

# SHADOW-AWARE OBJECT-BASED VIDEO PROCESSING

Andrea Cavallaro\*\*, Elena Salvador\*, Touradj Ebrahimi\*

\*\* Queen Mary, University of London (United Kingdom)

\* Swiss Federal Institute of Technology (Switzerland)

## Abstract

Local illumination changes due to shadows often reduce the quality of object-based video composition and mislead object recognition. This problem makes shadow detection a desirable tool for a wide range of applications, such as video production and visual surveillance. In this paper, we present an algorithm for the isolation of video objects from the local illumination changes they generate in real world sequences when camera, illumination and scene's characteristics are not known. The algorithm combines a change detector and a shadow detector with a spatio-temporal verification stage. Colour information and spatio-temporal constraints are embedded to define the overall algorithm. Colour information is exploited in a selective way. First, relevant areas to analyse are identified in each image. Then, the colour components that carry most of the needed information are selected. Finally, spatial and temporal constraints are used to verify the results of the colour analysis. The proposed algorithm is demonstrated on indoor as well as outdoor video sequences. Moreover, performance comparisons show that the proposed algorithm outperforms state-of-the-art methods.

## 1 INTRODUCTION

Advances in hardware and digital video analysis and editing technologies are favouring the rapid development of applications such as video post-production, immersive gaming, realistic video conferencing, natural human-computer interfaces, home video and corporate communications [1]. The diffusion of digital video cameras and powerful personal computers is also driving the introduction of authoring tools to the home and corporate markets for the creation of new and richer content [2]. A way to create new and richer content is by extracting natural objects from a scene and then composing them in a new scene with artificial objects, such as for immersive gaming and rich media presentations, or with other objects captured by multiple remote sensors, such as for immersive video conferencing. To extract natural objects, professional studio productions can afford spaces with controlled lighting as well as dedicated

cameras equipped with depth sensors or with ring of LEDs coupled with a special retro-reflective cloth [3]. The need for this equipment represents a limitation for the introduction of authoring tools to the home market.

In order to make authoring techniques affordable for home production too, the set-up of the scene from which the objects are extracted should be simplified. The ideal solution is to use a digital video camera to extract characters without the need of ad-hoc scenes or ad-hoc cameras. This simplification leads to the problem of segmenting video objects without using a blue screen and without being affected by illumination changes due to shadows [4]. In this work, we focus our attention on this problem and in particular on the problem of extracting and processing moving shadows. We propose an algorithm for the separation of video objects and their shadows which is based on colour information and spatio-temporal verification. Colour information is exploited by means of the RGB colour space and by means of photometric invariant features. A spatio-temporal verification stage is then introduced to refine the results. Although the shadow segmentation problem has been increasingly studied in the past years ([5]–[19]), there exists no generally accepted method to detect shadows in image sequences. The conclusion of the review presented in [23] is that different approaches to shadow detection should be taken when addressing different kind of scenes. In addition to this, we observed that some methods use the same information in a contrasting way as cue to detect shadows. An example is given by the use of Saturation which is in some cases described as increasing in shadows while in other cases as decreasing. This contradiction demonstrates the difficulty of defining a general model describing the effect of a shadow. According to the taxonomy presented in [23], the proposed method belongs to the deterministic non-model based class, which includes the techniques presented in [5, 9, 10, 11, 12, 17]. The methods presented in [5, 9] use gray-level information only to detect moving shadows. Colour information for detecting moving shadows is considered in [10, 11, 12, 17]. The user-oriented HSV space is used in [11], whereas the hardware-oriented YCbCr and YUV spaces are used in [10] and [12], respectively. A spatial veri-

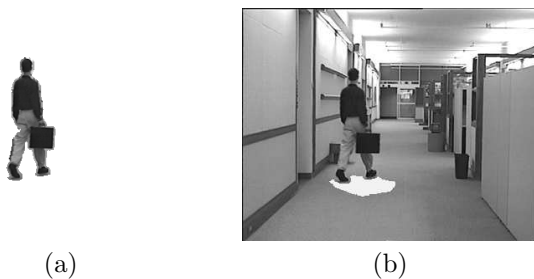
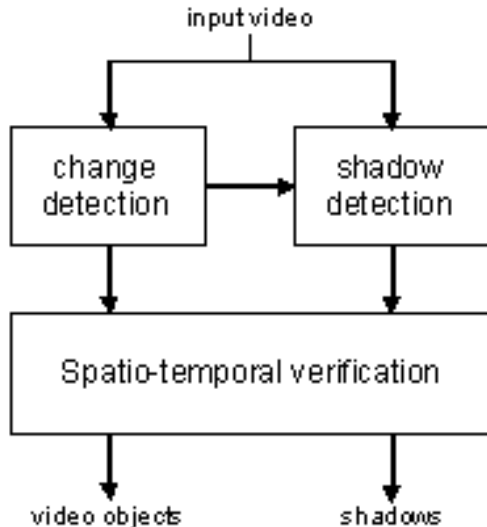


Figure 1: Block diagram of the proposed system for the separation of video objects (a) and their shadows (b)

fication process for refining object and shadow detection results is employed in [10]. The convex hull of the foreground edge points which are not candidate shadow points after colour analysis is used to extract the final object mask. This limits the performance of the method with non-convex objects, such as people. Finally, the dichromatic reflection model is adopted in [17] to model RGB changes in shadows and the body colour vector of candidate shadow regions is first estimated and then compared with the estimated body colour of the background to verify the results. This method requires a training phase to calculate the body colour of background surfaces that come under shadow in each different scene it is applied to. The method we present in this paper is designed to be able to work on different types of objects and when scene’s characteristics are not known. Its block diagram and an example of result are shown in Fig. 1.

The paper is organised as follows. Section 2 describes the change detection algorithm. The shadow detection algorithm is presented in Section 3. Next, Section 4 introduces the verification method. Experimental results and examples of applications are discussed in Section 5. Finally, in Section 6 we draw the conclusions and discuss future work.



