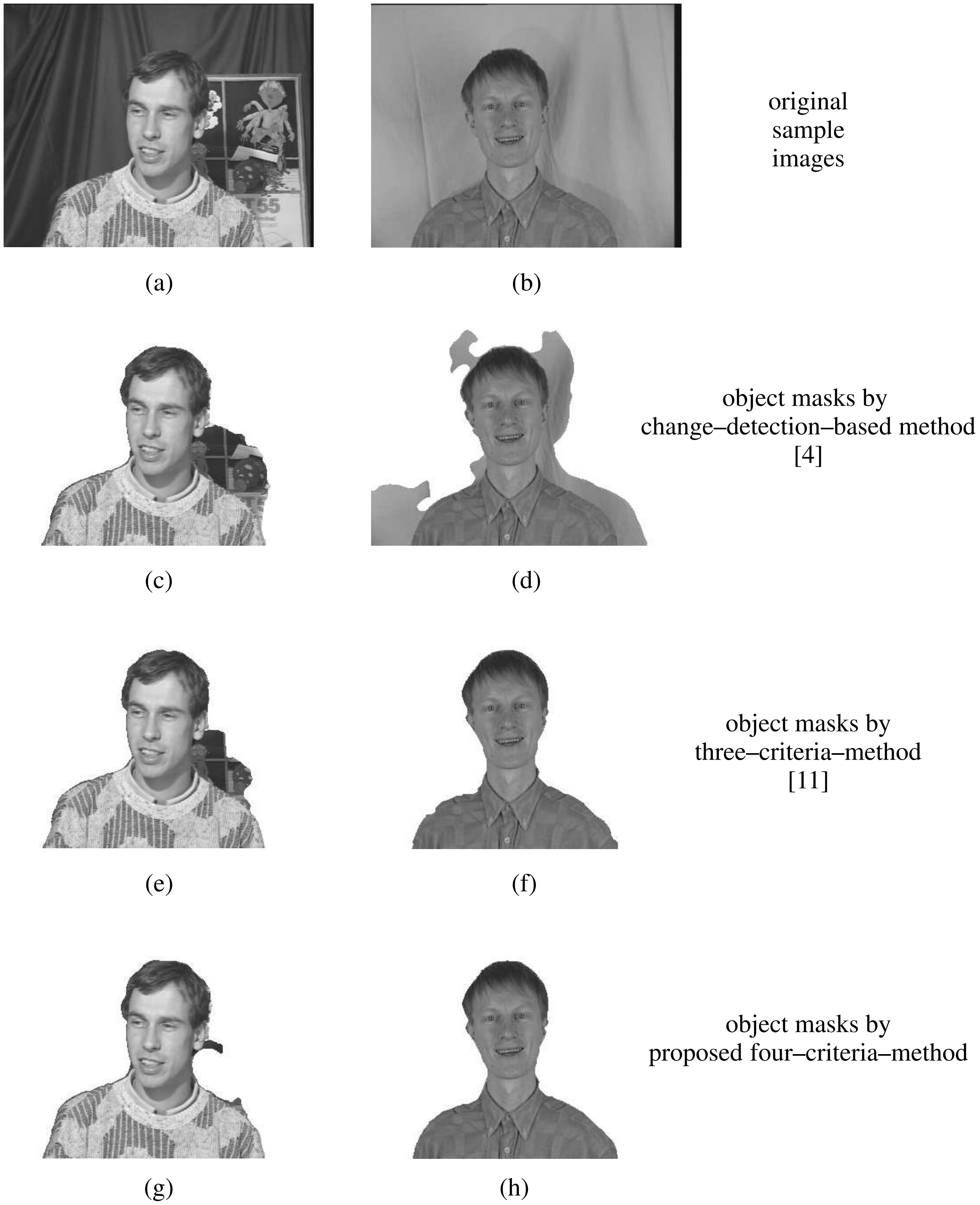

Fig 2 Sample image 42 and 23 for the image sequences (a) ”Erik” and (b) ”Jürgen”, respectively, and (c–f) obtained object masks.
Sample result (g) is the worse result of all 50 images of sequence ”Erik”. Sample result (h) is typical for all 25 images of sequence ”Jürgen”.