Spatio-temporal Shadow Segmentation and Tracking

Elena Salvador, Andrea Cavallaro and Touradj Ebrahimi

Signal Processing Institute (ITS)
Swiss Federal Institute of Technology (EPFL)
CH-1015 Lausanne, Switzerland

ABSTRACT

Shadow segmentation is a critical issue for systems aiming at extracting, tracking or recognizing objects in a given scene. Shadows can in fact modify the shape and colour of objects and therefore affect scene analysis and interpretation systems in many applications, such as video database search and retrieval, as well as video analysis in applications such as video surveillance. We present a shadow segmentation algorithm which includes two stages. The first stage extracts moving cast shadows in each frame of the sequence. The second stage tracks the extracted shadows in the subsequent frames. Tentative moving shadow regions are first identified based on spectral and geometrical properties of shadows. In order to confirm this tentative identification, shadow regions are then tracked over time. This second stage aims at exploiting the prior knowledge of a shadow detected in previous frames by evaluating its temporal behaviour. Shadow tracking is a difficult task, since colour, texture, and motion features in shadow regions cannot be used for solving the correspondence problem. Colour and texture change according to changes in the background's characteristics. The measurement of motion cannot be reliably computed for shadows. Therefore shadows may be described only by a limited amount of information. The proposed tracking algorithm makes use of this information and provides a reliability estimation of shadow recognition results of the first stage over time. This temporal analysis eliminates the possible ambiguities of the first stage and improves the efficiency of the overall shadow detection algorithm. The benefit of the proposed shadow segmentation and tracking algorithm is evaluated on both indoor and outdoor scenes. The obtained results are validated based on subjective as well as objective comparisons.

Keywords: Image Analysis, Shadows, Video Object, Tracking

1. INTRODUCTION

Shadows represent an interesting phenomenon for the analysis of images and image sequences. On one hand, they contain cues about the shape and the relative position of objects, as well as about the characteristics of surfaces and light sources in the scene. On the other hand, shadows can cause problems for algorithms which aim to track or recognize objects in the scene. A shadow cast by an object may either be in contact with the object, or disconnected from it. In the first case, the shadow distorts the object shape and colour, making the subsequent use of this information less reliable. In the second case, the shadow may be classified as a totally erroneous object in the scene.

Applications such as object-based coding and video databases, interactive video and remote surveillance are based on a representation of the video content in terms of video objects. The first step of object-based applications is the identification of the areas of a video sequence that correspond to meaningful objects. Many techniques to identify objects are based on motion information. Two main approaches can be identified that exploit motion information, namely motion segmentation and motion detection. Motion segmentation approaches classify clusters of pixels in the image that have similar motion. Motion detection approaches identify those pixels where motion exists. For both approaches, motion is measured by analysing the temporal changes of the intensity function in the image sequence. Since shadows generate temporal changes, the detection of a moving object may include the detection of its shadow or part of its shadow. The presence of shadows cast by moving objects misleads both motion segmentation and motion detection approaches. In order to obtain an accurate object identification, techniques for the explicit detection of shadows can be used.

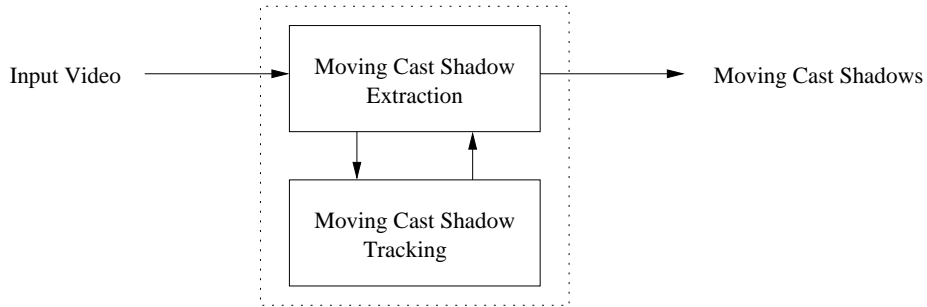E-mail addresses of the authors: {elena.salvador, andrea.cavallaro, touradj.ebrahimi}@epfl.ch

Figure 1. Block diagram of the proposed shadow segmentation and tracking algorithm

In this paper, an algorithm for the segmentation of moving cast shadows in video sequences recorded by a static camera is proposed. The algorithm includes two stages. The first stage extracts moving cast shadows in each frame of the sequence. The second stage tracks the extracted shadows in the subsequent frames. The complete system is depicted in Fig. 1 and described in the following sections. The presence of a shadow is first hypothesized by exploiting the assumption that a shadow region is darker than the same region if there was no shadow. Then this hypothesis is verified by analyzing photometric invariant color features and geometrical information. The shadow regions segmented in the first stage are then tracked over time. Tracking provides a reliability estimation for the moving cast shadow extraction stage. This allows us to remove shadows that have a low temporal reliability.