Figure 2: Examples of change detection results. Objects as well as their shadows are segmented

## 2 CHANGE DETECTION

Change detection aims at detecting moving objects by deciding whether the foreground signal corresponding to an object is present after background subtraction, while discarding the effect of the camera noise. To discount the residual effect of the camera noise, the frame difference between the current image and the background image is binarized by thresholding. Early change detection techniques fix the threshold empirically and perform sufficiently well on sequences where moving objects are well contrasted from the background. However, thresholds have to be tuned manually according to the characteristics of the sequence and often need to be updated from one frame to another along the sequence itself. These major drawbacks limit this approach for fully automatic applications. To overcome these problems and to obtain a much more flexible procedure, we employ a locally adaptive threshold ([22]). This dynamic thresholding strategy models the noise statistics and applies a significance test. The noise model is defined based on the following assumptions: all pixels in the neighbourhood have changed only because of noise (hypothesis  $H_0$ ), and each frame of the sequence is affected by an additive Gaussian noise with a certain mean and variance. Under these hypotheses the noise model is described by a  $\chi^2$  distribution, whose properties depend on the number of pixels in the neighbourhood and on the variance of the Gaussian noise affecting each frame of the sequence. Given the  $\chi^2$  distribution and a significance level  $\alpha$ , the adaptive value of the threshold  $\tau_\alpha$  can be computed automatically. The significance level  $\alpha$  is a stable parameter that does not need manual tuning. An example of change detection result from the test sequence *Hall Monitor* is presented in Fig. 2. It is possible to notice that shadows are detected by the change detector as part of the objects. A further shadow detection stage is therefore needed to improve the spatial accuracy of the segmented objects.

## 3 SHADOW DETECTION

Shadows are difficult phenomena to model. A shadow does not have its own appearance, but that

of the material it is cast upon. Furthermore, a number of factors influence the appearance of a shadow. The shadow segmentation problem is generally faced by embedding multiple constraints when processing a video: the problem is to define the most appropriate constraints and how to embed them in the shadow detection algorithm.

### 3.1 EMPIRICAL ANALYSIS

Shadows cannot be defined by a specific colour appearance. However, it is possible to characterise a shadow by considering its effect on the colour appearance of the region on which it is cast. We want to exploit this property of shadows to derive a number of general rules that hold for a large number of test sequences. In particular the algorithm should be designed to work when camera, illumination and scene’s characteristics are unknown. To this end, we compare each frame of a video sequence to a reference frame in order to characterise the presence or absence of a shadow (Fig. 3). Comparing the incoming video frame with a reference frame is a widely used approach in the related literature. For the applications addressed in this paper, it is reasonable to assume that a reference image is available, either as a snapshot of the scene or as a model resulting from a learning process [22]. The advantage of using a reference image representing the background compared to the alternative of using the previous frame is that it is possible to avoid the dependence on objects speed. We also assume that the camera is static, but the same procedure can be applied to a moving camera sequence after global motion compensation [9].

The comparison between the current and the reference frame is based on colour information. Colour analysis is performed in order to identify those pixels in the image that respect chromatic properties of a shadow. The RGB colour space as well as photometric invariant features are considered in the analysis. Photometric invariant features are functions describing the colour configuration of each image coordinate discounting local illumination variations, such as shadings and shadows. Examples of photometric invariant features are Hue and Saturation in the HSV colour space and the normalised-RGB colour space. The normalised-RGB colour space, the *rgb* colour system, was chosen for its fast computation since it can be obtained by dividing the R, G and B coordinates by their total sum. The transformation from the RGB coordinates to the normalized colour space is given by

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B}, \quad b = \frac{B}{R+G+B}. \quad (1)$$

This transformation projects a colour vector in the RGB cube into a point on the unit plane described by  $r + g + b = 1$ . Two out of the three *rgb* variables suffice to identify the coordinates of the colour point

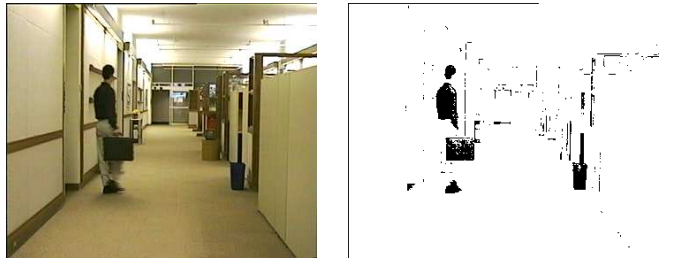


Figure 4: Pre-processing stage. Frame 37 of the test sequence *Hall Monitor* (left) and corresponding shadow selection map (right). Pixels discarded from the colour analysis stage are identified by the black label

in this plane.

The main rules derived from the empirical analysis are the following. First, a shadow darkens each colour component of the point on which it is cast. Second, the colour components do not change their order when a shadow occurs. Third, photometric invariant features do not change their value when an illumination change, such as a shadow, occurs; whereas they are likely to change their value in case of a material change. We use this information to define a strategy to detect shadows based on colour information. The details of the proposed method are described in the following sections.

### 3.2 PRE-PROCESSING

In order to reliably detect shadows, colour information is exploited in a selective way. First, the relevant parts to analyse are identified in each image. Then, the colour components that carry most of the needed information are selected. The relevant pixels that are suitable for the colour analysis are identified in a *selection map*. The computation of the map aims at eliminating the information which is useless or might mislead the colour analysis. In particular, the selection map identifies the pixels in achromatic parts of the scene. Typically a lower threshold is used to eliminate areas with low luminance values [11, 14]. Moreover, when considering *rgb* colour components in Eq. (1), the invariance is obtained at the cost of singularities and instabilities near the black vertex of the *RGB* cube. To avoid these instabilities, the volume close to the black vertex is not considered in the analysis process. The shadow selection map,  $m(x, y, t)$ , is a binary map indicating all the pixels whose colour components are smaller than the 20% of the colour dynamic range. Figure 4 shows an example of selection map.

