Detection and Removal of Chromatic Moving Shadows in Surveillance Scenarios

Ivan Huerta[†], Michael Holte[‡], Thomas Moeslund[‡], and Jordi Gonzàlez[†]

[†]Computer Vision Center & Department of Computer Science (UAB), Bellaterra, Spain
[‡]Computer Vision and Media Technology Laboratory, Aalborg Universtity, Aalborg, Denmark
ivan.huerta@cvc.uab.es

Abstract

Segmentation in the surveillance domain has to deal with shadows to avoid distortions when detecting moving objects. Most segmentation approaches dealing with shadow detection are typically restricted to penumbra shadows. Therefore, such techniques cannot cope well with umbra shadows. Consequently, umbra shadows are usually detected as part of moving objects. In this paper we present a novel technique based on gradient and colour models for separating chromatic moving cast shadows from detected moving objects. Firstly, both a chromatic invariant colour cone model and an invariant gradient model are built to perform automatic segmentation while detecting potential shadows. In a second step, regions corresponding to potential shadows are grouped by considering "a bluish effect" and an edge partitioning. Lastly, (i) temporal similarities between textures and (ii) spatial similarities between chrominance angle and brightness distortions are analysed for all potential shadow regions in order to finally identify umbra shadows. 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. Experimental results show the performance and accuracy of our approach in different shadowed materials and illumination conditions.

1. Introduction

A fundamental problem for all automatic video surveillance systems is to detect objects of interest in a given scene. A commonly used technique for segmentation of moving objects is background subtraction [12]. This involves detection of moving regions (i.e., the foreground) in an image by differencing the current image and a reference background image in a pixel-by-pixel manner. Usually, the background image is represented by a statistical background model, which is initialised over some time period.

An important challenge for foreground segmentation is the impact of shadows. Shadows can be devided into two catogories: *static shadows* and *dynamic (or 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 incorporated into the background model, while dynamic shadows have shown to be more problematic. Dynamic shadows are due to moving objects (e.g., people, vehicles, etc.). The impact of dynamic shadows can be crucial for the foreground segmentation, and cause objects to merge, distort their size and shape, or occlude other objects. This results in a reduction of computer vision algorithms' applicability for e.g, scene monitoring, object recognition, target tracking and counting.

Dynamic shadows can take any size and shape, and can be both *umbra* (dark shadow) and *penumbra* (soft shadow) shadows. Penumbra shadows exhibit low values of intensity but similar chromaticity values w.r.t. the background, while 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 from the chormaticity of the global background illumination, we define this as *chromatic shadow*. Consequently, umbra shadows are significatly more difficult to detect, and therefore usually detected as part of moving objects.

In this paper we propose an approach for detection and removal of chromatic moving shadows in surveillance scenarios. We present a novel technique based on gradient and colour models for separating chromatic moving shadows from detected moving objects. Firstly, both a chromatic invariant colour cone model and an invariant gradient model are built to perform automatic segmentation while detecting potential shadows. In a second step, regions corresponding to potential shadows are grouped by considering "a bluish effect" and an edge partitioning. Lastly, (i) temporal similarities between textures and (ii) spatial similarities between chrominance angle and brightness distortions are analysed for all potential shadow regions in order to finally identify umbra shadows.

The remainder of the paper is organised as follows. The state of the art in the field of shadow detection will be discussed in section 2, along with our contributions. 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. Finally, we present experimental results in section 5 and concluding remarks in section 6.

2. State of the Art and Our Contributions

Shadow detection is an extensive field of reseach within computer vision. Even though many algorithms have been proposed in the literature, the problem of detection and removal of shadows in complex environment is still far from being completely solved.

A common direction is to assume that shadows decrease the luminance of an image, while the chrominance stays relatively unchanged [1, 8]. However, this is not the case in many scenarios, e.g., in ourdoor scenes. Other approaches applies geometrical information. Onoguchi [15] uses two cameras to eliminate the shadows of pedestrians based on object height. However, objects and shadows must be visible to both cameras. Ivanov et al. [7] use a disparity model, which is invariant to arbitrarily rapid changes in illumination, for modeling background. However, to overcome rapid changes in illumination, at least three cameras are required. in [17], Salvador et al. use 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. It should be noted that most of the approaches which apply geometrical information normally requires shadows to be on a flat plane.

Another popular approach is to exploit colour differences between shadow and background in different colour spaces. In [2], Cucchiara et al. use the hypothesis that shadows reduce surface brightness and saturation while maintaining hue properties in the HSV colour space. While Schreer et al. [18] adopt the YUV colour space. In [5, 8], Horprasert et al. and Kim et al. build a model in the RGB colour space to express normalised luminance variation and chromaticity distortions. However, these methods requires all illumination sources to be white, and assumes shadow and non-shadow have similar chrominance. A number of approaches use textures to obtain a segmentation without shadows, such as Heikkila et al. [4] who uses Local Binary Patterns. However, it fails to detect umbra shadows.