The major contribution of this paper resides in the definition and implementation of a shadow tracking strategy and its use for reliability estimation. Given the nature of shadows, shadow tracking is a difficult task. Shadows do not possess invariant shape, colour nor texture properties. Colour and texture in shadow regions depend on the characteristics of the background the shadow is cast on and vary with time. They cannot be therefore exploited for establishing a correspondence between instances of shadows over frames. The techniques proposed in the literature for tracking video objects cannot therefore be directly extended to the problem of tracking shadows. This motivated us to develop a method for tracking moving cast shadows.

The paper is organised as follows. The state of the art in the framework of shadow detection is presented in Sec. 2. The proposed method is described in Sec. 3. Section 4 shows experimental results and finally Sec. 5 draws the conclusions.

2. STATE OF THE ART

Shadow detection techniques in the literature can be classified into two groups: model-based, and property-based techniques. Model-based techniques rely on models representing the *a priori* knowledge of the geometry of the scene, the objects, and the illumination. Property-based techniques identify shadows by using properties such as the geometry, brightness or colour of shadows.

Model-based techniques are designed for specific applications, such as aerial image understanding[1–6] and video surveillance.[7–9] They are based on matching sets of geometric features such as edges, lines or corners to 3D object models. Most of these model-based schemes handle, therefore, only simple objects. Many of the proposed approaches extract as shadows all the dark regions in an image, identified through thresholding, that lie next to an object in the direction of the illumination. These schemes can be quite successful when shadows conform to this hypothesis. However, they are generally only applicable to the specific application they are designed for.

To overcome the above mentioned limitations, spectral and geometric characteristics of shadows can be exploited. Luminance is exploited by analysing edges,[10, 11] histograms[12] or texture information.[13] In,[14] background regions which are covered or uncovered by a moving cast shadow are detected. The detection is performed by searching, from frame to frame, for static edges and uniform changes of shading in the background texture. The penumbra of the shadow needs also to be analysed. This may represent a limitation when shadows present

sharp edges, that is when the illumination source is far from the objects. This is a typical case in outdoor sequences.

Luminance, chrominance, and gradient density information are used in.[15] A combined shadow confidence score is derived for foreground regions that allows to separate a cast shadow from the object that is casting it. The separation is based on the hypothesis that shadow areas are not textured. This assumption is valid for applications like traffic surveillance when shadows are cast on the road. Hue and saturation are exploited, together with motion information, in[16] to detect moving objects, shadows and ghosts in video sequences. The detection of shadows is based on the observation that shadows change significantly the lightness of an area without greatly modifying the colour information.

A method for real-time shadow detection for video conference applications is proposed in.[17] The algorithm uses colour information in the YUV colour space in order to avoid time consuming colour transformations. The authors observed that a shadow reduces the YUV values of a point linearly with respect to the same point in light. This observation is used in the method to eliminate the shadows. In,[18] a statistical model of the change in appearance of a pixel when shadowed given its appearance when illuminated is exploited in an algorithm for separating moving objects from their cast shadows. The parameters of the method, which runs in real-time, are optimized for traffic monitoring.

A background subtraction algorithm based on a computational colour model which separates the brightness from the chromaticity component of a pixel is presented in.[19] Once the background has been modeled statistically on the basis of the proposed colour model, the algorithm classifies each pixel in the frame under analysis into four categories: original background, shadowed background, highlighted background, and moving object.

A physical-based approach to distinguish material changes from shadow boundaries in chromatic images is presented in.[20] This method takes into account the effect of ambient illumination which is usually neglected by other approaches. Similarly, colour ratios are used in.[21] An estimate of the illumination field from the integration of the edge information is also derived. In,[22] a method to process 3-band colour images to locate and subsequently remove shadows is presented. A camera calibration stage is needed to generate a 1-D illumination invariant shadow-free image, that is used together with the original image to locate shadow edges. Thresholding out image edges that are due to shadows and re-integrating them allows to obtain full colour shadow free images. Colour information is used also in,[23] where a physics-based approach for moving shadow detection in outdoor scenes is proposed. A spatio-temporal albedo normalization test and the dichromatic reflection model are used in estimation and verification process.

Geometric properties of shadows are generally less exploited than spectral properties, but they can provide valuable information for shadow identification. In,[24] both intensity and geometry constraints are used to detect and classify shadows in images of a constrained, simple environment. In,[25] geometry information is combined with colour information to detect cast shadows in still images. A limitation of the method is the active process required to determine the location of the light. The method cannot, therefore, be used to automatically recognize shadows. The approach proposed in[26] overcomes this limitation by presenting a method for the estimation of the projection of the light source in the image plane. This estimation is used to segment shadows in video sequences. The method requires the manual segmentation of the shadow in the first frame. Moreover, shadows are assumed to be fully included in the images.

Very little work can be found in the literature which addresses the problem of shadow tracking. Stauder[27] presents a method for detection and tracking of moving cast shadows in monocular video sequences. Temporal frame differences are detected and classified into regions covered and regions uncovered by moving shadows. Entire moving shadows are tracked by temporal integration of the covered background regions while subtracting the uncovered background regions. Portions of a cast shadow that are shadowed since the beginning of the sequence are not detected by the proposed method. Moreover, scenes with only one moving object are presented. This represents a limitation in real world scenes.