The second stage of the pre-processing reduces the set of colour components to be considered in the subsequent colour analysis. The rationale behind this stage is that a colour component with a small

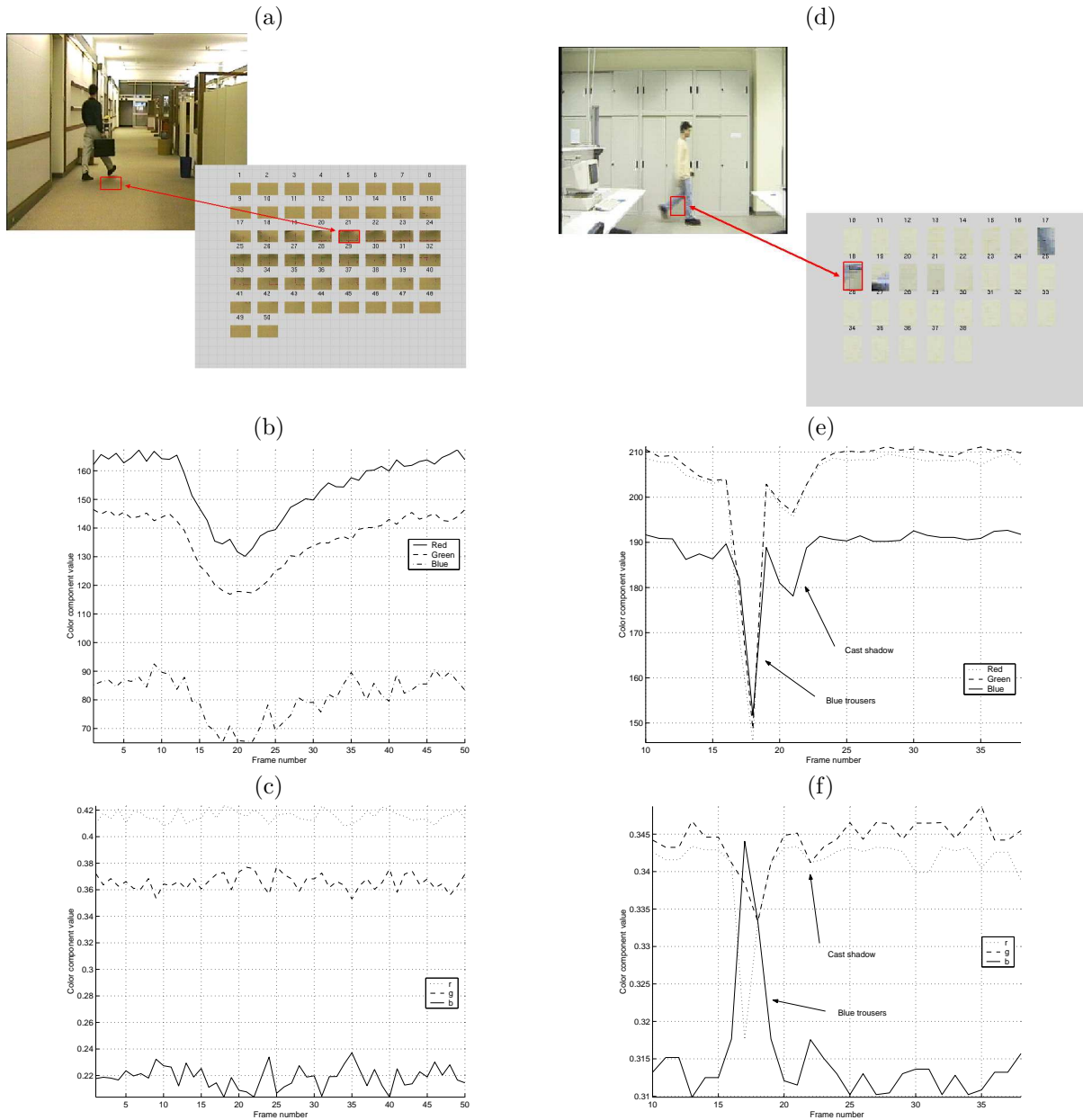


Figure 3: Comparison of the changes on the colour components of the RGB and  $rgb$  colour spaces at the passage of a shadow and an object. (a) Sample frame from the test sequence *Hall Monitor* and temporal region of interest (ROI) with a cast shadow. The passage of the cast shadow is highlighted for 50 frames. (b) Profile of the R, G, and B colour components in the selected ROI. (c) Profile of the  $r$ ,  $g$ , and  $b$  components in the same ROI. It is possible to notice the invariance properties of the  $rgb$  colour features. (d) Sample frame from the test sequence *Laboratory* and temporal region of interest (ROI) with the passage of an object and a cast shadow (highlighted for 29 frames). (e) Profile of the R, G, and B colour components in the selected ROI. The colour components decrease their value but do not change their order when a shadow occurs, whereas the colour components may change their order when an object passes. (f) Profile of the  $r$ ,  $g$ , and  $b$  components in the same ROI. It is possible to notice the different behaviour of the colour components in the presence of an object and in the presence of a shadow

value is highly influenced by the effects of reflections and ambient lighting and it is less discriminative for detecting shadows. This stage aims at choosing only those components that carry most of the information for the colour analysis. We observed that a shadow can be better characterised by analysing the behaviour of the colour components with the larger

values. This conclusion can guide the algorithm in eliminating from the analysis those colour features that do not add relevant information for the detection process. The smallest component is therefore not considered in the analysis.

Following the selection process in the pre-

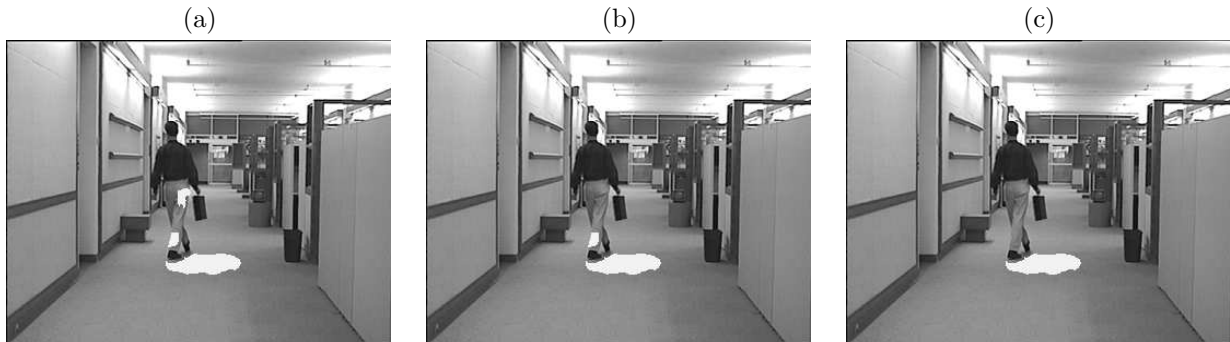


Figure 5: Shadow detection results with (a) colour analysis only; (b) colour analysis and spatial verification; (c) colour analysis and spatio-temporal verification

