

Chromatic Shadows Detection and Tracking for Moving Foreground Segmentation

Ivan Huerta^{a,b}, Michael B. Holte^c, Thomas B. Moeslund^c, Jordi González^d

^a*Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Parc Tecnològic de Barcelona, Llorens i Artigas 4-6, 08028 Barcelona, Spain (e-mail: ihuerta@iri.upc.edu).*

^b*DPDCE, University IUAV, Santa Croce 1957, 30135 Venice, Italy*

^c*Department of Architecture, Design and Media Technology, Aalborg University, 9220 Aalborg, Denmark (e-mail: mbh@create.aau.dk, tbm@create.aau.dk).*

^d*Computer Vision Center & Department of Computer Science (UAB), Edifici O, Campus UAB, 08193 Bellaterra, Spain (jordi.gonzalez@cvc.uab.cat).*

Abstract

Advanced segmentation techniques in the surveillance domain deal with shadows to avoid distortions when detecting moving objects. Most approaches for shadow detection are still typically restricted to penumbra shadows and cannot cope well with umbra shadows. Consequently, umbra shadow regions are usually detected as part of moving objects, thus affecting the performance of the final detection. In this paper we address the detection of both penumbra and umbra shadow regions. First, a novel bottom-up approach is presented based on gradient and colour models, which successfully discriminates between chromatic moving cast shadow regions and those regions detected as moving objects. In essence, those regions corresponding to potential shadows are detected based on edge partitioning and colour statistics. Subsequently (i) temporal similarities between textures and (ii) spatial similarities between chrominance angle and brightness distortions are analysed for each potential shadow region for detecting the umbra shadow regions. Our second contribution refines even further the segmentation results: a tracking-based top-down approach increases the performance of our bottom-up chromatic shadow detection algorithm by properly correcting non-detected shadows. To do so, a combination of motion filters in a data association framework exploits the temporal consistency between objects and shadows to increase the shadow detection rate. Experimental results exceed current state-of-the-art in shadow accuracy for multiple well-known surveillance image databases

which contain different shadowed materials and illumination conditions.

Keywords:

Detecting moving objects, Chromatic shadow detection, Temporal local gradient, Spatial and Temporal brightness and angle distortions, Shadow tracking.

1. Introduction

A fundamental problem for all automatic video surveillance systems is detecting objects of interest in a given scene [1]. A commonly used technique for segmentation of moving objects is background subtraction [2, 3]. Detection of moving regions (i.e., the foreground) is achieved by comparing the current image from a reference background image in a pixel-by-pixel manner. There have been many segmentation problems already tackled in the literature related to motion segmentation [4], such as bootstrapping [5, 6], changing background [7, 8, 9] and sudden illumination changes [10], to cite but a few, but one of the critical challenges is still shadow detection. Although this issue has been widely studied in [11, 12, 13], shadow segmentation is still far from being solved. The focus of this paper is to cope with the shadows problem.

Shadows can be divided into two categories: *static* and *dynamic (moving)* shadows. Static shadows occur due to static background objects (e.g., trees, buildings, parked cars, etc.) blocking the illumination from a light source. Static shadows can be successfully incorporated into the background model thus being properly detected. However, the impact of dynamic shadows is critical for foreground segmentation, since objects can be merged or hidden by other objects, and both their size and shape can be distorted. This results in a reduction of the performance of foreground detection approaches applied in scene monitoring, object recognition, target tracking and people counting.

Dynamic shadows are more problematic, since they are due to the moving objects (e.g., people, vehicles, etc.). Dynamic shadows can take any size and shape, and can be *penumbra* (soft) or *umbra* (hard) shadows. Penumbra shadows exhibit low values of intensity but similar chromaticity values w.r.t. the background. Umbra shadows can exhibit different chromaticity than the background, and their intensity values can be similar to those of any new object appearing in a scene. When the chromaticity of umbra shadows differs enough from the chromaticity of the global scene illumination, we define this

as *chromatic shadow*. Consequently, umbra shadows are significantly more difficult to detect, and therefore usually detected as a part of moving objects by current state-of-the-art approaches.

This paper presents an approach which successfully detects umbra and penumbra shadows. First a bottom-up approach for detection and removal of chromatic moving shadows in surveillance scenarios is presented based on our previous work [14]. We apply a multi-stage approach combining multiple cues, namely colour, gradient information, and shadow statistics. Secondly, a top-down architecture based on a tracking system is proposed to enhance the chromatic shadow detection presented in [14]. This step is required since shadows can be lost for a number of frames due to camouflage. In these cases the use of motion filters allows our proposed system to track shadows, thus improving the accuracy and robustness of the final foreground detection performance. Experimental results show that applying the top-down shadow tracking, the shadow detection rate is improved by approximately 13%.

Secondly, a top-down architecture based on a tracking system is proposed to enhance the chromatic shadow detection, where motion filters are used for tracking. This step is required since shadows can be lost for a number of frames due to camouflage, so in these cases the use of basic motion filters allows our proposed system to track shadows thus improving the accuracy and robustness of the final foreground detection performance.

The remainder of the paper is organised as follows. Next section reviews the field in shadow detection and tracking, along with our contributions to this subject. In section 3, the theoretical concept of our approach is outlined. The algorithm for foreground segmentation, along with the detection and removal of chromatic moving shadows are described in section 4. The top-down process used to enhance the shadow detection is described in section 5. Finally, we present experimental results in section 6 and concluding remarks in section 7.

2. Related Methodology

Shadow detection is a major field of research within computer vision. Even though many algorithms have been proposed [11, 12, 13], the problem of detection and removal of shadows in complex environments is still far from being completely solved. A common direction in the research is to assume that shadows decrease the luminance of an image, while the chrominance stays relatively unchanged [15, 16]. However, this is not the case in

many scenarios, e.g., in outdoor scenes. Other approaches apply geometrical information. Onoguchi [17] uses two cameras to eliminate the shadows of pedestrians based on object height, where objects and shadows must be visible to both cameras. Ivanov et al. [18] apply a disparity model, which is invariant to arbitrarily rapid changes in illumination, for modelling the background. However, to overcome rapid changes in illumination at least three cameras are required. In [19], Salvador et al. exploit the fact that a shadow darkens the surfaces on which it is cast, to identify an initial set of shadowed pixels. This set is then pruned by using colour invariance and geometric properties of shadows. Hsieh et al. [20] first separate the objects of interest and assume that the objects and their shadow have different orientation. Then, several features like orientation, mean intensity and center position of a shadow region are used to parametrize a shadow model. It should be noted that most of the approaches which apply geometrical information requires shadows to be cast on a flat plane, and give strong assumptions that need to be fulfilled.

Another popular approach is to exploit colour differences between shadow and background in different colour spaces. In [21], Cucchiara et al. consider the hypothesis that shadows reduce surface brightness and saturation while maintaining the hue properties in the HSV colour space. Liu et al. [22] propose another approach working in the HSV colour space, which combines local and global features for shadow removal. Schreer et al. [23, 24] adopt the YUV colour space, while Horprasert et al. [25], Kim et al. [16] and [26] build a model in the RGB colour space to express normalised luminance variation and chromaticity distortions. However, these methods require illumination sources to be white, and assume shadow and non-shadow have similar chrominance.

Several authors use textures to obtain a segmentation without shadows. The idea is that the structure of the texture/gradients/edges of regions lit by shadow are unchanged. Leone et al. [27] propose a texture-based approach using a preliminary procedure in order to evaluate the photometric information for all pixels marked as foreground. This process shows how much darker the segmented regions are with respect to the background model. Next, texture analysis is performed by projecting the neighbourhood of pixels onto a set of Gabor functions. Another algorithm for detection of moving cast shadows, based on a local texture descriptor called Scale Invariant Local Ternary Pattern (SILTP), is presented by Qin et al. [28]. Zhang et al. [29] use ratio edges to detect and locate where the shadows are, and in Chen et al. [30]

HOG descriptors of shadows are learned using SVMs to locate shadow regions. Sanin et al. [31] propose a method similar to Huerta et al. [14] where gradient magnitude and gradient orientation are used to detect shadows, but based on the gradient direction correlation. However, this paper [31] does not take into account cases where there are no texture/gradient in the background model. Whereas, Heikkila et al. [32] apply Local Binary Patterns. It still fails to detect umbra shadows.