3. PROPOSED METHOD

The proposed approach includes two stages. The first stage extracts moving cast shadows from each frame and is detailed in Sec. 3.1. The second stage tracks the extracted shadows and is presented in Sec. 3.2.

3.1. Moving Cast Shadows Extraction

The first stage of the proposed algorithm aims at extracting moving cast shadows in each frame of the sequence. This stage exploits the property that shadows cast on a surface reduce the surface intensities. To this end, changes in the intensity values of the three colour channels R, G, B are analysed. The variation of intensity values is checked by comparing the current frame and a reference frame. The reference frame represents the background of the scene, an image which does not contain dynamic objects nor shadows due to moving objects. Static shadows, that is shadow due to static objects such as buildings, parked cars, etc., can be present in the image.

The output of a background subtraction algorithm[28] are analysed to find candidate shadow points. The image difference $\mathbf{D}(x, y)$, computed as $\mathbf{D}(x, y) = \mathbf{I}(x_r, y_r) - \mathbf{I}(x, y)$ at each pixel (x, y) position belonging to the detected objects, is considered. The image $\mathbf{I}(x, y) = (R(x, y), G(x, y), B(x, y))$ and (x_r, y_r) belongs to the reference image and corresponds to the pixel (x, y) under analysis.

In a noise free case, the condition $\mathbf{D}(x, y) > \mathbf{0}$ would suffice to state that the pixel (x, y) belongs to a candidate shadow. In real situations, the noise introduced by the acquisition process alters the above test, so that it becomes $\mathbf{D}(x, y) > \mathbf{b}$, where $\mathbf{b} = (b_1, b_2, b_3)$ takes care of the distortions introduced by the noise. In addition, to obtain a more robust analysis for each pixel position, we extend the analysis to an observation window, $\mathcal{W}_{(x,y)}$, centered in (x, y). In $\mathcal{W}_{(x,y)}$, we analyse the sum of differences $\mathbf{D}_w(x, y)$, given by

$$\mathbf{D}_w(x, y) = \frac{1}{(2N+1)(2M+1)} \sum_{i=-N}^{N} \sum_{j=-M}^{M} \mathbf{D}(x+i, y+j) \tag{1}$$

If the sum of differences is larger than a threshold $\mathbf{b}$, then (x, y) belongs to a candidate shadow. The threshold $\mathbf{b}$ is content dependent and should be therefore tuned for each sequence. To avoid the tuning of the threshold we employ a statistical approach. This approach is based on the assumption that the noise in the signal $\mathbf{D}_w(x, y)$ respects a certain distribution. We analysed the difference $\mathbf{D}_w(x, y)$ in several sequences for indoor and outdoor scenes when no activity is present, and we derived that the measurement data follow a Gaussian distribution. The two parameters of the distribution are $\sigma = 2.0$ and $\mu = 0$. The statistical approach is based on a significance test.[28] The statistical test then checks the validity of the hypothesis that a sample $\mathbf{D}_w(x, y)$ comes from the Gaussian distribution. The pixel (x, y) is defined as candidate shadow pixel if the significance test is respected in all colour channels.

The first level of analysis leads to the detection of shadow pixels but also of object pixels. The analysis of changes in the invariant colour features allows to refine the hypothesis. Photometric colour invariants are functions which describe the colour configuration of each image point discounting shadings, shadows and highlights.[29] Therefore, the presence of a shadow does not alter the value of the invariant colour features. On the contrary, a material change modifies their value.

The identification of pixels satisfying this second evidence is achieved by analysing the difference in the invariant feature values, $\mathbf{d}(x, y)$, computed as

$$\mathbf{d}(x, y) = |\mathbf{Inv}(x_r, y_r) - \mathbf{Inv}(x, y)| \tag{2}$$

Among the different photometric invariant colour features, as stated in,[30] we adopted the $c_1 c_2 c_3$ model. Thus the function $\mathbf{Inv}(x, y)$ becomes