processing, each pixel  $(x, y, t)$  at time  $t$  which is not eliminated by the selection map  $m(x, y, t)$  is described by a feature vector  $\mathbf{f}(x, y, t)$  representing the value of the colour features  $f_i(x, y, t)$  that have been selected through the colour selection stage. The feature vector can be represented as

$$\mathbf{f}(x, y, t) = (f_1(x, y, t), f_2(x, y, t), f_3(x, y, t), f_4(x, y, t)), \quad (2)$$

where  $f_1(x, y, t)$  and  $f_2(x, y, t)$  are the selected components from the *RGB* space, and  $f_3(x, y, t)$  and  $f_4(x, y, t)$  are the selected components from the *rgb* space. On the other hand, the pixels eliminated from the colour analysis by the selection map are then treated separately in the post-processing.

### 3.3 COLOUR ANALYSIS AND POST-PROCESSING

The feature vector  $\mathbf{f}(x, y)$  selected in the pre-processing is analysed based on the results of the empirical analysis described in Section 3.1. The effect of a shadow is the darkening of each point on which it is cast. Let  $(x, y, t_R)$  be a background pixel and  $(x, y, t)$  the corresponding pixel in the image under test. The test in the RGB space is defined by

$$\begin{cases} f_1(x, y, t_R) > f_1(x, y, t) \\ f_2(x, y, t_R) > f_2(x, y, t) \end{cases} \quad (3)$$

This test is satisfied by a shadow, but also by an object which is darker than the corresponding background. To improve the selectivity, the colour information in the RGB space needs to be complemented by additional constraints. An additional constraint is based on exploiting the property of photometric invariant features. Photometric invariant features do not change their value when an illumination change occurs. On the other hand, they are likely to change their value in case of a material change. This constraint can be represented by the following condition

$$\begin{cases} f_3(x, y, t_R) \cong f_3(x, y, t) \\ f_4(x, y, t_R) \cong f_4(x, y, t) \end{cases} \quad (4)$$

The results of the analysis on  $f_1$  and  $f_2$ , and that on  $f_3$  and  $f_4$  are then fused to produce the shadow map. The fusion is the logical AND between the two partial results. The shadow map is then post-processed based on the selection map and on morphology in order to obtain spatially extended shadow areas (Fig. 5(a)).

## 4 VERIFICATION

Shadow detection based on colour analysis alone is not discriminative enough to allow for reliable segmentation, as shown in Fig. 5(a). The nature of the shadow detection problem provides us with additional information and therefore other constraints can be embedded in the algorithm based on contextual information. For this reason, after colour analysis, the shadow map undergoes a spatio-temporal verification process as described in the following.

### 4.1 SPATIAL VERIFICATION

Although the colour appearance of a shadow is undefined by nature, it is possible to define some constraints for its geometric appearance. The geometrical characteristics of shadows can be defined without any knowledge of the structure of the object or of the scene. In particular, the existence a line separating the shadow from the background is a necessary condition for the presence of a shadow. In order to check this condition, the relative position of the candidate shadow edge pixels is compared to the position of background and object pixels. Object pixels are pixels in the change detection mask that do not belong to the candidate shadow mask. This verification eliminates candidate shadows that are erroneously detected *inside* an object. In this case, shadow edge pixels are adjacent to object pixels *only* (Fig. 5(a), upper shadow). Shadows whose edge pixels are adjacent to object pixels as well as to background pixels are not eliminated (Fig. 5 5(b)). Further temporal processing is necessary to reduce this type of error.