To overcome some of these prior mentioned shortcomings, some authors use colour constancy methods, combine different techniques or use multi-stage approaches. In addition to scene brightness properties, [19] uses edge width information to differentiate penumbra regions from the background. In [3], Finlayson et al. use shadow edges along with illuminant invariant images to recover full colour shadow-free images. Nonetheless, a part of the colour information is lost in removing the effect of the scene illumination at each pixel in the image. Weiss [20] uses the reflectance edges of the scene to obtain an intrinsic image without shadows. However, this approach requires significant changes in the scene, and as a result the reflectance image also contains the scene illumination. An extension of Weiss is done by Matsushita et al. [11], which is also based on intrinsic images. Nevertheless, this method does not consider cast shadows of moving objects, only those cast by static objects such as buildings and trees. Martel et al. [10] introduce a nonparametric framework based on the physical properties of light sources and surfaces, and applies spatial gradient information to reinforce the learning of model parameters. Finally, [13] applies a multi-stage approach for ourdoor scenes, which is based on a spatio-temporal albedo test and dichromatic reflection model. A comparative and evaluation study of shadow detection techniques can be found in [16].

We also apply a multi-stage approach inspired by [13] but we use colour, gradient and textural information, together with known shadow properties. The contribution of this paper is threefold: (i) We combine an invariant colour cone model and an invariant gradient model to improve foreground segmentation and detection of potential shadows. (ii) We extend the shadow detection to cope with chromatic moving cast shadows by grouping potential shadow regions and considering "a bluish effect", edge partioning, spatial similarities between textures, and temporal similarities between chrominance angle and brightness distortions. (iii) Unlike other approaches, our method does not make any assumptions about camera location, surface geometries, surface textures, shapes and types of shadows, objects, and background.

3. Analysis of Shadow Properties

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.

$$\rho_a = \sigma \int E(\lambda)R(\lambda)Q_a(\lambda)d\lambda \tag{1}$$

The surface reflectance $R(\lambda)$ depends on the material. Hence, every material have different response to the same illumination change.

3.1. Applying textural information

By applying gradient information we can obtain knowledge about object bounderies, and thereby improve the for-

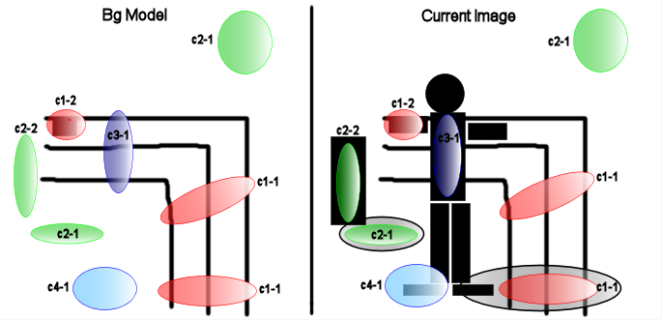

Figure 1. A sketch of the four main cases (c1-1 to c4-1) and two anormlies (c2-1 and c2-2) that can occur when performing foreground segmentation with the influence of shadows, and taking the temporal textures into account. The ellipses represent detection of potential chromatic shadows. They are grouped by considering an intensity reduction, "the bluish effect", and an edge partition.

ground segmentation. Additionally, the gradient can also provide textural information of both the background and foreground image. Although shadows will result in a reduction in the intensity of the illumnation, and the texture of a given object or the background will have lower gradient magnitude, the structure will still appear the same. Hence, the gradient orientation will be unchanged. This knowledge can be applied to identify shadows.

3.2. Applying the bluish effect

In outdoor scenes, the environment is illuminated by two light sources: a point light source (the sun) and a diffuse source (the sky) with different SPD $E(\lambda)$. Besides a reduction in the intensity, an outdoor cast shadow will result in a change of the chorminance. The illumination of the sky has higher power components in the lower wavelenghts λ (450 - 495 nm) of the visible spectrum, and it is therefore assumed bluish as argued in [13]. 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. This "bluish effect" and the chrominacne distortion can be exploited for shadow detection and grouping of potential shadow pixels.

3.3. Shadow scenaria and solutions

When performing fg. segmentation with the influence of shadows, and taking the temporal textures into account, four main cases can occur as illustrated in figure 1. The ellipses represent detection of potential chromatic shadows. They are grouped by considering an intensity reduction, "the bluish effect", and an edge partition. The entire shadow detection process will be explained in depth in section 4.