$$\mathbf{Inv}(x, y) = (c_1(x, y), c_2(x, y), c_3(x, y)) \tag{3}$$

As stated above, the condition $\mathbf{d}(x, y) = \mathbf{0}$ would suffice to state that the pixel (x, y) belongs to a candidate shadow in a noise free case. In real situations, the noise introduced by the acquisition process alters the above test. For this reason, the test becomes $\mathbf{d}(x, y) < \mathbf{f}$, where $\mathbf{f} = (f_1, f_2, f_3)$ takes care of the distortions introduced by the noise. As for the first level, we consider a window $\mathcal{W}_{(x,y)}$, centered in (x, y), and we analyse the sum of differences $\mathbf{d}_w(x, y)$. If $\mathbf{d}_w(x, y) < \mathbf{f}$, then (x, y) belongs to a candidate shadow. The setting of the threshold $\mathbf{f}$ is
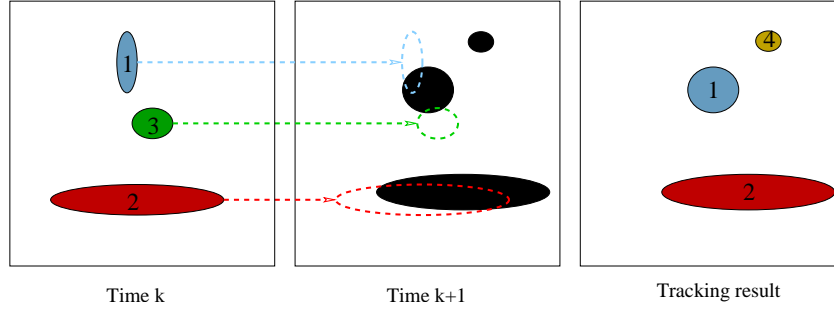
Figure 2. Tracking example

driven by experiments on different sequences. The threshold **f** is a stable parameter, which is different from the one used for RGB, because the dynamic range of the invariant is smaller than the one of the RGB components.

The last evidence about the existence of a shadow is derived from geometrical properties. This verification is based on analysing the boundary of the candidate shadow regions and testing the position of shadows with respect to objects. A necessary condition for the existence of a shadow is given by the presence of a line that separates the shadow pixels from the background pixels. Therefore, in case a hypothesized shadow is fully included in an object, the hypothesis is discarded.

3.2. Moving Cast Shadows Tracking

The second stage of the proposed algorithm aims at providing a coherent description of the segmented shadows over time. The goal is to track shadows from frame to frame, that is to establish a correspondence between instances of moving shadows over time. Tracking allows us to compute the life-span of each shadow. From each shadow's life-span and the relative position of objects and shadows, a reliability estimation is derived. This reliability estimation is used to validate or to discard each shadow detected in the first stage.

As commented in the introduction, shadow tracking is a difficult task. Traditional object tracking algorithms cannot be employed for shadow tracking. A shadow tracking algorithm has to be defined based on the limited amount of information available to describe a shadow and its evolution in time.

We assume that instances of the same shadow in consecutive frames overlap. This is a reasonable hypothesis for many video sequences. At each time instant, each extracted moving cast shadow is put in correspondence with previously extracted shadows. A correspondence between two shadows is established when the two corresponding regions overlap. Based on this rule, several cases may raise as shown in the example of Fig. 2. Let the frame at time k be composed of three extracted moving cast shadows, $S1, S2$, and $S3$. Once moving cast shadows are extracted at time $k+1$, their overlap with each shadow region at time k is computed. The possible configurations and the relative decisions are described below.

- If only one intersection between the regions at k and the one under analysis at $k + 1$ is found (as it is for the shadow $S2$ in the example), then the corresponding track in k is continued.

- If more than one intersection is found, that is more than one track at time k has an overlap with the region of S_{k+1} under analysis (as for shadows $S1$ and $S3$) then a conflict is encountered. The conflict is solved by updating the track that has the largest intersection with the considered region (track $S1$ in the example).

- If no overlap is found, a new track is initiated in $k + 1$ (as for shadow $S4$).

- If a splitting is encountered, that is when a shadow at time k is split in two disconnected regions at time $k + 1$, the two regions are considered as originated by the same track if both the regions have an overlap with the shadow region at time k.

- If a merging between two shadow regions is encountered, the track corresponding to the largest intersection with the merged region is continued.
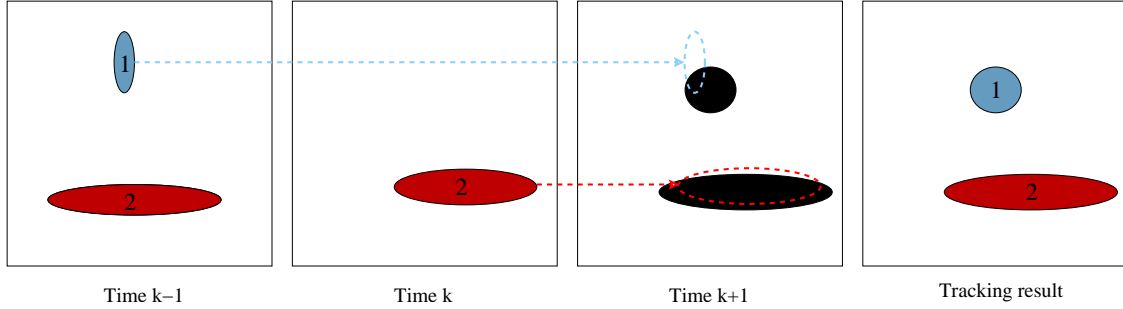
Figure 3. Tracking example

Given the mechanism described above, a *track duration* parameter is computed which defines the life-span of each shadow. The *track duration* is defined, at each frame, as the difference between the current frame number and the frame at which the track was initiated.