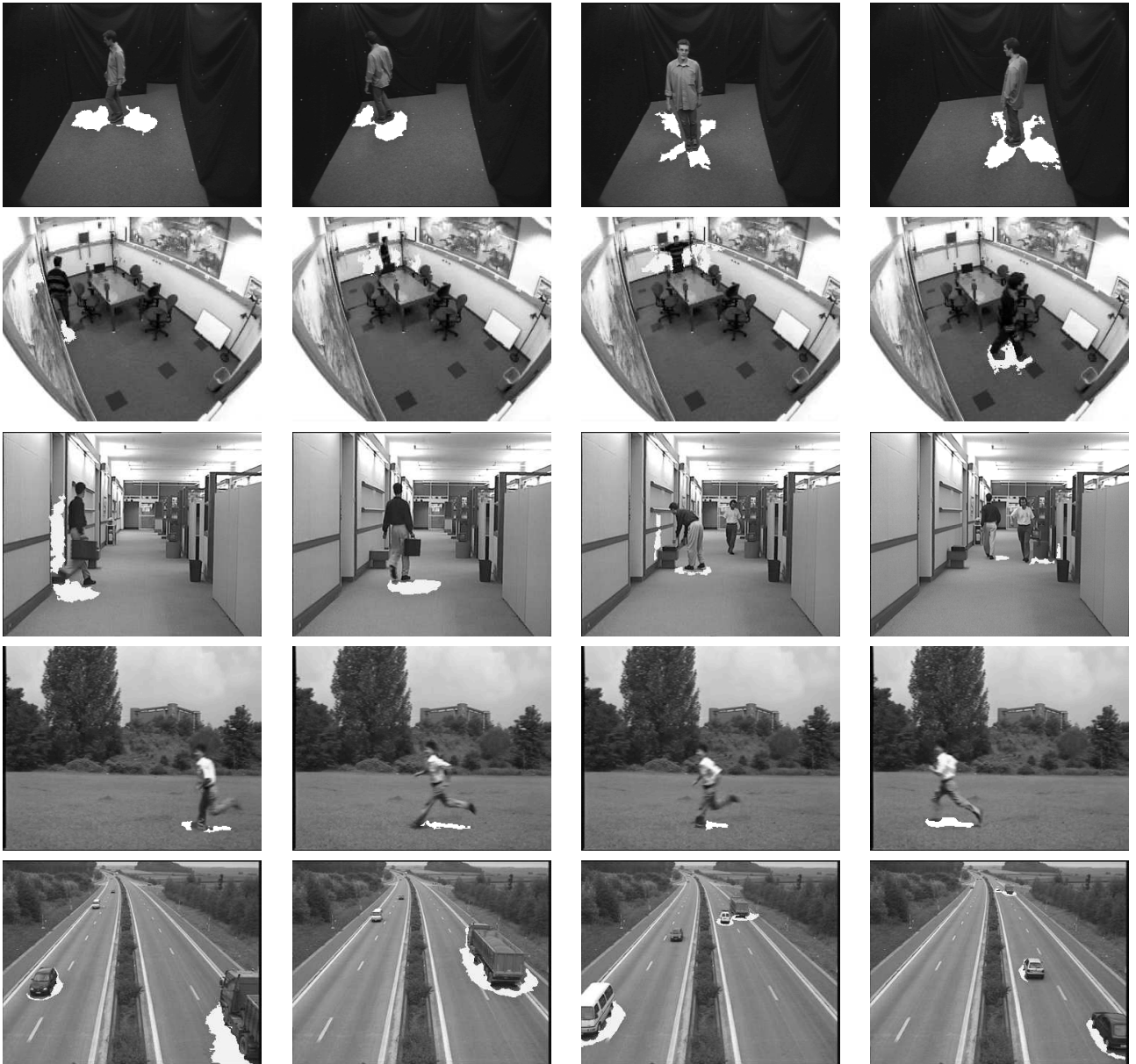


Figure 6: Shadow detection results for sample frames of the test sequence (from top to bottom) *Improvisation*, *Intelligent Room*, *Hall Monitor*, *Surveillance*, and *Highway*

## 4.2 TEMPORAL VERIFICATION

The final verification is based on the temporal consistency of shadows. Temporal verification 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. Shadow tracking is a difficult task because shadows are not characterised by their own appearance and their shape may change rapidly. For this reason, most of the 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 over time. To this end, we

employ a temporal filter based on Nearest Neighbour ([21]). Tracking allows us to compute the life-span of each shadow. From each shadow’s life-span, a heuristic rule is used to derive a temporal reliability estimation, which is used to validate or to discard each shadow detected in the previous stage. We observed in fact that short-lived shadows are highly likely to be due to a shadow detection error. This simple tracking mechanism allows one to remove shadows with a low temporal reliability, as shown in Fig. 5(c).

## 5 RESULTS

The results of the object and shadow detectors described in this paper as well as examples of applications of the proposed method are presented in this

section. Subjective and objective evaluation and comparison with state-of-the-art techniques are introduced in order to evaluate the performance by comparison with alternative methods. The results are evaluated subjectively by showing the detected shadows superimposed over the original image and colour-coded in white (see Fig. 1(b)). Furthermore, objective evaluation is performed with respect to a ground-truth segmentation by comparing the results of the change detector enhanced with the shadow detector (see Fig. 1(a)).

## 5.1 Shadow isolation

Test sequences from the MPEG-4 and MPEG-7 data set are used, as well as test sequences from the test set of the ATON project and the European project *art.live* (Figure 6). The sequences are in CIF format ( $288 \times 352$  pixels) and the frame rate is 25 Hz, unless otherwise stated in the remainder of the section. The sequence *Improvisation* (Figure 6, first row) is taken from the data set of the IST European project *art.live*. This scene represents a typical situation for video production where an actor moves in the scene generating multiple shadows. It is possible to notice that the shadow detector correctly identifies the shadows and the segmentation results are stable over time. We would like to highlight here that the detection is performed without any model of the scene, the illumination, or the captured object. A different scene set-up is shown for the test sequence *Intelligent Room* in the second row. The format of this sequence is  $320 \times 240$ . Here the scene is more complex compared to the previous sequence and the object casting shadows is smaller. However, shadows cast on the floor as well as on the walls are correctly detected. Similarly, the third row shows the shadow detection results for the MPEG-4 test sequence *Hall Monitor*. To demonstrate the performance of the proposed method in outdoor scenes, the test sequences *Surveillance* (fourth row, format  $352 \times 240$ ) and *Highway* (fifth row) are considered. The fourth row shows the shadow detection results in case of a deformable object, which can illustrate, for instance, a situation of outdoor video production. Finally, the MPEG-7 test sequence *Highway* illustrates how the proposed method can work for fast moving objects.

## 5.2 Performance comparison

The detection of shadows can be exploited to improve the performance of algorithm extracting video objects which are based on change detection. The objective evaluation is performed here by computing the deviation of the object segmentation with respect to a ground-truth segmentation. Defining the ground-truth segmentation of video objects is more reliable than defining the ground-truth segmentation of shadows. The generation of a ground-truth

for shadows is ambiguous: the outer boundary of a shadow occurs at points of infinitesimal decrease in the amount of illumination. As a result, the exact boundary of a shadow cannot be manually determined in a reliable way. In addition to this, a segmentation quality measure depends on the needs of the application in which the segmentation algorithm is used. For these reasons, we quantify the performance of shadow detectors when used in combination with a change detector.