Case 1: Textures are present in the background model and in the current image, and they are similar. By examining similarities between the textures, and the fact that there is no foreground object in the current image, potential shadows can be detected and identified as shadow regions (case 1-1). However, if a foreground object is present, it can be miss-classified as shadow if the textures of the background and the foreground object are similar (case 1-2).

Case 2: There is no available background model nor texture 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 (case 2-1). However, a foreground object can be miss-classified as shadow if the foreground object has no texture. Furthermore, the chrominance angle distortion can also be similar among the pixels in the region of the object (case 2-2).

Case 3: Texture is present in the background model but not in the current image. By examining similarities between temporal textures, potential shadow can be correctly detected as a foreground object if there is background texture and a new foreground object in the current image.

Case 4: Texture is present in the current image but not in the background model. Then there must be a new foreground object in the current image. In this case, the texture in the current image is employed to detect shadow regions. Hence, there is no need to analyse the potential region further.

The described characteristics are not sufficient to address these anomalies in case 1-2 and case 2-2. Therefore, we take futher precautions and apply some additional steps, which will be explained in section 4. Furthermore, it should be noted that these additional steps also improves the shadow detection in some of the four main cases.

4. Chromatic Shadow Detection

The approach, depicted in Fig. 2, is a multi-stage approach. Our approach is mainly divided in two parts: the first stage obtains the potential chromatic shadow regions (4.1 to 4.3), and the second stage classifies them into foreground or chromatic shadow (4.4 to 4.8). In particular, the first step (4.1) obtains the pixels classified as foreground including the chromatic shadow regions. The second step, the shadow reduction step (4.2), reduces the number of these pixels which cannot be shadow, and the bluish effect step (4.3) reduces those foreground pixels which cannot be chromatic shadow, thereby avoiding false positive and false negative pixels. In the second part (4.4 to 4.8 steps): the fourth step (4.4) divide the pixels obtained in the previous steps

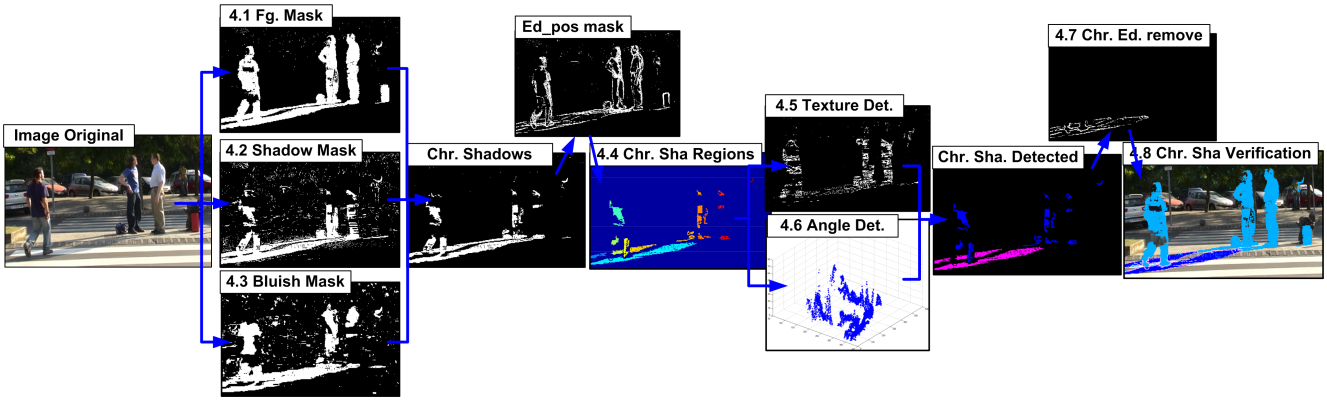

Figure 2. An overview of the chromatic shadow detection approach. Each of the numbers added to the image captions corresponds to the sub-sections in section 4.

in regions of potential shadows. Chromatic shadow detection is realised in steps (4.5) and (4.6) based on textures and chrominance angles. The last two steps (4.7 and 4.8) removes the edges of the chromatic shadows detected, and avoids foreground regions detected wrongly as chromatic shadows, respectively.

4.1. Moving foreground segmentation

In this stage foreground objects, shadows, and some erroneous pixels are segmented. In order to achieve the moving foreground segmentation an improved hybrid approach based on [6], which fuses colour and gradient information, is used. Note that the results from this approach can cope with several motion segmentation anomalies, among them it can cope with penumbra shadows because it is based on a chromatic colour model [5].

We use similar architecture and automatic threshold selection as the hybrid approach in [6]. This architecture provides the highest detection rate in comparison to other motion segmentation approaches [6]. However, the colour and gradient model are changed in order to get more accurate segmentation and use them for the next stages.