To overcome these shortcomings, a number of approaches apply colour constancy methods, combine different techniques or use multi-stage approaches. In addition to scene brightness properties, Stauder et al. [33] extract edge width information to differentiate penumbra regions from the background. In [34], Finlayson et al. use shadow edges along with illuminant invariant (intrinsic) images to recover full colour shadow-free images, and in [35] the authors propose an entropy-based approach. Even so, a part of the colour information is lost in removing the effect of the scene illumination at each pixel in the image. Weiss [36] computes the reflectance edges of the scene to obtain an intrinsic image without shadows. However, this approach requires significant changes in the scene, and the reflectance image also contains the scene illumination. Huang et al. [37] adopt a bi-illuminant model and apply a Gaussian Mixtures Model (GMM) and confidence-rated learning, while Martel et al. introduce a parametric approach based on Gaussian mixtures (GMSM) [38]. Additionally, they propose a nonparametric framework based on the physical properties of light sources and surfaces, and apply spatial gradient information to reinforce the learning of the model parameters [39]. Nadimi et al. [40] propose a multi-stage approach for outdoor scenes, which is based on a spatio-temporal albedo test and dichromatic reflection model. Finally, a number of authors have introduced methods for shadow removal using only one single image [41, 42, 43, 44, 45, 46], however, the focus of this work is shadow detection and removal in video sequences. Comparative studies of shadow detection techniques can be found in [11, 47].

Exploiting temporal information for shadow detection has been rarely used. Liu et al. [22] use temporal information in order to avoid misclassifying the object under segmentation as shadow. To solve this, a nearest neighbour match method is used to track the object by checking if it is a foreground object in the previous frames. However, the authors only use tracking to recover miss-detected objects and thus lose the shadow information. In our case, we take advantage of this information to track shadows, resulting in both improved object and shadow detection. In the previous

papers, when a shadow has successfully been detected it is usually removed instantly, since only the object is of interest for further processing and not the shadow. As a result, the shadow information is lost. Our idea is to use this information to improve other aspects of object and shadow detection and tracking. Specifically, if a detected shadow is tracked over time instead of being discarded, it could be used to improve the shadow detection and possibly the object detection and tracking as well.

Summarizing the contributions of this paper compared with the reviewed literature: (i) a shadow detection technique is designed to cope with chromatic moving cast shadows, by grouping potential shadow regions and considering the bluish effect, edge partitioning, temporal similarities between local gradient, and spatial similarities between chrominance angle and brightness distortions; (ii) both objects and shadows are tracked, thereby establishing data association between them. We use the mutual information and data association of both object and its shadow for improving the robustness of the detection by recovering misdetected shadows due to shadow camouflages. Consequently, foreground detection does not take into account neither penumbra and umbra shadows, thus improving current state-of-the-art performance without any assumption about camera location, surface geometries, surface textures, shapes and types of shadows, objects and background properties.

3. Analysis of Shadow Properties

The colour information ρ at a given pixel a obtained from a recording camera supposing Lambertian surfaces depends on four components: the Spectral Power Distribution (SPD) of the illuminant denoted $E(\lambda)$, the surface reflectance $R(\lambda)$, the sensor spectral sensitivity $Q(\lambda)$ evaluated at each pixel a and a shading factor σ .

$$\rho_a = \sigma \int E(\lambda)R(\lambda)Q_a(\lambda)d\lambda \quad (1)$$

The surface reflectance $R(\lambda)$ depends on the material, i.e., materials have different response to the same illumination.

3.1. The bluish effect

In outdoor scenes, the environment is mainly illuminated by two light sources: a point light source (the sun) and a diffuse source (the sky) with

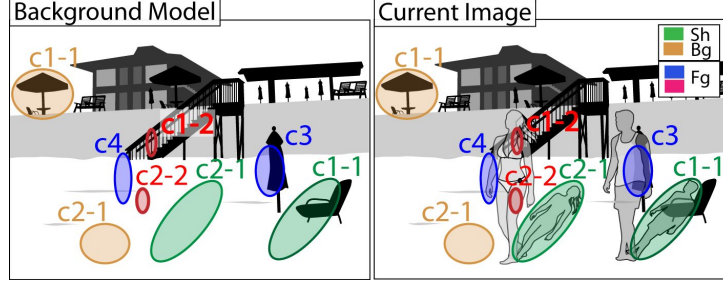


Figure 1: A sketch of the four main cases (c1 to c4: blue ellipses) and two anomalies (c1-2 and c2-2: red ellipses) that can occur, when performing foreground segmentation with the influence of shadows, using the temporal local gradients. The ellipses represent detection of potential chromatic shadows. They are grouped by considering an intensity reduction, the bluish effect and an edge partition.

different SPD $E(\lambda)$. Besides a reduction in the intensity, an outdoor cast shadow results in a change of the chrominance. The illumination of the sky has higher power components in the lower wavelengths λ (450 - 495 nm) of the visible spectrum, and it is therefore assumed bluish as argued in [40]. When the direct illumination of the sun is blocked and a region is only illuminated by the diffuse ambient light of the sky, materials appears to be more bluish.

3.2. Temporal local gradient information

By applying gradient information we can obtain knowledge about object boundaries, and thereby improve the foreground segmentation. Additionally, the gradient provides textural information about both the background and foreground image. Although shadows result in a intensity reduction of the illumination, and the texture of a given object or the background has lower gradient magnitude, the structure remains the same, i.e., the gradient orientation is unchanged.

3.3. Shadow scenario and solutions

When performing foreground segmentation with the influence of shadows, and taking the temporal local gradients into account, four main cases can occur as illustrated in Fig. 1. The ellipses represent detection of potential chromatic shadows. They are grouped by considering an intensity reduction, the bluish effect and an edge partition.

Case1: Similar local gradient structures are present in the background model and in the current image. By examining similarities between

the local gradients, and the fact that there is no foreground object in the current image, potential shadows can be detected and identified as shadow regions (*case1-1*). However, if a foreground object is present, it can be misclassified as shadow if the gradients of the background and the foreground region considered as possible shadow are similar (*case1-2*).

Case2: There are no available local gradients in the background model nor in the current image. Since, the change in illumination of all the potential shadow regions has to be similar, temporal and spatial similarities between chrominance angle and brightness distortions within the potential regions are analysed to detect chromatic shadows (*case2-1*). However, a foreground object can be misclassified as shadow if the foreground object has no gradients, and the chrominance angle distortion is similar among the pixels in the region of the object (*case2-2*).

Case3: Local gradient structure is present in the background model but not in the current image. If there is no gradient in the current image but in the Bg. model there is, there must be a new foreground object in the current image. So, by examining similarities between temporal gradients, the potential shadow will be considered a foreground object.

Case4: Local gradient structure is present in the current image but not in the background model. Therefore, there must be a new foreground object in the current image. In this case, the gradients in the current image are employed for object detection. Hence, there is no need to analyse the potential region further.

The described characteristics are not sufficient to address the anomalies in *case1-2* and *case2-2*. Therefore, we make further assumptions and apply some additional steps, which are explained next.

4. Bottom-Up Chromatic Shadow Detection

Our approach, depicted in Fig. 2 is a multi-stage approach. The first three stages (4.1, 4.2 and 4.3) remove the pixels which cannot be shadow. The fourth step (4.4) divides the regions of potential shadows. Chromatic shadow detection is performed in stage 5 (4.5) and 6 (4.6) based on gradients and chrominance angles, respectively. In stage 7 (4.7) the edges belonging to

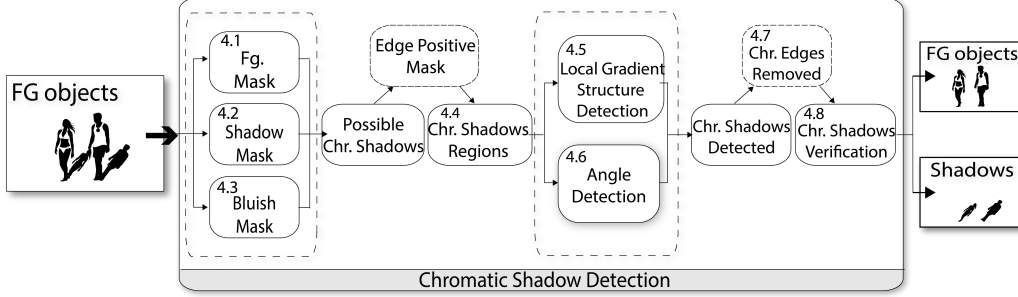


Figure 2: A schematic overview of the chromatic shadow detection approach.

the shadows deleted in stage 4 (4.4) are considered shadows again. The last step (4.8) avoids foreground regions being erroneously detected as chromatic shadows. An example is given in Fig. 3.