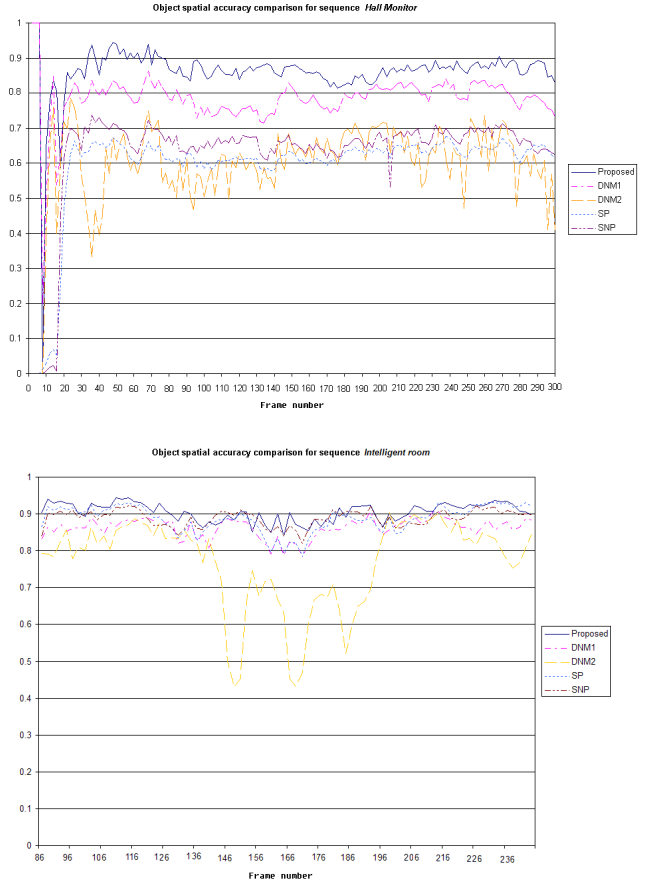


Figure 7: Comparison of the proposed algorithm with state-of-the-art methods. The plots represent the spatial accuracy  $\nu(t)$  of object segmentation. (Top) Test sequence *Hall Monitor*. (Bottom) Test sequence *Intelligent room*

Two types of errors can be defined in each frame of the sequence  $t$ , namely false positives  $\epsilon_p(t)$ , and false negatives  $\epsilon_n(t)$ . False positives are pixels incorrectly detected as belonging to the object mask. False negatives are pixels belonging to the object that are not detected. If  $\text{card}(C(t))$  represents the number of pixels detected as object pixels at frame  $t$ , and  $\text{card}(C_g(t))$  the number of pixels belonging to the ground-truth, then, when  $\text{Card}(C(t)) \neq 0 \wedge \text{Card}(C_g(t)) \neq 0$ , we compute the deviation from the reference segmentation as

$$\epsilon(t) = \frac{\epsilon_n(t) + \epsilon_p(t)}{\text{Card}(C(t)) + \text{Card}(C_g(t))}, \quad (5)$$

where  $\epsilon(t) \in [0, 1]$ . The spatial accuracy of the segmentation result,  $\nu(t)$ , is then

$$\nu(t) = 1 - \epsilon(t), \quad (6)$$

where  $\nu(t) \in [0, 1]$ . The larger  $\nu(t)$ , the higher the spatial accuracy. If  $\nu(t) = 1$ , then there is a perfect match between segmentation results and ground-truth. The results of the objective comparison for the test sequence *Hall Monitor* and *Intelligent room* are presented in Fig. 7. The symbols in the legend refer to the shadow detection technique used in the object extraction process: DNM1 [11], DNM2 [9], SP [13], SNP [14]. The mean values of accuracy corresponding to the plots in Fig. 7 are summarised in Table 1. It is possible to notice that the proposed combination of change and shadow detection (PROP) provides a more accurate segmentation than state-of-the-art methods.

Table 1: Mean object segmentation accuracy for the test sequences *Hall Monitor (HM)* and *Intelligent room (IR)*.

Sequence	DNM1	DNM2	SP	SNP	PROP
<i>HM</i>	0.78	0.60	0.59	0.63	0.86
<i>IR</i>	0.86	0.77	0.89	0.89	0.90

### 5.3 Example of applications

We conclude the results section with two examples of application of the proposed method in which it is possible to appreciate the benefits introduced by the shadow detector.

The first example of application of the proposed method is the detection of multiple simultaneous objects. Figure 8 shows a scene with three people walking in a room and casting several shadows which are caused by their interaction with multiple light sources. In this scene, a model based method for shadow recognition would fail due to the complexity of the scene. The proposed method is based on shadow properties and therefore can be applied to complex scenes, when shadows and object occlude each other. Cast shadows are typically attached to the shadow-casting object and cause segmentation errors by creating false adjacency between objects. An example of this problem is shown by two objects getting close to each other (Figure 8(a), first row) that are erroneously extracted as a single object when using the change detector *only*. When the shadow segmentation algorithm is applied (Figure 8(b)), the identification of shadow regions allows to solve the object segmentation problem and each person is assigned a separate bounding box (Figure 8(c)).

The second example of application of the proposed method is video composition without the use



Figure 9: Home set-up for video composition. Sample frame from the test sequence used for video composition (a) and background image (b)

of studio equipment. Figure 9(a) shows a sample frame from a test sequence recorded with a digital video camera in an ordinary room illuminated by a table lamp and by the light entering from the windows. The aim is to create a composite video emulating a weather forecast bulletin by placing the extracted object over a weather forecast map (Figure 9(b)). Figure 10 (top) shows some sample frames of the composite scene obtained without using shadow information. The absence of illumination effects due to shadows gives the impression of a flat 2D scene. The human brain does not in fact receive strong cues to infer depth information in the scene. Illumination effects have an important role in the perception of visual scenes and discarding them limits the visual quality of the scene composition due to a lack of naturalness. Shadows cast by objects on a background surface are informative about the shape of the object, the background and the spatial arrangement of the object relative to the background, especially when the shadow is in motion [24]. To improve the quality of the scene, shadows are explicitly segmented in the original scene and then rendered in the composite scene. This results in an augmented naturalness, as shown in Figure 10, bottom. The hand now appears clearly positioned in a 3D space and a realistic result is obtained, as if the hand had been filmed directly in front of the background image. The effect is more evident when the entire sequence is viewed since the object and shadow motion enhances the depth perception.