The chromatic cylinder model employed in many motion segmentation approaches [6, 5, 8] is changed into a new chromatic cone model. It uses chrominance angle distortion instead of chromatic distortion. For the same chromaticity line the chromatic distortion used in the above mentioned papers depends on the brightness distortion, while the chrominance angle distortion is invariant to the brightness, as it can be seen in Fig. 3 (the chromatic distortion δ increases propotional to the brightness distortion α, while the chrominance angle distortion β is equal). The invariant chromatic cone model is more robust towards chromatic shadows because these shadows (umbra shadows) modifies both the brightness and the chromaticity.

A new invariant gradient model is employed in order to identify the different textures of the scene. As argued in [10, 14], the gradient model has to be invariant towards global and local illuminations changes, such as shadows. The gradient model presented in this paper uses a new combination of gradient magnitudes and gradient directions which is invariant to illumination changes.

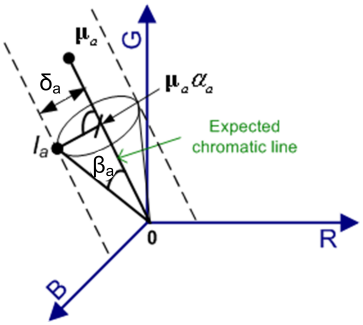

Figure 3. A colour cone model, where $\boldsymbol{\mu}_a$ represents the expected RGB colour value for a pixel a, while $\mathbf{I}_a$ is the current pixel value. The line $\overline{\mathbf{0}\boldsymbol{\mu}_a}$ shows the expected chromatic line, and all colours along this line have the same chrominance but different brightness. α_a and β_a give the current brightness and chrominance angle distortion, respectively.

4.1.1 Invariant colour cone model

The Background Colour Model (BCM) is computed according to the chromatic invariant cone representation shown in Fig. 3. First, the RGB mean $\boldsymbol{\mu}_a = \left(\mu_a^R, \mu_a^G, \mu_a^B\right)$ and standard deviation $\boldsymbol{\sigma}_a = \left(\sigma_a^R, \sigma_a^G, \sigma_a^B\right)$ of every image pixel a during the time period $t = [1 : T_1]$ are computed.

Once each RGB component is normalised by their respective standard deviation σ_a^c, two distortion measures are established during the training period: the brightness distortion, $\alpha_{a,t}$, and the chrominance angle distortion, $\beta_{a,t}$. The brightness distortion can be computed by minimising the distance between the current pixel value $\mathbf{I}_{a,t}$ and the chromatic line $\overline{\mathbf{0}\boldsymbol{\mu}_a}$. The angle between $\overline{\mathbf{0}\boldsymbol{\mu}_a}$ and $\overline{\mathbf{0}\boldsymbol{I}_a}$ is, in fact,

the chrominance angle distortion. Thus, the brightness and the chrominance angle distortions are given by:

$$\alpha_{a,t} = \frac{\frac{I^R_{a,t}\mu^R_a}{(\sigma^R_a)^2} + \frac{I^G_{a,t}\mu^G_a}{(\sigma^G_a)^2} + \frac{I^B_{a,t}\mu^B_a}{(\sigma^B_a)^2}}{\left(\frac{\mu^R_a}{\sigma^R_a}\right)^2 + \left(\frac{\mu^G_a}{\sigma^G_a}\right)^2 + \left(\frac{\mu^B_a}{\sigma^B_a}\right)^2} \quad (2)$$

$$\beta_{a,t} = \arcsin \frac{\sqrt{\sum_{c=R,G,B}\left(\frac{I^c_{a,t} - \alpha_{a,t}\mu^c_a}{\sigma^c_a}\right)^2}}{\sqrt{\sum_{c=R,G,B}\left(\frac{I^c_{a,t}}{\sigma^c_a}\right)^2}} \quad (3)$$

Finally, the Root Mean Square over time of both distortions for each pixel are computed: $\bar{\alpha}_a$ and $\bar{\beta}_a$, respectively:

$$\bar{\alpha}_a = RMS\left(\alpha_{a,t} - 1\right) = \sqrt{\frac{1}{T_1}\sum_{t=0}^{T_1}\left(\alpha_{a,t} - 1\right)^2} \quad (4)$$

$$\bar{\beta}_a = RMS\left(\beta_{a,t}\right) = \sqrt{\frac{1}{T_1}\sum_{t=0}^{T_1}\left(\beta_{a,t}\right)^2} \quad (5)$$

where 1 is subtracted from $\alpha_{a,t}$, so that the brightness distortion is now distributed around zero: positive values represent brighter pixels, whereas negative ones represent darker pixels, w.r.t. the learnt values. These values are used as normalising factors so that a single threshold can be set for the whole image. This 4-tuple $BCM =< \boldsymbol{\mu}_a, \boldsymbol{\sigma}_a, \bar{\alpha}_a, \bar{\beta}_a >$ constitutes the pixel-wise colour background model.