Shadows may be occluded by an obstacle for a certain duration. Furthermore, the algorithm for moving cast shadow extraction may fail to deliver stable results. In this second case, a previously extracted shadow is not detected in the current frame. This case is illustrated in the example reported in Fig. 3. Here, the same shadow $S1$ is present at time $k - 1$ and at time $k + 1$, but not at time k. In this case the above described mechanism would terminate a track at time k and initiate a new track at time $k + 1$. To avoid this kind of error and deal with the occlusion problem, a *track absence* parameter is defined. The *track absence* parameter is defined as the number of consecutive frames where the track is not associated with any of the extracted shadow regions. Since the reason for the disappearance of the shadow is unknown, the decision to terminate the corresponding track is delayed. If the track absence exceeds a certain period of time, *absence threshold t*, then the track is deleted. During the absence of corresponding shadows, the missing shadow region is assumed to maintain its area and position. An alternative approach would be to predict the position of the missing shadow region by projecting its area based on its motion. This solution would introduce errors in case of objects changing their direction. The overlap analysis is applied in this case as previously described. If a temporary occlusion (limited in time) causes the disappearance of the shadow and the track absence is less that the absence threshold, then the reappeared shadow is associated with the correct track. If the occluded shadow region changes its dynamic behaviour during the occlusion and no intersection is found when the shadow reappears, a new track is created when the shadow reappears. This error will have an influence in the computation of the *track duration* of the shadow under analysis.

During the tracking stage, the identity and durations of all target shadows are stored for each frame. Once the complete sequence has been analysed, the overall duration in the sequence is computed for each shadow. These parameters are processed off-line to improve the recognition results of the first stage of the algorithm. We observed, in fact, that short-lived shadows are very likely to be due to failure of the shadow recognition algorithm. Therefore, a temporal filtering of shadow recognition results is performed to eliminate all shadows whose duration is smaller than a certain threshold, τ. The value of τ is tuned according to the content of the sequence, as commented in the next section.

4. RESULTS

Similation results with various indoor and outdoor sequences have been obtained. The sequences *Hall Monitor* from the MPEG-4 test set, *Highway* from the MPEG-7 test set, *Group* from the European project *art.live*[31] test set, and the sequence *Laboratory* from ATON project[32] test set have been used. The reference image used in the analysis and described in Sec. 3.1 is the first frame of the sequence, acquired before the objects enter the field of view. For all the sequences, the same set of parameters has been used, with the only exception of τ. The threshold τ has been emprically determined as function of the level of activity. If objects are stationary for a long number of frames, the value of τ increases, if objects change their position rapidly or stay in the sequence

(a)   *(b)*

Figure 4. Shadow segmentation results for test sequence *Hall Monitor*. (a) Moving cast shadow extraction. (b) Final result after tracking.

for a short period of time, the value of τ decreases. The values used to obtain the results in this section are 38, 20, 4 and 5 for *Hall Monitor*, *Highway*, *Group* and *Laboratory*, respectively.

4.1. Moving cast shadow segmentation

Figures 4, 5, 6 and 7 illustrate the moving cast shadows extracted by the proposed method. The border of the shadow has been superimposed to the original image and the shadow region has been darkened to make it more visible. For each test sequence, the first column shows the results of the first stage of the shadow segmentation process described in Sec. 3.1. The second column shows the final results after the off-line processing performed thanks to the tracking stage, as described in Sec. 3.2.

Experimental results show that the proposed approach correctly identifies moving cast shadows, both in contact with objects and disconnected from them (for this second case see Fig. 4, third row). Two types of errors can be identified in the results of the first columns. Parts of moving objects are misclassified as moving shadows in portions of the object that are self shadowed and where the colour of the object is similar to that of the background. Moreover, portions of the objects that are very dark are erroneously detected as shadows. Examples of the first class of errors can be seen in Fig. 4, in Fig. 5, second row, Fig. 6, and in Fig. 7, second row. In Fig. 5, first and third row, and Fig. 7, third row, examples of the second type of errors are visible. These figures indicate difficult cases for the proposed approach. In the first case the misclassified object parts have characteristics that are similar to those of a shadow, while in the second case the use of photometric invariant features becomes unreliable. Photometric invariants are, in fact, unstable for R,G,B values that are close to the origin of the RGB cube.[29]