## 6 CONCLUSIONS

We addressed the problem of isolating video objects and their shadows by exploiting three sources of information, namely colour, spatial, and temporal constraints. The proposed algorithm for object and shadow segmentation is designed to work when the imaging conditions and the scene set-up are not under control. This algorithm does not require the knowledge of objects or scene models, nor requires external intervention. The selective use of colour, the use of photometric invariants, and the integra-



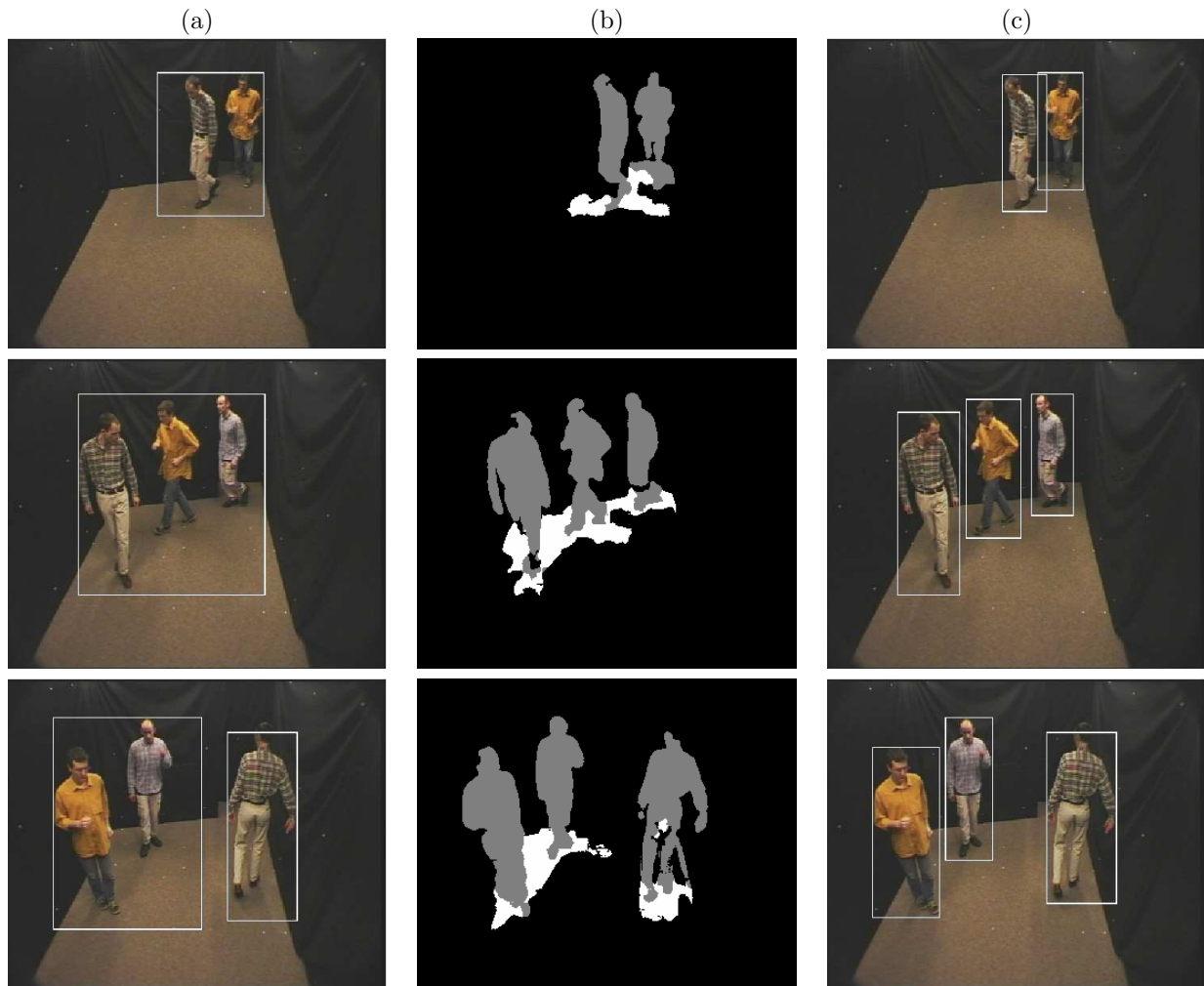


Figure 8: Isolating objects from their shadows enables the separation of multiple simultaneous objects. (a) Multiple moving objects, represented by their bounding boxes, are extracted as a single object by change detection in three sample frames of the test sequence *Group*. (b) Classification into object, shadow and background. Objects are displayed in grey, shadows in white. (c) Using the classification in (b), the segmentation errors are reduced and each object is represented by one bounding box

tion of spatial and temporal information allowed us to improve the performance of state-of-the-art methods.

Given the modularity of the proposed method, the different stages can be modified according to speed or to accuracy requirements. For example, colour analysis alone could be used for applications that do not require high accuracy, such as for the identification of the direction of the light source. Current work includes the optimisation of the code to reach real-time performance and the extension of the method to a multi-camera environment.

## References

- [1] A. Cane, "Lights. Action. Animate!," *Financial Times - IT review*, p. 1, 7 July 2004.
- [2] T. Glover, "Cut! Edit! Why digital video editing is the next big thing," *The Business*, **280**, pp. 1-6, July 2003.
- [3] O. Grau, "A Studio Production System For Dynamic 3D Content," in *Proc. of Visual Communications and Image Processing*, 2003.
- [4] X. Marichal and T. Umeda "Real-Time Segmentation of Video Objects for Mixed-Reality Interactive Applications," in *Proc. of Visual Communications and Image Processing*, 2003.
- [5] 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. Conf. on Signal Processing*, pp. 1216-1264, 1998.
- [6] K. Onoguchi, "Shadow elimination method for moving object detection," in *Proc. of IEEE Int. Conf. on Pattern Recognition (ICPR)*, pp. 583-587, 1998.
- [7] J.M. Scanlan *et al.*, "A shadow detection and removal algorithm for 2-d images," in *Proc. of IEEE*