4.1.2 Invariant gradient model

The Background Edge Model (BEM) is built as follows: first the Sobel edge operator is applied to each colour channel in horizontal and vertical directions. This yields both a horizontal $G^c_{x,a,t} = S_x * I^c_{a,t}$ and a vertical $G^c_{y,a,t} = S_y * I^c_{a,t}$ gradient image for each frame during the training period $t = [1 : T_1]$, where $c \in \{R, G, B\}$ denotes the channel.

Next, the gradient of each background pixel is modelled using the gradient mean $\mu_{Gx,a} = (\mu^R_{Gx,a}, \mu^G_{Gx,a}, \mu^B_{Gx,a})$ and $\mu_{Gy,a} = (\mu^R_{Gy,a}, \mu^G_{Gy,a}, \mu^B_{Gy,a})$, and the gradient standard deviation $\sigma_{Gx,a} = (\sigma^R_{Gx,a}, \sigma^G_{Gx,a}, \sigma^B_{Gx,a})$ and $\sigma_{Gy,a} = (\sigma^R_{Gy,a}, \sigma^G_{Gy,a}, \sigma^B_{Gy,a})$ computed for each channel for all the training frames.

Then, the magnitudes of the gradient mean μ_G and standard deviation σ_G are computed in order to build the BEM. The orientation of the gradient (μ_θ and σ_θ) is also computed to avoid false edges created by illumination changes.

$$\mu^c_{G,a} = \sqrt{(\mu^c_{Gx,a})^2 + (\mu^c_{Gy,a})^2};\ \sigma^c_{G,a} = \sqrt{(\sigma^c_{Gx,a})^2 + (\sigma^c_{Gy,a})^2} \quad (6)$$

$$\mu^c_{\theta,a} = \arctan\left(\frac{\mu^c_{Gy,a}}{\mu^c_{Gx,a}}\right);\ \sigma^c_{\theta,a} = \arctan\left(\frac{\sigma^c_{Gy,a}}{\sigma^c_{Gx,a}}\right) \quad (7)$$

where $c \in \{R, G, B\}$ denotes the colour channel. Thus, $BEM =< \mu^c_{G,a}, \sigma^c_{G,a}, \mu^c_{\theta,a}, \sigma^c_{\theta,a} >$.

The thresholds employed for the segmentation task are automatically computed for each model, as shown in [6].

4.1.3 Image segmentation

The colour segmentation is achieved by following the same rules as [6]. However, edge segmentation is achieved based on the following premises:

1. Illumination changes modify the gradient magnitude but not the gradient orientation.
2. The gradient orientation is not feasible where there are no edges.
3. An edge can appear in a place where there were no edges before.

Assuming the first two premises, the gradient orientations will be compared instead of the gradient magnitudes for those pixels which have a minimum magnitude, in order to avoid the false edges due to illumination changes:

$$F_\theta = \left((\tau^c_{e,a} < V^c_{G,a,t}) \wedge (\tau^c_{e,a} < \mu^c_{G,a})\right) \wedge (\tau^c_{\theta,a} < |V^c_{\theta,a,t} - \mu^c_{\theta,a}|) \quad (8)$$

For those pixels satisfying the third premise, their gradient magnitudes are compared instead of their orientations:

$$F_G = \left(\neg\left((\tau^c_{e,a} < V^c_{G,a,t}) \wedge (\tau^c_{e,a} < \mu^c_{G,a})\right)\right) \wedge (\tau^c_{G,a} < |V^c_{G,a,t} - \mu^c_{G,a}|) \quad (9)$$

where $V^c_{\theta,a,t}$ and $V^c_{G,a,t}$ are the gradient orientation and magnitude for every pixel in the current image, respectively.

The use of the invariant models provides a high detection rate in comparison to other motion segmentation approaches. After the initial detection, moving foreground objects, chromatic shadows and some isolated pixels are contained in a binary mask named *M1*. Furthermore, the mask obtained using the gradient model is divided into two masks, which are used for the next steps. The *Edneg* mask corresponds to the foreground pixels belonging to the bg. model. While the *Edpos* mask corresponds to the foreground pixels belonging to the current image. A third mask is also created called *Edcom*, which contains the common edges detected in the bg. model and in the current image.

4.2. Shadow intensity reduction

In this step the *M1* mask from step 1 is reduced in order 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. Then, a new mask for this step is created according to the next equation:

$$M2_{a,t} = (I^R_{a,t} < \mu^R) \wedge (I^G_{a,t} < \mu^G) \wedge (I^B_{a,t} < \mu^B) \quad (10)$$

where *a* corresponds to the pixel location in the *M1* mask.

