Chapter 1
Moving Cast Shadows Detection Methods for Video Surveillance Applications

Ariel Amato, Ivan Huerta, Mikhail G. Mozerov, F. Xavier Roca and Jordi Gonzàlez

Abstract Moving cast shadows are a major concern in today's performance from broad range of many vision-based surveillance applications because they highly difficult the object classification task. Several shadow detection methods have been reported in the literature during the last years. They are mainly divided into two domains. One usually works with static images, whereas the second one uses image sequences, namely video content. In spite of the fact that both cases can be analogously analyzed, there is a difference in the application field. The first case, shadow detection methods can be exploited in order to obtain additional geometric and semantic cues about shape and position of its casting object ('shape from shadows') as well as the localization of the light source. While in the second one, the main purpose is usually change detection, scene matching or surveillance (usually in a background subtraction context). Shadows can in fact modify in a negative way the shape and color of the target object and therefore affect the performance of scene analysis and interpretation in many applications. This chapter wills mainly reviews shadow detection methods as well as their taxonomies related with the second case, thus aiming at those shadows which are associated with moving objects (moving shadows).

———————————

Ariel Amato
Computer Vision Center, Universitat Autònoma de Barcelona, e-mail: `aamato@cvc.uab.es`

Ivan Huerta
Institut de Robòtica i Informàtica industrial, Universitat Politècniva de Catalunya, e-mail: `ihuerta@iri.upc.edu`

Mikhail G. Mozerov
Computer Vision Center, Universitat Autònoma de Barcelona, e-mail: `mozerov@cvc.uab.es`

F. Xavier Roca
Computer Vision Center, Universitat Autònoma de Barcelona, e-mail: `xavir@cvc.uab.es`

Jordi Gonzàlez
Computer Vision Center, Universitat Autònoma de Barcelona, e-mail: `poal@cvc.uab.es`
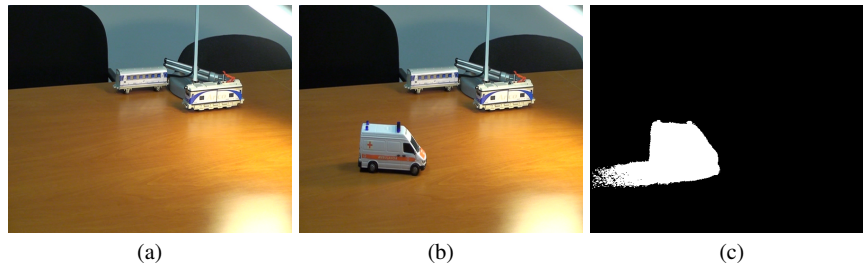
(a)   (b)   (c)

Fig. 1.1 Motion Segmentation: (a) background Image; (b) current image; and (c) Segmented Image.

1.1 Introduction

Video Surveillance has been in our society for a long time [6, 13]. It began in the twentieth century to assist prison officials in the discovery of escape methods. However, it was not until the late-twentieth century that surveillance expanded to include the security of property and people. Video surveillance is more prevalent in Europe than anywhere in the world. For instance, in the past decade, successive UK governments have installed over 2.4 million surveillance cameras (about one for every 14 people). [1] The average Londoners are estimated to have their picture recorded more than three hundred times a day [2]. Traditionally video surveillance was used to display images on monitors inspected by guards or operators. This fact has allowed the observation of an increase number of places using less people and also to perform patrolling duties from the safety of a control room. However, a single operator can only monitor a limited amount of scenes simultaneously and for a limited amount of time, because the process of manual surveillance is very time-consuming and is a really tedious task.

The new breakthroughs in technology have led to a new generation of video surveillance. The current generation of video surveillance systems uses digital computing and communication technologies to improve the design of the original architecture, with the ultimate goal to create an automatic video surveillance system.

Recent trends in computer vision has delved into the study of cognitive vision systems, which uses visual information to facilitate a series of tasks on sensing, understanding, reaction and communication. In other words, video surveillance systems aim to automatically identify people, objects or events of interest in different kinds of environments. Although video surveillance is probably one of the most popular areas for research in the field of computer vision, and much effort has been made to achieve an automatic system, this goal has yet to be reached.