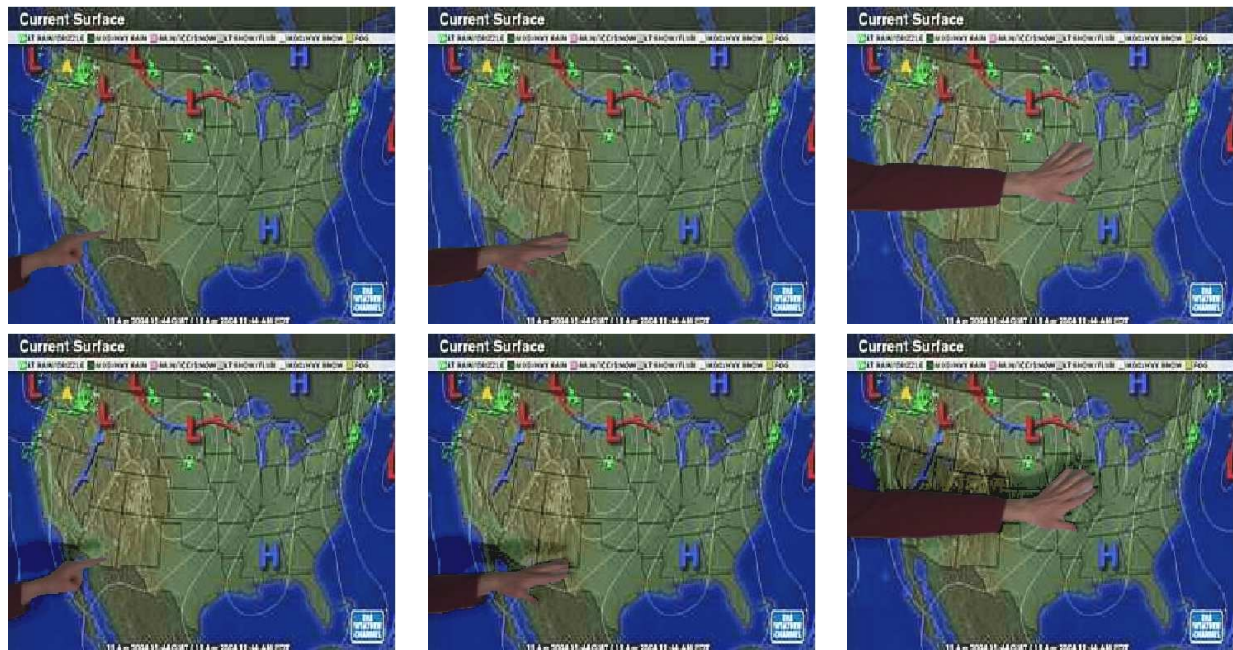


Figure 10: Scene composition without considering the shadow (top) and including the shadow (bottom)

- Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 2057–2060, 1990.
- [8] M. Adjouadj, “Image analysis of shadows, depressions, and upright objects in the interpretation of real world scenes,” in *Proc. of IEEE Int. Conf. on Pattern Recognition (ICPR)*, pp. 834–838, 1986.
- [9] J. Stauder, R. Melch, and J. Ostermann, “Detection of moving cast shadows for object segmentation,” *IEEE Trans. on Multimedia*, **1**(1), pp. 65–77, 1999.
- [10] G.S.K. Fung *et al.*, “Effective moving cast shadows detection for monocular colour image sequences,” in *Proc. of 11th Int. Conf. on Image Analysis and Processing (ICIAP)*, pp. 404–409, September, 2001.
- [11] R. Cucchiara *et al.*, “Detecting objects, shadows and ghosts in video streams by exploiting colour and motion information,” in *Proc. of 11th Int. Conf. on Image Analysis and Processing*, pp. 360–365, September 2001.
- [12] O. Schreer, I. Feldmann, U. Goelz, and P. Kauff, “Fast and robust shadow detection in videoconference applications,” in *Proc. of VIPromCom 2002*, pp. 371–375, June, 2002.
- [13] I. Mikic, P. C. Cosman, G. T. Kogut, and M. M. Trivedi, “Moving shadow and object detection in traffic scenes,” in *Proc. of IEEE Int. Conf. on Pattern Recognition (ICPR)*, pp. 321–324, 2000.
- [14] T. Horprasert, D. Harwood, and L. S. Davis, “Statistical approach for real-time robust background subtraction and shadow detection,” in *Proc. of IEEE Int. Conf. on Computer Vision*, Sept. 1999
- [15] K. Barnard and G. Finlayson, “Shadow identification using colour ratios,” in *IS&T/SID 8th Colour Imaging Conference: Colour Science, Systems and Appl.*, pp. 97–101, 2000.
- [16] G. Finlayson, S. Hordley, and M. S. Drew, “Removing shadows from images,” in *European Conf. on Computer Vision*, pp. 823–836, 2002.
- [17] S. Nadimi and B. Bhanu, “Moving shadow detection using a physics-based approach,” in *Proc. of IEEE Int. Conf. on Pattern Rec.*, **2**, pp. 701–704, 2002.
- [18] C. Jiang and M. O. Ward, “Shadow segmentation and classification in a constrained environment,” *CVGIP: Image Underst.*, **59**(2), pp. 213–225, 1994.
- [19] G. Funke-Lea and R. Bajcsy, “Combining colour and geometry for the active, visual recognition of shadows,” in *Proc. of IEEE Int. Conf. on Computer Vision*, pp. 203–209, 1995.
- [20] S. A. Shafer, “Using colour to separate reflection components,” *Colour Research Applications*, **10**(4), pp. 210–218, 1985.
- [21] E. Salvador, A. Cavallaro, T. Ebrahimi, “Spatio-temporal Shadow Segmentation and Tracking,” in *Proc. of Visual Comm. and Image Processing*, 2003.
- [22] A. Cavallaro and T. Ebrahimi, “Video object extraction based on adaptive background and statistical change detection,” in *Proc. of Visual Comm. and Image Processing*, pp. 465–475, 2001.
- [23] A. Prati *et al.*, “Detecting moving shadows: Algorithms and evaluation,” *IEEE Trans. on Pattern Analysis and Machine Int.*, **25**, pp. 918–923, 2003.
- [24] P. Mamassian, D.C. Knill, and D. Kersten, “The perception of cast shadows,” *Trends in Cognitive Sciences*, **2**, pp. 288–295, 1998.