4.3. The bluish effect

The effect of illuminants which are different than white light provokes chromaticity changes because the changes in the intensity are different for every channel. In outdoor sequences the main illuminants are the sky and the sun (any of them white illuminant). The sky is the only source of illumination on shadowed regions, and the sky is assumed to be bluish as argued in [13]. Therefore, the intensity changes in the red and green channels are bigger than in the blue channel. This knowledge can be used to reduce the shadow region detected in the previous step (*M2*):

$$M2_{a,t} = (I^R_{a,t} - \mu^R) > (I^B_{a,t} - \mu^B) \wedge (I^G_{a,t} - \mu^G) > (I^B_{a,t} - \mu^B) \quad (11)$$

where *a* corresponds to the pixel location in the *M2* mask. Obviously, the bluish effect cannot be applied in indoor sequences.

4.4. Potential chromatic shadow regions

It is supposed that shadow regions have the same intensity change for each channel, since the illuminant is similar for all the shadowed region. However, different surfaces have different reflectance characteristics. Hence, the change in intensity depends on the surfaces material for the given shadow pixels. However, edges can show the changes between continuous pixels. Therefore, using the foreground edges detected in the current image, mask *Edpos*, the potential shadow regions can be separated from the moving foreground objects.

$$M3_{a,t} = M2_{a,t} \wedge (\neg Edpos_{a,t}) \quad (12)$$

A minimum area morphology is applied in order to avoid smaller regions which do not contain enough information for the subsequent steps of the shadow analysis.

4.5. Chromatic shadow texture detection

In this step the temporal textures of the regions detected in the previous mask M3 are analysed, in order to identify in which case of the theoretical shadow analysis (see section 3) each of the regions complies with. A region will be considered as a shadow if it complies with case 1. Negative foreground edges (Edneg mask) inside of the region are compared to the common foreground edges (Edcom mask), in order to prove if the region is a shadow and avoid the anomaly case 1-2. Furthermore, it also test if the negative edges are noise (larger regions have a higher probability to contain negative edges from noise):

$$Tx_b = \left(\frac{\sum_{a \in R_b} (R_b \wedge Edneg)}{|R_b \wedge Edtot|} \cdot k_n < \frac{\sum_{a \in R_b} (R_b \wedge Edcom)}{|R_b \wedge Edtot|} \right) \wedge \left(\frac{\sum_{a \in R_b} (R_b \wedge Edneg)}{|R_b|} < k_s \right) \quad (13)$$

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, which is equal to the probability of the region belongs to a shadow or a foreground object; and k_s is used to measure if the negative edges corresponds to noise. Again larger regions have a higher probability to contain negative edges from noise.

4.6. Chromatic shadow angle and brightness detection

In this step the 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 as a shadow if it complies with case 2. The only regions analysed in this section will be those that does not have textures, neither in the background model nor in the current image. If the pixels do not have texture, nor similar chrominance angle distortion and do not have a significant brightness distortion, then the region will be classified as shadow.

$$ABd_b = \left(\frac{\sum_{a \in R_b} (R_b \wedge Edtot)}{|R_b|} < k_t \right) \wedge (\sigma(R_b \wedge \breve{\alpha}) < k_a) \wedge \left(\sigma(R_b \wedge \breve{\beta}) < k_b \right) \quad (14)$$

where σ is the standard deviation of $\breve{\alpha}$ and $\breve{\beta}$ which are the chrominance angle and brightness normalised distortions calculated for each pixel in the region number b (R_b), respectively; k_t is a confidence region to avoid noise textures; k_a and k_b is a minimum threshold used to determine if the angle and brightness distortion are similar amoung the pixels of the evaluated region.

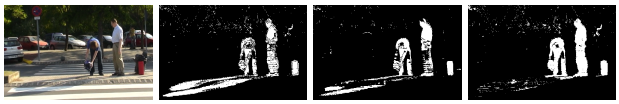

Figure 4. An original image from the Outdoor_Cam1 sequence, and foreground results after shadow removal using the Huerta et al. approach [6], the zivkovic et al. approach [21] using a shadow detector [16], and our approach, respectively.

4.7. Chromatic shadow edge removal

Pixels from the potential shadow regions, which were neglected in section 4.4 because they were part of the *Edpos* mask, have to be included again in the regions destected as shadow.

4.8. Shadow position verification

A moving cast shadow is always caused by a moving foreground object. Therefore, in this section it is tested if a detected shadow has an associated fg. 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 chrominance angles between the detected shadows and the fg. objects are calculated. After, the most probable chrominance angle obtained in the training period is used to discard detected shadows, which do not have any foreground object in the direction of the chrominance angle.