The results illustrated in the second columns of the figures show the improvements achieved thanks to tracking. The described failures of the moving cast shadow extraction stage have been eliminated by the analysis of the temporal behaviour of shadows.
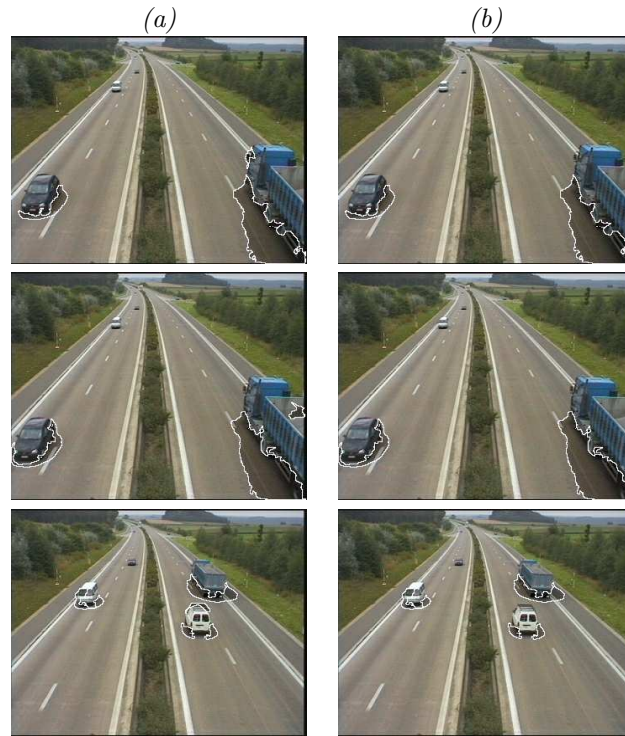
Figure 5. Shadow segmentation results for test sequence *Highway*. (a) Moving cast shadow extraction. (b) Final results after tracking.
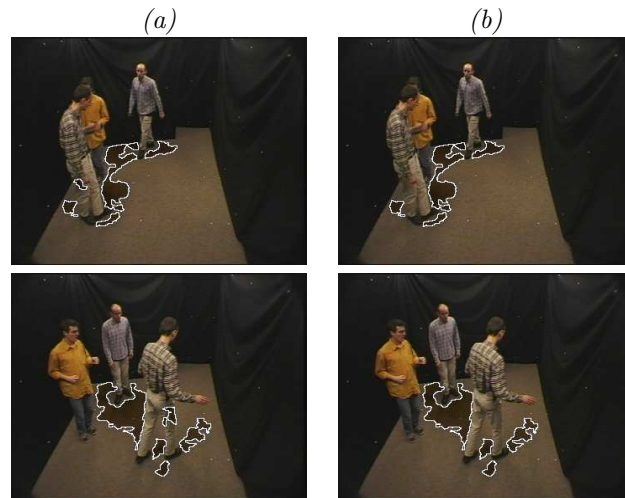

Figure 6. Shadow segmentation results for test sequence *Group*. (a) Moving cast shadow extraction. (b) Final results after tracking.

A failure of the tracking stage is shown in Fig. 8. The moving cast shadow extraction algorithm has correctly detected the shadows cast by the person on the background closed to his left hand and right foot. However, since the majority of edge pixels of these shadow regions are connected to the object and their temporal duration is very short, they have been removed by the off-line processing.
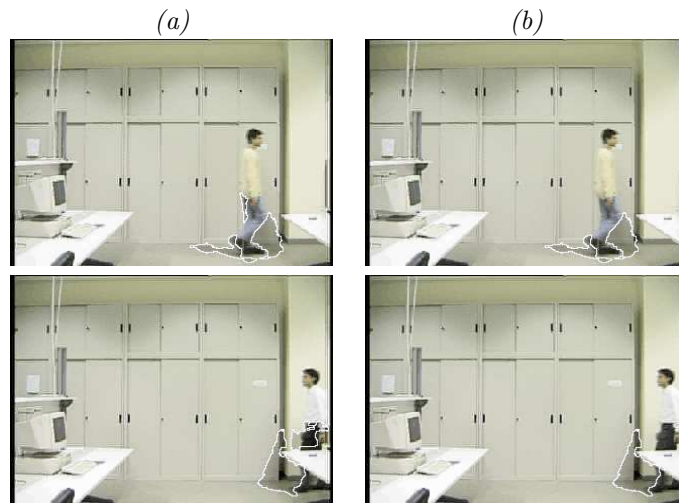
Figure 7. Shadow segmentation results for test sequence *Laboratory*. (a) Moving cast shadow extraction. (b) Final results after tracking.

Figure 8. Shadow segmentation results for test sequence *Laboratory*. (a) Moving cast shadow extraction. (b) Final results after tracking.

4.2. Application of moving cast shadow segmentation to the refinement of video object extraction

Since we are interested in shadow segmentation to improve video object extraction, we have compared the results of video object extraction obtained with the change detection method described in[28] to those obtained by removing from the object the cast shadow detected by the proposed algorithm.

In order to obtain an objective evaluation, we analysed the deviation of object masks with respect to a ground-truth segmentation. Two types of errors can be defined in each frame of the sequence n, namely false positives $\epsilon_p(n)$, and false negatives $\epsilon_n(n)$. False positives are pixels uncorrectly detected as belonging to the object mask, while false negatives are pixels belonging to the object but not detected. If $card(C(n))$ represents the number of pixels detected as object pixels at frame n, and $card(C_g(n))$ the number of pixels belonging to the ground-truth, then we compute the deviation from the reference segmentation as:

$$\epsilon(n) = \begin{cases} 0 & \text{if } Card(C(n)) = 0 \text{ and } Card(C_g(n)) = 0 \\ \frac{\epsilon_n(n) + \epsilon_p(n)}{Card(C(n)) + Card(C_g(n))} & \text{otherwise} \end{cases} \tag{4}$$

where $\epsilon(n)$ is in $[0,1]$. The spatial accuracy of the segmentation result is then quantified by:

$$\nu(n) = 1 - \epsilon(n) \tag{5}$$

that takes again values in $[0,1]$. If $\nu(n) = 1$ then there is a perfect match between segmentation results and ground-truth.