4.1. Moving foreground segmentation

In this stage, foreground objects, shadows and some erroneous pixels are segmented. In order to achieve moving foreground segmentation, a hybrid approach presented in [48], which fuses invariant colour and gradient models, is used (please refer to [48] for further details). This approach can cope with several motion segmentation challenges, e.g., illumination changes, achromatic shadows, since it is based on a chromatic colour model [25]. Note that any other algorithm can also be used to generate this mask [4, 5, 6, 7, 8, 9, 10], however, the method proposed in [48] has been selected, since it is able to avoid achromatic shadows by fusion of different cues (chromatic, colour cone and gradient models). Additionally, the colour and gradient models presented in [48] are used for the next steps of our chromatic shadow detection, which requires the use of the gradient model in order to create the edge masks.

After this initial detection, moving foreground objects, chromatic shadows and some isolated pixels are represented by a binary mask named $M1$. Furthermore, a mask is created using the gradient model in [48] and divided into three masks ($Edneg$, $Edpos$, and $Edcom$), which are used for the next steps. The $Edneg$ mask contains the edges from the background model which are occluded by the foreground object. The $Edpos$ mask contains the edges of the foreground object for the current image, which are not in the background model. The $Edcom$ mask contains the edges that are shared between the foreground object and the edges occluded in the background model for the current image.

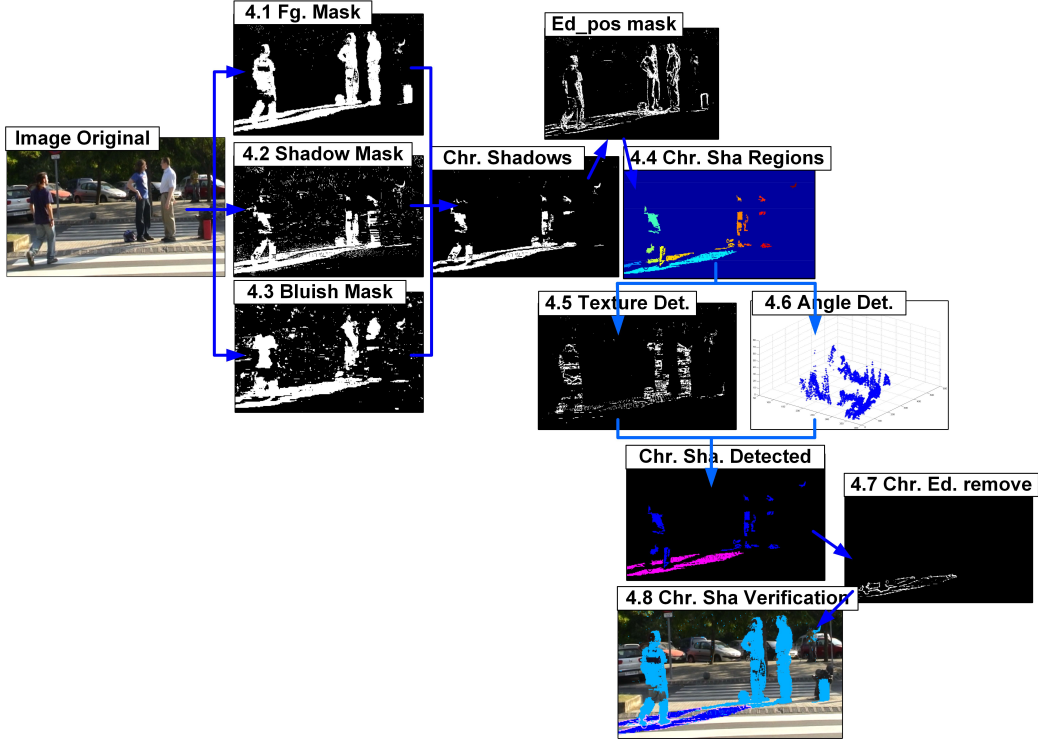


Figure 3: An example of chromatic shadow detection. The numbers added to the image captions correspond to the respective sub-sections in section 4.

4.2. Shadow intensity reduction

In this step the $M1$ mask is reduced to avoid pixels which cannot be shadows. A foreground pixel cannot be a shadowed pixel if it has a higher intensity than the background model. Hence, a new mask $M2$ is created according to equation 2.

$$M2_{a,t} = M1_{a,t} \wedge (I_{a,t}^R < \mu_a^R) \wedge (I_{a,t}^G < \mu_a^G) \wedge (I_{a,t}^B < \mu_a^B) \quad (2)$$

where I corresponds to the image, a to a pixel position, and μ is the median value for the RGB pixel of the background model.

4.3. The bluish effect

The effect of illuminants, which are different than white lights, provokes chromaticity changes, since the intensity varies differently for each color channel. In outdoor sequences the main illuminants are the sky and the sun (neither of them white illuminant). The sky is the only source of illumination on

shadowed regions, and it is assumed to be bluish, as argued in [40]. Therefore, the intensity changes in the red and green channels are larger than in the blue channel. This knowledge is used to reduce the potential shadow region detected in the previous step:

$$M3_{a,t} = M2_{a,t} \wedge (I_{a,t}^R - \mu_a^R) > (I_{a,t}^B - \mu_a^B) \wedge (I_{a,t}^G - \mu_a^G) > (I_{a,t}^B - \mu_a^B) \quad (3)$$

where a corresponds to a pixel position. Obviously, the bluish effect cannot be applied for indoor sequences, however the entire approach can be applied without this mask for indoor sequences as it has shown in the experimental results with the Hallway sequence.

4.4. Potential chromatic shadow regions

It is assumed that shadow regions have similar intensity changes for each channel, since the illuminant is the same. However, different surfaces have different reflectance characteristics, hence, the intensity change depends on the surface material. Therefore, we apply edges to describe region borders. Specifically, we build a new mask $M4$ using the foreground edges detected in the current image ($Edpos$) to separate the potential shadow regions from the moving foreground objects:

$$M4_{a,t} = M3_{a,t} \wedge \neg Edpos_{a,t} \quad (4)$$

The \neg symbol before $Edpos$ in equation 4 means negation. $M3$ contains the foreground blobs after being corrected by the shadow intensity and the bluish effect, and $Edpos$ contains only the edges of foreground objects in the current image, not the edges of the background model. Since shadow regions do not contain any edges which are not in the background model, we use the $Edpos$ mask to divide the foreground blobs into smaller regions, and to evaluate if these regions are shadows or not. Consequently, based on $Edpos$, the holes in $M3$ will not correspond to shadows. A minimum area morphology is applied to avoid smaller regions, which do not contain enough information for the subsequent steps of the shadow analysis.

4.5. Chromatic shadow gradient detection

Next, the temporal gradients of the regions in $M4$ are analysed to identify which case of the theoretical shadow analysis (see section 3) each of the regions complies with. A region will be considered a shadow if it complies with case 1. The negative foreground edges ($Edneg$) of the region are compared

to the common foreground edges ($Edcom$) to check if the region is a shadow and avoid the anomaly case 1-2.

$$Tx_b = \begin{cases} 1 & \text{if } \left(\frac{\sum_{a \in R_b} (R_{b,a} \wedge Edneg_a)}{|R_b \wedge Edtot|} \cdot k_n < \frac{\sum_{a \in R_b} (R_{b,a} \wedge Edcom_a)}{|R_b \wedge Edtot|} \right) \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where a is the pixel position; R_b is the evaluated region and b is the number of the region; $|R_b|$ denotes the number of pixels of region b ; $|R_b \wedge Edtot|$ denotes the number of pixels representing the edges detected in the background model and the current image; k_n corresponds to a confidence region, set at 1.5 for all the different sequences in the experimental results, which is equal to the probability that the region belongs to a shadow or a foreground object based on the local gradients of the image. Note that Tx_b is not a mask, but a boolean expression that will be 1 if the region R_b is a shadow, and 0 otherwise. If the region is not a shadow, the evaluation will be performed by the chromatic shadow angle and brightness detection in the next stage (4.6).