5. Experimental Results and Discussion

The results presented in this section are all 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 chosen test sequences are relatively long and umbra and penumbra shadows are cast by multiple foreground objects. The sequences analysed are Outdoor_Cam1 (800 frames, 607x387 PX), HigwayIII (2227 frames, 320x240 PX), and HallwayI[1] (1800 frames, 320x240 PX).

Figure 4, 5 and 6 show the results when comparing our approach with other approaches from the state-of-the-art [6, 8, 21, 16, 10]. As it can be seen in these figures our approach outperforms the other analysed methods. However, in a few cases the gradient masks cannot be accurately build due to camouflage and noise problems. Thus, the separation of a foreground object and a shadow region can fail. Occasionally, when the anomaly in case 2-2 occurs and a part of the foreground object or the shadow is not segmented due to segmentation problems, the shadow position can misclassify the shadow as a foreground object.

To evaluate our approach in a quantitative way, it is compared with the approaches [9, 10] using the most employed

[1] http://vision.gel.ulaval.ca/ CastShadows/

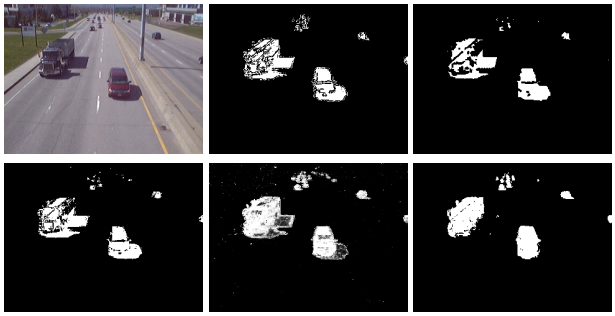

Figure 5. An original image from the HighwayIII sequence, and foreground results after shadow removal using the Huerta et al. approach [6], the Kim et al. approach [8], the zivkovic et al. approach [21] using a shadow detector [16], the Martel et al. approach [10], and our approach, respectively (read row-wise).

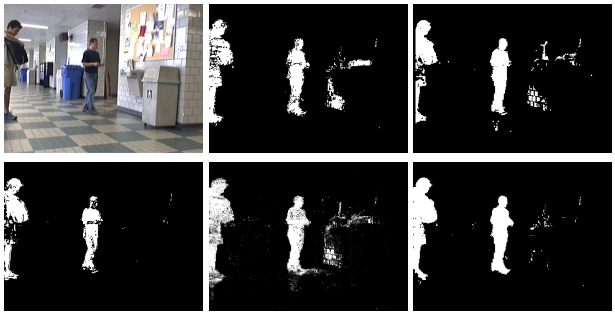

Figure 6. An original image from the HallwayI sequence, and foreground results after shadow removal using the Huerta et al. approach [6], the Kim et al. approach [8], the zivkovic et al. approach [21] using a shadow detector [16], the Martel et al. approach [10], and our approach, respectively (read row-wise).

Method	HallwayI	
	SR	SD
GMSM	0.605	0.870
Physical model	0.724	0.867
Our approach	0.807	0.907

Table 1. SR and SD results for our approach and two of the most successful methods: Gaussian Mixture Shadow Models (GMSM) [9] and a physical model of light and surfaces [10].

quantitative expressions utilized to evaluate the shadow detection performance: the Shadow Detection Rate (SR) and the Shadow Discriminate Rate (DR), refer to [16] for the exact equations. Results show that our method outperforms both the parametric approach based on Gaussian mixtures GMSM [9] and the nonparametric physical model [10]. Note that the results for the GMSM [9] and the physical model [10] on the sequence Hallway have been obtained directly from [10].

It should be noted that our approach needs a reasonable resolution to work correctly. Futhermore, shadow regions need to have a minimum area for analysis or there might

not be enough information for a proper shadow detection.

6. Conclusions

In this paper, we have presented an approach for detection and removal of chromatic moving shadows in surveillance scenarios. The approach apply a novel technique based on gradient and colour models for separating chromatic moving shadows from detected moving objects. We extend and improve well-known colour and gradient models to an invariant colour cone model and an invariant gradient model, respectively. Furthermore, we combine colour, gradient and textural information, together with known shadow properties to improve the shadow detection. The resulting shadow detection can detect and remove chromatic moving shadows (umbra shadows) and penumbra shadows, while several other methods are restricted to the latter. Qualitative and quantitative results of tests for both outdoor and indoor sequences from well-known databases validate the presented approach. Overall, our approach gives a more robust and accurate shadow detection and foreground segmentation compared to state-of-the-art methods.