The results of application of moving cast shadow segmentation to the refinement of video object extraction are shown in Fig. 9. Objective as well as subjective comparison are presented.

The first row reports the objective evaluation of video object extraction on the test sequence *Hall Monitor*. The ground-truth segmentation used in the experiments is provided by the European project COST 211.[33] Here, we compare the following results. First, the spatial accuracy of results for the change detection. Second, the spatial accuracy of results for the change detection combined to the moving cast shadow extraction method of the first stage of the proposed algorithm. Third, the spatial accuracy obtained by using the tracking.

In the second row, object masks obtained with the three different methods for frame 125 are shown for subjective comparison. This demonstrates that the results of a change detection can be progressively improved by first extracting moving shadows and by then tracking them.

5. CONCLUSIONS

We proposed a shadow segmentation and tracking algorithm based on two stages. The first stage extracts moving cast shadows in each frame of a sequence. The second stage tracks the extracted shadows over time. The presence of a shadow is first hypothesized based on the assumption that shadows are darker than the corresponding area without shadow. The validity of detected regions as shadows is then verified by making use of the invariance of photometric invariant colour features. The value of a photometric colour invariant is not altered by the presence of a shadow. Next, geometrical information is considered. The boundaries of the detected regions and the position with respect to objects in the scene are analysed to check their consistency with the characteristics of shadow boundaries. Only the detected regions which satisfy all the previous hypotheses are considered as tentative shadow regions. The tentative shadow regions are finally tracked over time. The tracking stage evaluates the temporal behavior of a shadow and provides a reliability estimation of each shadow over time. This temporal reliability analysis eliminates the ambiguities still present after the shadow segementation stage thus improving the efficiency of the overall algorithm.

Several simulations have been carried out to evaluate the performance of the proposed shadow segmentation and tracking algorithm. Furthermore, the benefits introduced by the tracking stage have been quantified for indoor as well as for outdoor sequences. The obtained results validate the efficiency of the method for detection of moving shadows for a large class of scenes.

Future developments include the automatic computation of the threshold τ based on measurements of the level of activity in each sequence.
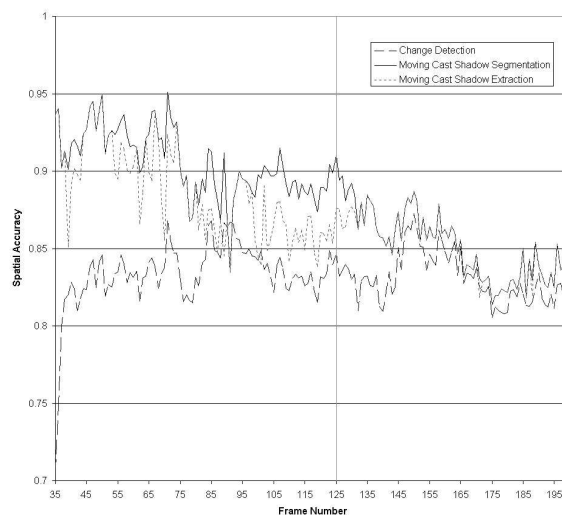
Figure 9. Comparison of video object extraction results on test sequence *Hall Monitor*. First row: objective spatial accuracy comparison among change detection results, change detection results combined to the proposed moving cast shadow extraction method, and to the complete moving cast shadow segmentation and tracking method. Second row: subjective object extraction comparison among the three methods for frame 125.

REFERENCES

1. M. H. Loew, L. Hambrick, and R. Carroll, "The entry-exit method of shadow-boundary segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* , September 1987.

2. A. Huertas and R. Nevatia, "Detecting building in aerial images," *Computer Vision Graphics Image Processing,* **41**, pp. 131–152, 1988.

3. R. Irvin and D. Mckeown, "Methods for exploiting the relationship between buildings and their shadows in aerial imagery," *IEEE Transactions on Systems, Man, Cybernetics,* **19**, pp. 1564–1575, 1989.

4. C. Wang, L. Huang, and A. Rosenfeld, "Detecting clouds and cloud shadows on aerial photographs," *Pattern Recognition Letters,* **12**, pp. 55–64, 1991.

5. Y. Liow and T. Pavlidis, "Use of shadows for extracting buildings in aerial images," *Computer Vision Graphics Image Processing,* **49**, pp. 242–277, 1991.

6. M. Bejanin, A. Huertas, G. Medioni, and R. Nevatia, "Model validation for change detection," in *2nd Int. IEEE Workshop on Applications of Computer Vision, Los Alamitos, USA*, pp. 160–167, December, 1994.

7. D. Koller, K. Danilidis, and H.-H. Nagel, "Model-based object tracking in monocular image sequences of road traffic scenes," *International Journal of Computer Vision,* **10**(3), pp. 257–281, 1993.