4.6. Chromatic shadow angle and brightness detection

In this step temporal and spatial similarities of the chrominance angle and brightness distortion for all pixels belonging to regions which have so far not been classified as shadow, are analysed. A region will be considered a shadow if it complies with case 2. The only regions analysed in this stage are those without gradients, neither in the background model nor in the current image. If the pixels do not have a significant gradient, but have similar chrominance angle distortion and similar brightness distortion, the region is classified as shadow.

$$ABd_b = \begin{cases} 1 & \text{if } \left(\left(\left(\frac{\sum_{a \in R_b} (R_{b,a} \wedge Edneg_a)}{|R_b \wedge Edtot|} \wedge \frac{\sum_{a \in R_b} (R_{b,a} \wedge Edcom_a)}{|R_b \wedge Edtot|} \right) = 0 \right) \right. \\ & \vee \left(\frac{\sum_{a \in R_b} (R_{b,a} \wedge Edtot_a)}{|R_b|} < k_t \right) \\ & \wedge (\sigma(R_b \wedge \check{\alpha}) < k_a) \wedge (\sigma(R_b \wedge \check{\beta}) < k_b) \\ 0 & \text{otherwise} \end{cases}$$

where σ is the standard deviation of $\check{\alpha}$ and $\check{\beta}$ which are the chrominance angle and brightness normalised distortions calculated for each pixel in the region R_b , respectively; k_t is a confidence region to avoid noise gradients; k_a and k_b are minimum thresholds used to determine if the angle and brightness distortion are similar among the pixels of the evaluated region. From the experiments of different sequences, a stable detection has been achieved using the following range of values: $k_a = [3-7]$, $k_b=35$, and $k_t = 0.1$ so if more than 10% of the pixels are considered edges then it cannot be noise.

4.7. Chromatic shadow edge removal

Pixels of the potential shadow regions, which were neglected in section 4.4, since they were part of the *Edpos* mask, are included again in the new set of shadow regions.

4.8. Shadow position verification

A moving cast shadow is always caused by a moving foreground object. Therefore, in this section it is ascertained if a detected shadow has an associated foreground object, in order to avoid the anomaly in case 2-2. Only shadows detected in the chrominance angle and brightness distortion analysis (section 4.6) will be tested. During a training period T_2 , the angles between the detected shadows and the foreground objects are calculated. The training period T_2 is used to learn the initial position of shadows. In the first frames of a shadow detection, the angle between the detected shadow and the foreground object is calculated based on the shadows detected in stage (4.5). In order to guarantee that these angles will be consistent with the illumination changes during long periods of time, the angle is updated using a first order recursive filter. Hereafter, the most probable angle obtained in the training period is used to discard detected shadows, which do not have any foreground object in its direction.

5. Top-Down Shadow Tracking

When a shadow has successfully been detected it is usually removed, since only the *Fg.* object is interesting for further processing. As a result, the shadow information is lost. Our idea is to use this information a posteriori, in order to improve the shadow detection when it fails (e.g., due to camouflage problems). Specifically, if a detected shadow is tracked over time instead of

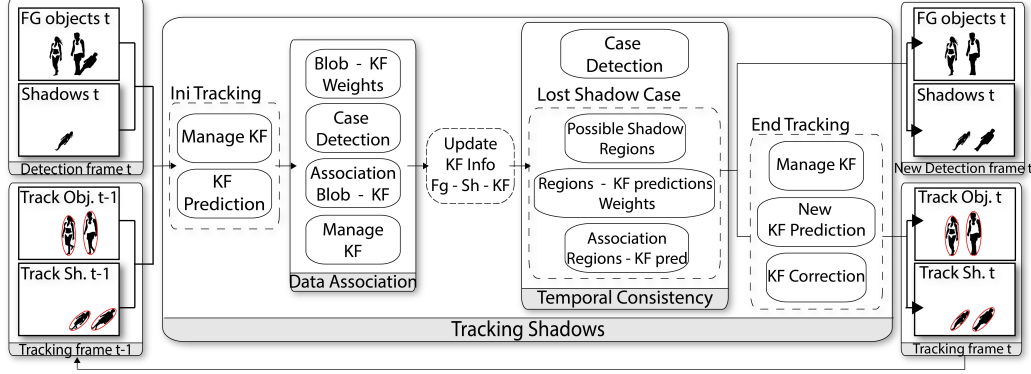


Figure 4: A schematic overview of the top-down shadow tracking process to enhance the chromatic shadow detection.

being discarded, it can be used to recover misdetracted shadows, and hereby improve shadow detection.

In this section a top-down approach is applied to enhance the chromatic shadow detection using a Kalman Filter (KF) based tracking. We choose a basic Kalman filter because we are interested in showing how the temporal consistence in the data association step used in tracking can be exploited to improve shadow detection. However, our proposed method can be applied to other different (non-linear) tracking methods. Hence, Extended Kalman filter or Unscented Kalman filter which are non-linear could be perfectly used with almost no modification of the proposed method.

Fig. 4 shows an overview of the top-down tracking process. First, the tracking module tracks objects and shadows through the scene. As input, the tracking module receives a binary mask from the object and shadow detection described in the previous section. The standard tracker and Probabilistic Appearance Models (PAMs) are explained in sections 5.1, and 5.2, respectively. The output of the tracker is a list of tracks for each object and shadow, and their mutual association. Secondly, in section 5.3 the association between foreground objects (FG) and shadows (SH) is described and updated for the Kalman filters (KFs). Thirdly (sec. 5.4), temporal consistency is investigated in the association between FG and SH blobs, and their assigned KFs, in order to identify possible lost shadows. Once the shadows are detected and tracked, the information is used as feedback for the chromatic shadow detection to recover misdetracted shadows in the original image (sec. 5.5). Finally, the KF and the PAM are updated by taking the infor-

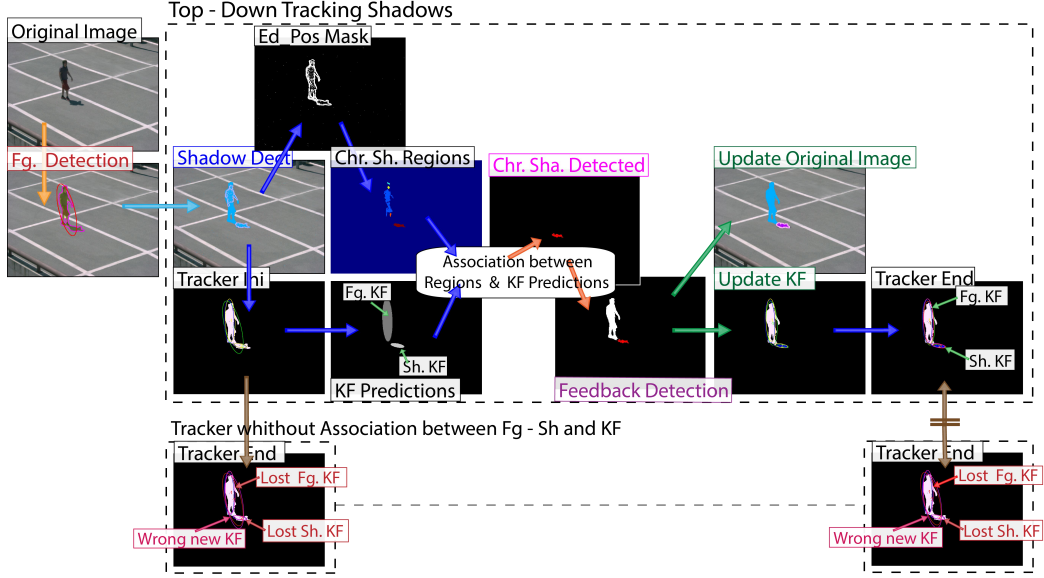


Figure 5: An example of the top-down tracking process. The figure illustrates the steps, within the tracking and motion segmentation, applied to enhance the shadow detection when a shadow is lost. See the main body text for further details.

mation from the new data association into account, and used for tracking in the next frames (sec. 5.6). An example of the entire process can be seen in the Fig. 5.

5.1. Tracking using Kalman filters

The detected foreground objects and shadows are tracked using first order Kalman filters. The tracking and data association are based on a number of estimated parameters for the detected objects and shadows:

- The centroid of an ellipse fitting.
- The major and minor axis length of the ellipse.
- The probabilistic Appearance Model.