In future work, edge-linking or B-spline techniques can be used to avoid the partial loss of foreground borders due to camouflage, and thereby improve the edge model. Another interesting aspect is to use the direction of penumbra to umbra for a cast shadow as discrimination between foregrounds and shadows, which does not have texture nor similar temporal and spatial chrominance angle and brightness distortions. Finally, high-level information such as tracking information for a detected shadow can be applied to enhance the detection process in subsequent frames.

Acknowledgments

This work is supported by EC grants IST-027110 for the HERMES project and IST-045547 for the VIDI-Video project, and by the Spanish MEC under projects TIN2006-14606 and CONSOLIDER-INGENIO 2010 MIPRCV CSD2007-00018, and by he Big Brother project (Danish National Research Councils - FTP).

References

[1] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati. Detecting moving objects, ghosts, and shadows in video streams. *IEEE TPAMI*, 25(10):1337–1342, October 2003. 2

[2] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirotti. Improving shadow suppression in moving object detection with hsv color information. In C. A.F., editor, *Proc. IEEE ITSC*, page 334339, Oakland, CA, USA, 2001. 2

[3] G. Finlayson, S. Hordley, C. Lu, and M. Drew. On the removal of shadows from images. *IEEE TPAMI*, 28(1):59–68, January 2006. 2

[4] M. Heikkila and M. Pietikainen. A texture-based method for modeling the background and detecting moving objects. *IEEE TPAMI*, 28(4):657–662, 2006. 2

[5] T. Horprasert, D.Harwood, and L.S.Davis. A statistical approach for real-time robust background subtraction and shadow detection. In *IEEE Frame-Rate Applications Workshop*, Kerkyra, Greece, 1999. 2, 4

[6] I. Huerta, A. Amato, J. Gonzlez, and J. Villanueva. Fusing edge cues to handle colour problems in image segmentation. In *Proc. AMDO'08*, volume 5098, pages 279–288, Andratx, Mallorca, Spain, 2008. Springer LNCS. 4, 5, 7

[7] Y. Ivanov, A. Bobick, and J. Liu. Fast lighting independent background subtraction. *IJCV*, 37(2):199–207, June 2000. 2

[8] K. Kim, T. Chalidabhongse, D. Harwood, and L. Davis. Real-time foreground-background segmentation using codebook model. *Real-Time Imaging*, 11(3):172–185, June 2005. 2, 4, 7

[9] N. Martel-Brisson and A. Zaccarin. Learning and removing cast shadows through a multidistribution approach. *IEEE TPAMI*, 29(7):1133–1146, 2007. 7

[10] N. Martel-Brisson and A. Zaccarin. Kernel-based learning of cast shadows from a physical model of light sources and surfaces for low-level segmentation. In *IEEE CVPR'08*, pages 1–8, June 2008. 2, 4, 7

[11] Y. Matsushita, K. Nishino, K. Ikeuchi, and M. Sakauchi. Illumination normalization with time-dependent intrinsic images for video surveillance. *IEEE TPAMI*, 26(10):1336–1347, 2004. 2

[12] T. B. Moeslund, A. Hilton, and V. Kruger. A survey of advances in vision-based human motion capture and analysis. *CVIU*, 104:90–126, November-December 2006. 1

[13] S. Nadimi and B. Bhanu. Physical models for moving shadow and object detection in video. *IEEE TPAMI*, 26(8):1079–1087, August 2004. 2, 3, 6

[14] R. O'Callaghan and T. Haga. Robust change-detection by normalised gradient-correlation. In *IEEE CVPR'07*, pages 1–8, June 2007. 4

[15] K. Onoguchi. Shadow elimination method for moving object detection. In *ICPR*, volume 1, page 583587, 1998. 2

[16] A. Prati, I. Mikic, M. Trivedi, and R. Cucchiara. Detecting moving shadows: Algorithms and evaluation. *IEEE TPAMI*, 25(7):918923, July 2003. 2, 7

[17] E. Salvador, A. Cavallaro, and T. Ebrahimi. Cast shadow segmentation using invariant color features. *CVIU*, 95(2):238–259, August 2004. 2

[18] O. Schreer, I. Feldmann, U. Goelz, and P. Kauff. Fast and robust shadow detection in videoconference applications. In *Proc. IEEE VIPromCom*, pages 371–375, 2002. 2

[19] J. Stauder, R. Mech, and J. Ostermann. Detection of moving cast shadows for object segmentation. *IEEE Trans. Multimedia*, 1(1):65–76, March 1999. 2

[20] Y. Weiss. Deriving intrinsic images from image sequences. In *Proc. ICCV'01*, volume 02, pages 68–75, Vancouver, Canada, 2001. 2

[21] Z. Zivkovic and F. Heijden. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recognition Letters*, 27(7):773–780, 2006. 7