8. Y. Sonoda and T. Ogata, "Separation of moving objects and their shadows, and application to tracking of loci in the monitoring images," in *Proc. of IEEE Int. Conference on Signal Processing*, pp. 1216–1264, 1998.

9. K. Onoguchi, "Shadow elimination method for moving object detection," in *Proc. of IEEE International Conference on Pattern Recognition (ICPR)*, pp. 583–587, 1998.

10. A. P. Witkin, "Intensity-based edge classification," in *Proc. of National Conference on on Artificial Intelligence*, pp. 46–41, 1982.

11. J. M. Scanlan, D. M. Chabries, and R. Christiansen, "A shadow detection and removal algorithm for 2-d images," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 2057–2060, 1990.

12. M. Nagao, T. Matsutyama, and Y. Ikeda, "Region extraction and shape analysis in aerial photographs," *Computer Vision Graphics Image Processing,* **10**, pp. 195–223, 1979.

13. M. Adjouadj, "Image analysis of shadows, depressions, and upright objects in the interpretation of real world scenes," in *Proc. of IEEE Int. Conference on Pattern Recognition (ICPR)*, pp. 834–838, 1986.

14. J. Stauder, R. Melch, and J. Ostermann, "Detection of moving cast shadows for object segmentation," *IEEE Transactions on Multimedia,* **1**(1), pp. 65–77, 1999.

15. G. S. K. Fung, N. H. C. Yung, G. K. H. Pang, and A. H. S. Lai, "Effective moving cast shadows detection for monocular color image sequences," in *Proc. of 11th International Conference on Image Analysis and Processing (ICIAP)*, pp. 404–409, September, 2001.

16. R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting objects, shadows and ghosts in video streams by exploiting color and motion information," in *Proc. of 11th International Conference on Image Analysis and Processing (ICIAP)*, September, 2001.

17. O. Schreer, I. Feldmann, U. Goelz, and P. Kauff, "Fast and robust shadow detection in videoconference applications," in *Proc. of VIPromCom 2002, 4th EURASIP IEEE International Symposium on Video Processing and Multimedia Communications*, (Zadar, Croatia), June, 2002.

18. I. Mikic, P. C. Cosman, G. T. Kogut, and M. M. Trivedi, "Moving shadow and object detection in traffic scenes," in *Proc. of IEEE International Conference on Pattern Recognition (ICPR)*, (Barcelona, Spain), 2000.

19. T. Horprasert, D. Harwood, and L. S. Davis, "Statistical approach for real-time robust background subtraction and shadow detection," in *Proc. of IEEE International Conference on Computer Vision, FRAME-RATE Workshop*, (Kerkyra, Greece), September, 1999.

20. R. Gershon, A. D. Jepson, and J. K. Tsotsos, "Ambient illumination and the determination of material changes," *J. Opt. Soc. Am. A,* **3**(10), pp. 1700–1707, 1986.

21. K. Barnard and G. Finlayson, "Shadow identification using colour ratios," in *IS&T/SID 8th Color Imaging Conference: Color Science, Systems and Appl.*, pp. 97–101, 2000.

22. G. Finlayson, S. Hordley, and M. S. Drew, "Removing shadows from images," in *European Conference on Computer Vision*, (Cogenhagen), 2002.

23. S. Nadimi and B. Bhanu, "Moving shadow detection using a physics-based approach," in *Proc. of IEEE International Conference on Pattern Recognition*, **2**, pp. 701–704, 2002.

24. C. Jiang and M. O. Ward, "Shadow segmentation and classification in a constrained environment," *CVGIP: Image Understanding,* **59**(2), pp. 213–225, 1994.

25. G. Funka-Lea and R. Bajcsy, "Combining color and geometry for the active, visual recognition of shadows," in *Proc. of IEEE International Conference on Computer Vision*, pp. 203–209, 1995.

26. J. Pinel and H. Nicolas, "Estimation of 2d illuminant direction and shadow segmentation in natural video sequences," in *Proceedings of VLBV*, pp. 197–202, (Athens, Greece), October, 2001.

27. J. Stauder and R. Mech, "Detection and tracking of moving cast shadows," in *COST211 Workshop on Image Analysis for Multimedia Interactive Services*, (Berlin, Germany), May-June, 1999.

28. A. Cavallaro and T. Ebrahimi, "Video object extraction based on adaptive background and statistical change detection," in *Proc. of Visual Communications and Image Processing*, pp. 465–475, 2001.

29. T. Gevers and A. W. M. Smeulders, "Color-based object recognition," *Pattern Recognition,* **32**, pp. 453–464, 1999.

30. E. Salvador, A. Cavallaro, and T. Ebrahimi, "Shadow identification and classification using invariant color models," in *Proc. of IEEE Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2001.

31. http://www.tele.ucl.ac.be/PROJECTS/art.live/.

32. http://cvrr.ucsd.edu:88/aton/.

33. http://www.tele.ucl.ac.be/EXCHANGE/.