Each track is associated with these parameters, and a standard Kalman filter is used to predict the object's location using a first order motion model. Hence, the target state is defined by $x_t = (posx_t, posy_t, velx_t, vely_t, maj_t, min_t, \theta_t)$, which establishes a state vector for every observation, and

adds the target speed and the size deformation rate at time t . Where $posx_t$ and $posy_t$ define the position (the centroid of the ellipse); $velx_t$ and $vely_t$ are velocity components; maj_t and min_t are the major and minor axis of the ellipse, respectively; and θ_t is the orientation. The KFs are initialised based on the detected foreground and shadow blobs, and the uncertainties are empirically estimated according to the precision of the detector. The foreground blobs extracted and classified as object or shadow (see section 4), are associated with a list of possible Kalman filters using a modification of the stable marriage algorithm [49], similar to the tracking system presented in [50], see them for details.

5.2. Probabilistic appearance models

Probabilistic appearance models inspired by [50] are applied for data association and to solve inter-object occlusion. Each track has its own PAM, which consists of an RGB colour model with an associated probability mask. The colour model, which is denoted $M_{RGB}(\mathbf{x})$, shows the appearance of each pixel of an object. $P_c(\mathbf{x})$ denotes the probability mask and represents the probability of the object being observed at that pixel. The use of PAMs can be viewed as weighted template matching, where the template is $M_{RGB}(\mathbf{x})$ and the weights are given by $P_c(\mathbf{x})$. The coordinates of \mathbf{x} are expressed using the coordinate system of the model, which is normalized to the object centroid. For each new track, a new PAM is created. In the object match situation, a track refinement step is applied before updating the model by finding the best fit in a small neighbourhood, e.g. 5×5 pixels. Track refinement increases the accuracy of the model; especially the colour model becomes sharper. Detail on building the model can be found in [50].

5.3. Object-shadow association

After the blobs (belonging to a FG or a SH) have been assigned to the KF, the association between which shadow belongs to which FG and vice versa is saved in the KF info for use in the next frames. This information is used to identify the possible cases in the association between FG and SH. An example showing the data association between the blobs and the KFs, and how the data association between FG and SH is saved in the KF info, can be seen in Fig. 6. The first image of Fig. 6 represents the FG detection provided in section 4.1. The second image shows the shadow detection presented in previous section 4, and how the FG segmentation is further analysed and divided into FG and SH blobs. These blobs are associated, since both are

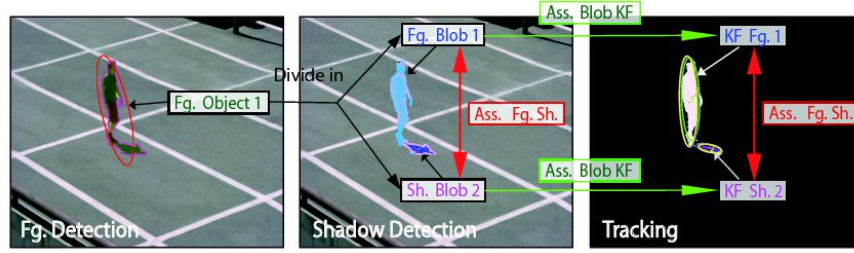


Figure 6: An example of the data association between FG, SH and the assigned KFs. The first image represents the FG detection. The second image shows the shadow detection, and how the analysed FG is divided into FG and SH blobs. In the third image the tracker has assigned one KF to each blob.

part of the same FG object. In the third image the tracking system has assigned one KF to each blob, and the data association between the FG and the SH blobs is saved in the KF info.

5.4. Temporal consistency in the data association

The information related to the association between FG and SH saved in the KF is analysed to check the possible data association cases, e.g., if a shadow has been lost.

When testing for temporal consistency in the data association between FG and SH with their respective KF, three main situations can occur:

- FG and SH match:
The association created at time $t-1$ continues at time t , which is the ideal case.
- New shadow: FG/SH splitting.
A new association between the FG and SH is created at time t . Temporal association of a splitting shadows is applied to avoid misdetected shadows in posterior frames.
- Lost shadow: FG-SH merging.
The association between the FG and SH at time $t-1$ is lost at time t , since the shadow is misdetected (Fig. 5).

The three cases are illustrated in Fig. 7. It is possible that a new shadow appears or a shadow is lost without merging of the FG object. However,

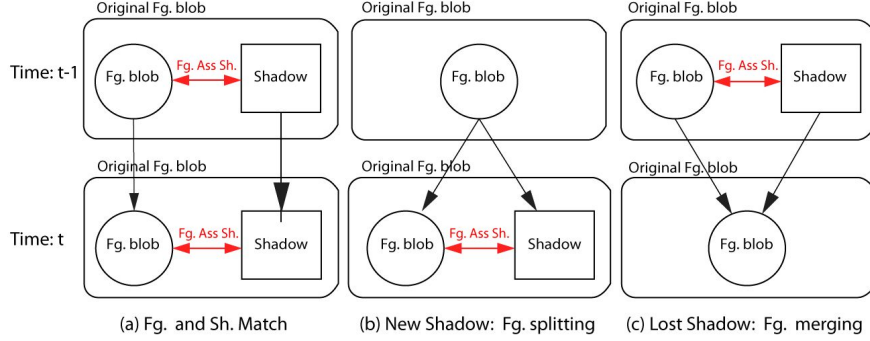


Figure 7: The three possible data association situations between FG, SH and their KFs. A rounded rectangle illustrates an original FG blob before shadow detection, a circle illustrates a FG and a square illustrates a SH from the shadow detection. A double red arrow indicates an association between FG and SH, and a black arrow indicates an association between FG and SH in the next frame.

these cases are not of interest, since they do not have any FG-SH association, hence, they will be tracked in the usual manner by the KFs.

The shadow region can be recovered by evaluating the FG blob (which contains the merged FG and SH), the blob prediction for the FG KF, and the blob prediction for the lost SH KF. The next subsections explain the recovery process for the lost shadow case and an example of the entire process is shown in Fig. 5.

5.4.1. Recovering lost shadows

The FG blob which belongs to the FG KF, associated in the previous frame with the shadow considered lost, is analysed in order to recover the possible shadow region. To do so, the mask of the positive edges (*Edpos* mask) plus morphological operators are applied, to divide the FG blob into FGs and possible shadow regions. Multiple regions can be found but theoretically only one is the shadow. This happens because the positive edges are used to divide the image, and these edges come from the current image. As explained in sec. 3, one of the characteristics of shadows is that they can only have negative edges, i.e., the edges from the background image. Therefore, theoretically several FG regions can be found but only one SH region. Fig. 5 shows how the original FG blob, detected as described in section 4, is subdivided into possible chromatic shadow regions using the *Edpos* mask (image Chr.Sh.Regions in the figure; the regions in the image are shown in differ-

ent colours). In the following, the new divided blobs from the regions are associated with the predictions of the KFs to recover the chromatic shadows.

5.4.2. Correspondence matrix for the new divided blobs

The weights of the blob prediction for the FG KF and the SH KF are calculated w.r.t. all possible regions found in the previous step. Therefore, two correspondence matrix are computed, where one contains the euclidean distance between the new blobs and the FG and SH KF predictions, and the other the overlapping (matching) between the new blobs and the FG and SH KF predictions. Next, these weights are applied to associate the SH and FG KF predictions with the blobs.

5.4.3. Association between KF predictions and the new blobs

The best match (shortest distance and best overlap) between the SH KF predictions and the blob will be considered as the shadow region, while the other blobs will be considered as FG blobs, since only one region can be shadow. Hence, the other blobs have to be FGs. In this way, by using the tracking information, the original FG blob can be segmented into FG and SH regions, and thereby recover misdetected chromatic shadows. This information is used as a feedback from the tracking to the shadow detection step. Fig. 5 shows how the blobs extracted from the divided regions are associated with the predictions of the KFs to detect the chromatic shadow.

5.5. Feedback to the chromatic shadow detection

Once the chromatic shadows are detected, the original image (original FG blob) is divided into one blob for the FGs and one blob for the SHs, using the information from the positive edges and the shadow tracking. Hence, the FG blob will be associated with the FG KF and the SH blob will be associated with the lost SH KF. Next, the original image is updated, so the misdetected chromatic shadows are now marked as detected. Fig. 5 shows how the detected chromatic shadow is updated according to the original blob after the feedback from the shadow tracking.

5.6. Managing and updating KFs and PAMs

Finally, the information related to the new associations between the FG and SH blobs, and their respective KFs must be updated. Additionally, the KFs must be updated with the new associated blobs, and the PAMs must



Figure 8: Moving foreground segmentation images; mask $M1$ in sec. 4.1 from [48] copes with achromatic shadows as it can be seen in ATON_Intelligentroom, ATON_Campus, ATON_Laboratory, CAVIAR, and NEMESIS database images (from left to right), because these databases do not present chromatic shadows. See text for details.

be updated considering the new blobs. It is possible that a new KF is erroneously created because one object together with its shadow were considered a new object. Therefore, the new KFs created in the data association between the blobs and the KFs must be checked. Any unused KFs, which were assigned to blobs considered as lost shadows, are deleted. In Fig. 5 also can be seen the output of the tracker with and without our Top-down approach (the last two images at the left called “Tracker End”). Top-down Tracker End image shows how the system recovers the detected chromatic shadows, hence the FG KF and SH KF are correctly updated. On the other hand, the image of the tracker without Top-down approach shows how the FG and SH KF are lost, and a new wrong KF is created.

Consequently, due to the data association between FG and SH we have achieved: (i) an enhancement of the chromatic shadow detection by recovering misdetecting shadows, which were incorrectly detected by the shadow detector (sec. 4). (ii) An improvement of the segmentation for high level processes, such as detection and tracking, by avoiding shadows. (iii) A more robust tracking by exploiting FG and SH association.

6. Experimental Results

The results presented in this section are from tests conducted on datasets selected from well-known databases. Our approach is tested on sequences of outdoor and indoor scenarios, and compared to other statistical approaches when results are available. The test sequences are relatively long and umbra and penumbra shadows are cast by multiple foreground objects. We have considered two different types of databases. On one hand, shadows datasets like ATON_Intelligentroom, ATON_Campus, ATON_Laboratory, CAVIAR, and NEMESIS do not exhibit chromatic shadows as it is stated in [14, 13, 51].

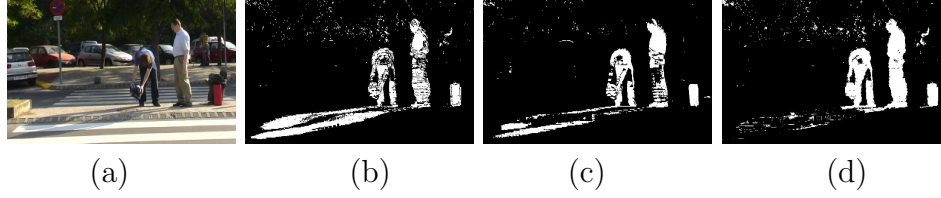


Figure 9: (a) An original image from the Outdoor_Cam1 sequence, and foreground results after shadow removal using (b) Huerta et al. approach [48], (c) Zivkovic et al. approach [5] using a shadow detector [15] and (d) using only our bottom-up approach.

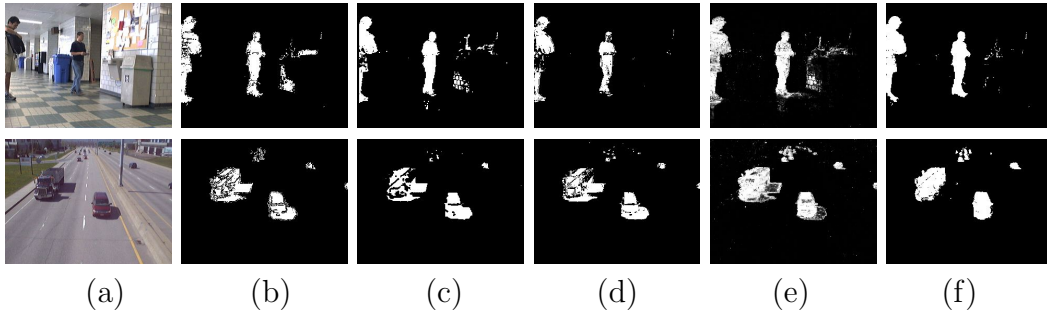


Figure 10: (a) Original images from the HallwayI and HighwayIII sequences, and foreground results after shadow removal using (b) [48], (c) [16], (d) [5] with a shadow detector [15], (e) [39] and (f) using only our bottom-up approach.

In these scenarios, the moving foreground segmentation presented in section 4.1 already copes well with such shadows, as already demonstrated in [48]. See Fig. 8 for a selected frames. On the other hand, we have considered another datasets like Outdoor_Cam1 (800 frames, 607x387 pixels), HighwayIII¹ (2227 frames, 320x240 pixels), HallwayI² (1800 frames, 320x240 pixels), HERMES_ETSEentrance (6500 frames, 640x480 pixels), CDW14 Bungalows² (1700 frames, 360x240 pixels), and CDW14 PeopleInShade (1199 frames, 380x244 pixels) which do exhibit chromatic shadows, and for which the whole proposed framework presented here is fully tested, as detailed next.

Figure 9 and Figure 10 show qualitative results when comparing our shadow detector with state-of-the-art approaches [16, 15, 39, 48]. As shown in these figures our approach outperforms the other analysed methods.

To evaluate our approach in a quantitative way, a comparison to several

¹<http://vision.gel.ulaval.ca/~CastShadows/>

²<http://changedetection.net/>

Sequence	HallwayI	
Method	η	ξ
Bottom-up approach	0.836	0.913
Qin et al. [28]	0.823	0.911
Martel-Bri [39]	0.724	0.867
Martel-Bri [38]	0.605	0.870

Table 1: Quantitative results: shadow detection rate (η) and shadow discrimination rate (ξ) for our bottom-up chromatic shadow detection approach and [28], [39], and [38]. Results using ground-truth from [38].

other shadow detection approaches is conducted, using the most employed quantitative expressions utilized to evaluate the shadow detection performance: the *Shadow Detection Rate* (η) and the *Shadow Discriminate Rate* (ξ). Please refer to [11] for the exact equations. Table 1 and 2 show the computed η and ξ values of each test sequence, comparing our Bottom-up (chromatic shadow detection) approach with other successful methods from the state of the art. Table 1 shows a comparative with [28, 39, 38] approaches, η and ξ are from their respective papers and the ground-truth is from [38]. Table 2 shows a comparative with [31, 15, 37, 20, 27] approaches, the code and ground-truth are from [47].

The results in Table 2 show that our bottom-up method outperforms the other approaches achieving *0.75* shadow average rate over all the sequences tested. Furthermore, It gives the highest shadow detection rate average ($\bar{\eta}$) *0.60* and the highest shadow discriminate rate average ($\bar{\xi}$) *0.91*, in other words it is the most stable and robust approach of all the methods.

For the HallwayI sequence our bottom-up method outperforms the other approaches. Sanin et al.’s method [31], which is also based on gradient information, gives very similar performance, while Qin et al.’s method [28] produces a similar shadow detection rate but at the cost of a markable reduction in the shadow discriminate rate, and Hsieh et al.’s [20] method gives a very poor performance for this particular sequence because of his geometrical approach. Notice this sequence is an indoor scenario, therefore the bluish effect mask is not used. Nevertheless, our bottom-up method obtains the highest values outperforming the other approaches.

For the HighwayIII sequence, our method performs the best when it comes to the shadow discriminate rate, however Sanin et al.’s method [31] gives a better shadow detection. This is because our approach needs a minimum

Sequence		Method					
		Bottom-up	[31]	[15]	[37]	[20]	[27]
HallwayI	η	0.954	0.952	0.938	0.826	0.487	0.961
	ξ	0.933	0.911	0.599	0.718	0.511	0.311
HighwayIII	η	0.478	0.622	0.477	0.439	0.427	0.057
	ξ	0.864	0.792	0.285	0.320	0.513	0.802
HERMES ETSE	η	0.586	0.023	0.178	0.068	0.800	0.147
	ξ	0.983	0.462	0.549	0.506	0.775	0.634
CDW14 Bun	η	0.362	0.546	0.043	0.149	0.677	0.059
	ξ	0.876	0.353	0.464	0.669	0.079	0.448
CDW14 People	η	0.289	0.348	0.078	0.07	0.667	0.368
	ξ	0.947	0.926	0.920	0.911	0.116	0.658
Mean	$\bar{\eta}$	0.53	0.50	0.34	0.31	0.61	0.32
	$\bar{\xi}$	0.92	0.69	0.56	0.62	0.40	0.57
	$\bar{\eta, \xi}$	0.73	0.59	0.45	0.47	0.51	0.44

Table 2: Quantitative results: shadow detection rate (η) and shadow discrimination rate (ξ) for our bottom-up chromatic shadow detection approach and a number of other successful methods tested on three challenging video sequences. $\bar{\eta, \xi}$ is the total average of both the shadow detection rate and the shadow discrimination rate. Our bottom-up approach is the most stable and robust for all the sequences. See text for details.

area for the shadow regions, or there might not be enough information for a proper shadow detection and classification. Our method produces a lower detection rate in this sequence because the moving cars are very small when they are far away from the camera.

For the HERMES_ETSEentrance sequence, Sanin et al.’s [31] method gives a very poor performance, since there is no or very little gradient information in the background, which is crucial for this solely texture-based approach. Our method performs the best when it comes to the shadow discriminate rate. However, the approach proposed by Hsieh et al. [20] gives a better shadow detection. This geometric approach performs very well for this particular sequence, since the people and shadows have different orientations and are clearly defined. However, this method usually fails when the objects and shadows are divided, or there are multiple shadows, or the objects and the shadows have the same orientation. This can be seen in the lower values obtained for the other test sequences.

Overall, our multi-stage approach, which takes advantage of known shadow properties, texture and invariant colour models, and spatial geometry gives the most robust and stable performance for all indoor and outdoor scenarios.

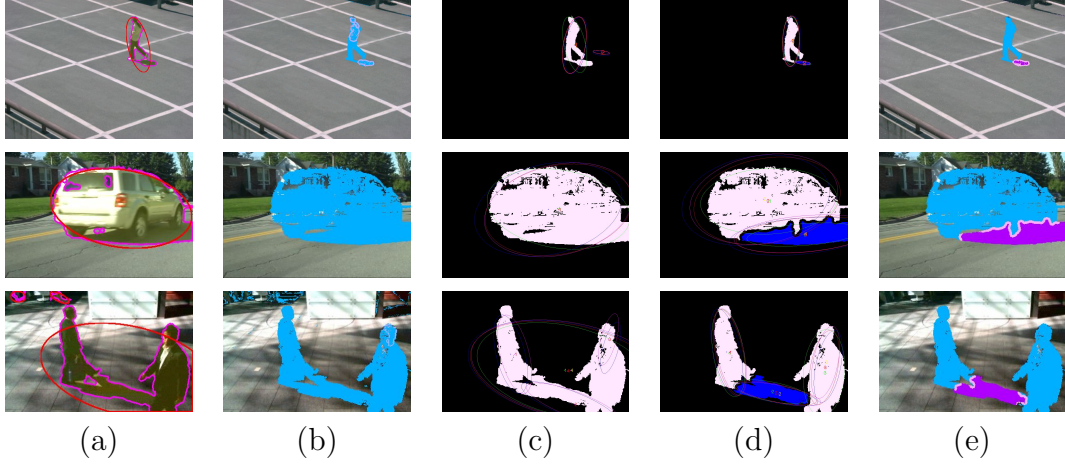


Figure 11: Top-down shadow recovery examples for the HERMES_ETSEentrance, CDW14_bungalows and CDW14_PeopleInShade sequences. (a) foreground detection images. (b) bottom-up chromatic shadow detection results (shadows are not correctly detected). (c) output of the tracker without applying our top-down approach (SH KFs are lost and incorrectly updated). (d) output of the tracker using our top-down process (shadows are detected, FG KFs and the SH KFs are correctly updated, and none of them are lost). (e) final chromatic shadow detection, where pink pixels correspond to shadows.

The top-down process assists the chromatic shadow detector when it fails to detect shadows, as shown in Figure 11.c. The tracking system is able to track the shadows, and use this information as feedback to the chromatic shadow detector. Hence, the misdetected shadows can be recovered and correctly detected. Figure 11 presents examples of shadow recovery using our top-down approach for the HERMES_ETSEentrance (Fig. 11 first row), CDW14_bungalows (Fig. 11 second row), and CDW14_PeopleInShade (Fig. 11 third row) sequences. Fig. 11.a shows the foreground detection results from section 4.1 (mask $M1$). In Fig. 11.b the chromatic shadow detection results from our bottom-up approach are shown. Note that the shadows are not correctly detected. Fig. 11.c shows the output of the tracker without applying our top-down approach. The KFs associated with the shadows are lost and therefore falsely updated. While Fig. 11.d shows the output of the tracker, after the chromatic shadows are recovered, using our top-down process. The shadows are accurately detected, therefore the FG KFs and the SH KFs are correctly updated, and none of them are lost. The a posteriori state is depicted with a red ellipse. Finally, Fig. 11.e shows how the chromatic shadows are accurately detected after feedback from the

Sequence	HERMES ETSE		CDW14 Bungalows		CDW14 PeopleInSh	
Method	η	ξ	η	ξ	η	ξ
Bottom-up	0.586	0.983	0.362	0.876	0.289	0.947
Top-down	0.669	0.975	0.445	0.839	0.527	0.940

Table 3: Quantitative results: shadow detection rate (η) and shadow discrimination rate (ξ) comparing our bottom-up: chromatic shadow detection with our top-down: shadow tracking. With our shadow tracking the shadow detection rate has been improved by approximately 13%.

tracker to the chromatic shadow detector.

To evaluate quantitatively the results of our shadow tracking we have hand-segmented 1500 frames from the HERMES_ETSEentrance sequence ³. Table 3 shows the results when applying our top-down shadow tracking to the HERMES_ETSEentrance, CDW14 Bungalows and CDW14 PeopleInShade from HERMES and Changedetection datasets, the shadow detection rate is improved by approximately 13%, which effectively means that we detect more shadow regions using shadow tracking. Additionally, the shadow discriminate rate is pretty stable.

First, a discussion of the limitations of the current approach is presented. Later, some remarks on the computational complexity and the execution time for a possible real-time application are discussed. The tracking method employed is a basic Kalman filter, which is limited in some aspects (linear, not multi-hypothesis, not able to track multiple objects in crowded scenes). However, the main goal of the proposed approach is not to cope with advanced tracking problems, but to show how shadow detection can be improved by exploiting the temporal consistence of the data association between foreground objects and shadows. Other more complex trackers, like the ones presented in [52], could easily be used with slightly changes of the approach, in order to tackle problems such as multiple occlusions.

In terms of computational complexity, motion segmentation has a cost that is linear to the number of pixels in the image. The specific implementation used in the experiments [48] runs at about 3 fps in Matlab. However,

³Original sequence, foreground and shadows hand-segmented ground truth are available at <http://www.cvc.uab.es/~ivanhc/shadows/huertaShaDows.html> for comparison purposes

a faster reimplementaion or the use of other algorithms [11, 5] can lead to realtime performance. The shadow detection (bottom-up) part runs at about 28 fps in Matlab. Since the models are already computed from [48], so there is no need to calculate them again. If another faster motion segmentation approach is used, then the colour and edge models should be calculated, which runs at about 6 fps in Matlab. The shadow tracking (top-down) part runs at around 10 fps in Matlab. The time will increase with the number of elements to track, but tracking a shadow will have the same computational cost as tracking a new object.

7. Conclusion

In this paper, we have presented two main contributions for surveillance analysis: (i) a novel bottom-up approach removes chromatic moving shadows, and (ii) a top-down approach based on motion filters keeps track of both objects and shadows to handle chromatic shadow misdetections. Unlike other approaches, our method does not make any a-priori assumptions about camera location, surface geometries, surface textures, shapes and types of shadows, objects, and background properties.

The bottom-up algorithm exploits gradient and colour information for separating chromatic moving shadows from moving objects. Invariant colour cone and invariant gradient models are used to detect potential shadow regions. Subsequently, these regions are grouped by considering the bluish effect together with edge partitioning. Lastly, (i) temporal similarities between local gradient structures and (ii) spatial similarities between chrominance angle and brightness distortions are analysed to finally identify those umbra shadows. Unfortunately, there are still shadows incorrectly classified, mainly due to camouflage.

In order to cope with the camouflage problem, the top-down technique tracks both objects and shadows using basic motion filters. By performing data association and exploiting the knowledge from mutual information between object and shadows, we are able to detect more shadows thus improving the final segmentation performance compared with the state-of-the-art. Experimental results obtained in well-known outdoor and indoor sequences demonstrate the advantages of our proposed approach since its performance is proven to be more robust and accurate compared to current state-of-the-art techniques.

Acknowledgments.

This work has been partially funded by the European project CargoANTs (FP7-SST-2013- 605598), by the Spanish CICYT project DPI2013-42458-P and CICYT TIN2012-39051.

References

- [1] I. Huerta, M. Pedersoli, J. González, A. Sanfeliu, Combining where and what in change detection for unsupervised foreground learning in surveillance, *Pattern Recognition* 48 (3) (2015) 709–719.
- [2] T. B. Moeslund, A. Hilton, V. Kruger, A survey of advances in vision-based human motion capture and analysis, *CVIU* 104 (2006) 90–126.
- [3] T. B. Moeslund, A. Hilton, V. Kruger, L. Sigal, *Visual Analysis of Humans - Looking at people*, Springer, ISBN:978-0-85729-996-3, 2011.
- [4] K. Toyama, J.Krumm, B.Brummitt, B.Meyers, Wallflower: Principles and practice of background maintenance, in: *Proc. ICCV’99*, Vol. 1, Kerkyra, Greece, 1999, pp. 255–261.
- [5] Z. Zivkovic, F. Heijden, Efficient adaptive density estimation per image pixel for the task of background subtraction, *PRL* 27 (7) (2006) 773–780.
- [6] O. Barnich, M. V. Droogenbroeck, Vibe: A universal background subtraction algorithm for video sequences, *TIP* 20 (6) (2011) 1709–1724.
- [7] Y. Sheikh, M. Shah, Bayesian modeling of dynamic scenes for object detection, *IEEE TPAMI* 27 (11) (2005) 1778–1792.
- [8] L. Maddalena, A. Petrosino, A self-organizing approach to background subtraction for visual surveillance applications, *IEEE TIP* 17 (7) (2008) 1168–1177.
- [9] S. Liao, G. Zhao, V. Kellokumpu, M. Pietikainen, S. Z. Li, Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes, in: *CVPR*, France, 2010.
- [10] J. Pilet, C. Strecha, P. Fua, Making background subtraction robust to sudden illumination changes, in: *ECCV’08*, France, 2008.

- [11] A. Prati, I. Mikic, M. Trivedi, R. Cucchiara, Detecting moving shadows: Algorithms and evaluation, *TPAMI* 25 (7) (2003) 918–923.
- [12] W. Zhang, Q. J. Wu, X. Z. Fang, Moving cast shadow detection, in: *VSSPR*, 2007.
- [13] A. Amato, I. Huerta, M. Mozerov, F. Roca, J. González, Moving cast shadows detection methods for video surveillance applications, *Augmented Vision and Reality*, Springer Berlin Heidelberg, 2013, pp. 1–25.
- [14] I. Huerta, M. Holte, T. Moeslund, J. González, Detection and removal of chromatic moving shadows in surveillance scenarios, in: *ICCV*, 2009.
- [15] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, Detecting moving objects, ghosts, and shadows in video streams, *TPAMI* 25 (10) (2003) 1337–1342.
- [16] K. Kim, T. Chalidabhongse, D. Harwood, L. Davis, Real-time foreground-background segmentation using codebook model, *Real-Time Imaging* 11 (3) (2005) 172–185.
- [17] K. Onoguchi, Shadow elimination method for moving object detection, in: *ICPR*, 1998.
- [18] Y. Ivanov, A. Bobick, J. Liu, Fast lighting independent background subtraction, *IJCV* 37 (2) (2000) 199–207.
- [19] E. Salvador, A. Cavallaro, T. Ebrahimi, Cast shadow segmentation using invariant color features, *CVIU* 95 (2) (2004) 238–259.
- [20] J.-W. Hsieh, W.-F. Hu, C.-J. Chang, Y.-S. Chen, Shadow elimination for effective moving object detection by gaussian shadow modeling, *Image and Vision Computing* 21 (6) (2003) 505–516.
- [21] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, S. Sirotti, Improving shadow suppression in moving object detection with hsv color information, in: *Intelligent Transportation Systems*, 2001.
- [22] Z. Liu, K. Huang, T. Tan, L. Wang, Cast shadow removal combining local and global features, in: *CVPR*, 2007.

- [23] N. Martel-Brisson, A. Zaccarin, Moving cast shadow detection from a gaussian mixture shadow model, in: CVPR, 2005.
- [24] O. Schreer, I. Feldmann, U. Goelz, P. Kauff, Fast and robust shadow detection in videoconference applications, in: VIPromCom, 2002.
- [25] T. Horprasert, D. Harwood, L.S. Davis, A statistical approach for real-time robust background subtraction and shadow detection, in: Frame-Rate Applications Workshop, 1999.
- [26] F. Porikli, J. Thornton, Shadow flow: a recursive method to learn moving cast shadows, in: ICCV, 2005.
- [27] A. Leone, C. Distanto, Shadow detection for moving objects based on texture analysis, PR 40 (4) (2007) 1222–1233.
- [28] R. Qin, S. Liao, Z. Lei, S. Z. Li, Moving cast shadow removal based on local descriptors, in: ICPR, 2010.
- [29] W. Zhang, X. Z. Fang, X. K. Yang, Q. J. Wu, Moving cast shadows detection using ratio edge, Trans. Multimedia 9 (6) (2007) 1202–1214.
- [30] C.-C. Chen, J. K. Aggarwal, Human shadow removal with unknown light source, in: ICPR, 2010.
- [31] A. Sanin, C. Sanderson, B. Lovell, Improved shadow removal for robust person tracking in surveillance scenarios, in: ICPR, 2010.
- [32] M. Heikkila, M. Pietikainen, A texture-based method for modeling the background and detecting moving objects, TPAMI 28 (4) (2006) 657–662.
- [33] J. Stauder, R. Mech, J. Ostermann, Detection of moving cast shadows for object segmentation, Trans. Multimedia 1 (1) (1999) 65–76.
- [34] G. Finlayson, S. Hordley, C. Lu, M. Drew, On the removal of shadows from images, TPAMI 28 (1) (2006) 59–68.
- [35] G. D. Finlayson, M. S. Drew, C. Lu, Entropy minimization for shadow removal, IJCV 85 (1) (2009) 35–57.

- [36] Y. Weiss, Deriving intrinsic images from image sequences, in: ICCV, 2001.
- [37] J. Huang, C. Chen, Moving cast shadow detection using physics-based features, in: CVPR, 2009.
- [38] N. Martel-Brisson, A. Zaccarin, Learning and removing cast shadows through a multidistribution approach, TPAMI 29 (7) (2007) 1133–1146.
- [39] N. Martel-Brisson, A. Zaccarin, Kernel-based learning of cast shadows from a physical model of light sources and surfaces for low-level segmentation, in: CVPR, 2008.
- [40] S. Nadimi, B. Bhanu, Physical models for moving shadow and object detection in video, TPAMI 26 (8) (2004) 1079–1087.
- [41] J. Tian, J. Sun, Y. Tang, Tricolor attenuation model for shadow detection, TIP 18 (10) (2009) 2355–2363.
- [42] J. Zhu, K. G. Samuel, S. Z. Masood, M. F. Tappen, Learning to recognize shadows in monochromatic natural images, in: CVPR, 2010.
- [43] R. Guo, Q. Dai, D. Hoiem, Single-image shadow detection and removal using paired regions, in: CVPR, 2011.
- [44] J.-F. Lalonde, A. A. Efros, , S. G. Narasimhan, Estimating the natural illumination conditions from a single outdoor image, IJCV, Preprint.
- [45] E. Arbel, H. Hel-Or, Shadow removal using intensity surfaces and texture anchor points, TPAMI 33 (6) (2011) 1202–1216.
- [46] X. Huang, G. Hua, J. Tumblin, L. Williams, What characterizes a shadow boundary under the sun and sky?, in: ICCV, 2011.
- [47] A. Sanin, C. Sanderson, B. Lovell, Shadow detection: A survey and comparative evaluation of recent methods, PR 45 (4) (2012) 1648–1695.
- [48] I. Huerta, A. Amato, F. X. Roca, J. González, Exploiting multiple cues in motion segmentation based on background subtraction, Neurocomputing 100 (6) (2013) 183–196.

- [49] D. Gusfield, The stable marriage problem: structure and algorithms, MIT Press.
- [50] A. Senior, A. Hampapur, Y.-L. Tian, L. Brown, S. Pankanti, R. Bolle, Appearance models for occlusion handling, *Image and Vision Computing* 24 (11) (2006) 1233–1243.
- [51] A. Amato, M. Mozerov, A. Bagdanov, J. Gonzalez, Accurate moving cast shadow suppression based on local color constancy detection, *TIP* 20 (10) (2011) 2954–2966.
- [52] A. Yilmaz, M. S. O. Javed, Object tracking: A survey, in: *Acm computing surveys (CSUR)*, 2006.