Nowadays, the task of a video surveillance system aims to provide support to the human operator. The system warns an operator when an event, e.g., possible risks or potential dangerous situations, is detected. Despite the fact that the long-term goal

---

[1] http://news.bbc.co.uk/2/hi/uk_news/6108496.stm

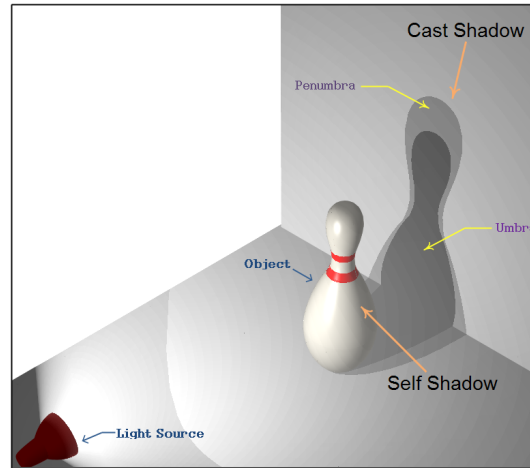[2] http://epic.org/privacy/surveillance/

Fig. 1.2 Shadows Types: Self and Cast (Umbra and Penumbra).

is to build a completely automated systems, the short-term one is to increase the robustness of the current systems in order to reduce false alarms. This goal can only be achieved if the systems are able to interpret the interaction of events in the scene. To reach this goal low and high level tasks must be performed. A method that is able to perform the low level tasks, namely detection, localization and tracking with high accuracy, can highly benefit the process of scene understanding.

In video surveillance moving object detection plays an important role [15, 43]. Along decades, different methods have been developed to extract moving region in the scene. However, the most common, simple and effective approach to moving object segmentation is Background Subtraction, where a stationary camera is used to observe dynamic events in a scene.

In moving object detection algorithms, moving cast shadows have a high probability to be misclassified as moving objects (foregrounds). Such an error is due to the fact that a moving object and its moving shadow share similar motional characteristics. An example of motion segmentation image based on background subtraction process is shown in Fig. 1.1(c). The segmented image shows that the shadow was also segmented as a part of the object (foreground).

A shadow is a photometric phenomenon that occurs when an object partially or totally blocks the direct light source. Shadows can take any size and shape. In general, shadows can be divided into two major classes: self and cast shadows. A self shadow occurs in the portion of an object that is not illuminated by direct light. Cast shadows are the areas projected on a surface in the direction of direct light. Cast shadows can be further classified into umbra and penumbra. The region where the direct light source is totally blocked is called the umbra, while the region where it is partially blocked is known as the penumbra. These definitions are visually represented in Fig. 1.2.
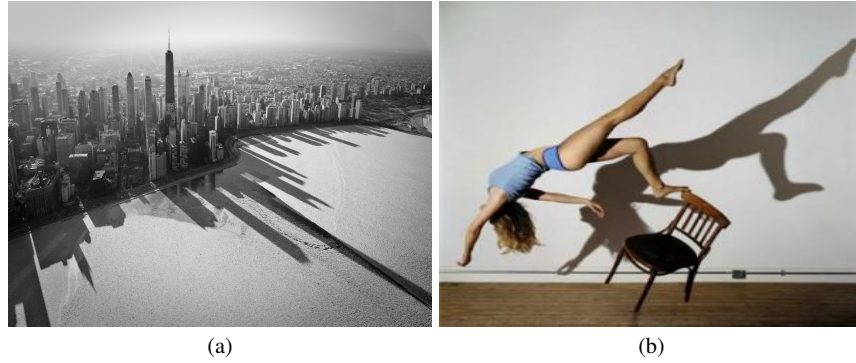
(a)   (b)

Fig. 1.3 Shadows from static and moving objects. (a) Shadows casted by static objects. (b) Shadow casted by a moving object.

Shadows in images are generally divided into static and dynamic shadows (see Fig. 1.3). Static shadows are shadows due to statics objects such as building, parked cars, trees, etc. Moving object detection methods do not suffer from static shadows since static shadows are modeled as a part of background. In contrary, dynamic (moving) shadows, the subject of interest in this chapter, are harmful for moving object detection methods. Shadows can be either in contact with the moving object, or disconnected from it (see Fig. 1.4). In the first case, shadows distort the object shape, making the use of subsequent shape recognition methods less reliable. In the second case, the shadows may be wrongly classified as an object in the scene. Typical problems caused by moving shadows in surveillance scenarios are shown in Fig. 1.5. In Fig. 1.5-(I), a traffic surveillance scene, shadows cause merging of multiple objects; in Fig. 1.5(II), an indoor scenario, shadows are projected on the floor and on the wall. In this case a false positive foreground (shadow casted on the wall) occurs; and in Fig. 1.5(III), a long shadow causes a severe object shape distortion in an outdoor scenario. Clearly, in many image analysis applications, the existence of moving cast shadows may lead to an inaccurate object segmentation. Consequently, tasks such as object description and tracking are severely affected, thus inducing an erroneous scene analysis.

This chapter is organized in the following Sections. Section 1.2 some relevant background subtraction techniques are revised. Reported taxonomies on moving cast shadow methods are described in Section 1.3. Section 1.4 gives a literature review on moving cast shadow detection methods. A discussion of open issues and main difficulties in the area of moving cast shadow detection is presented in Section 1.5. Section 1.6 gives an analysis of the tools for the performance evaluation on moving cast shadow detection algorithms. Finally, section 1.7 briefly reviews the topics discussed in the different sections of this chapter and establishes the final concluding remarks of this work.
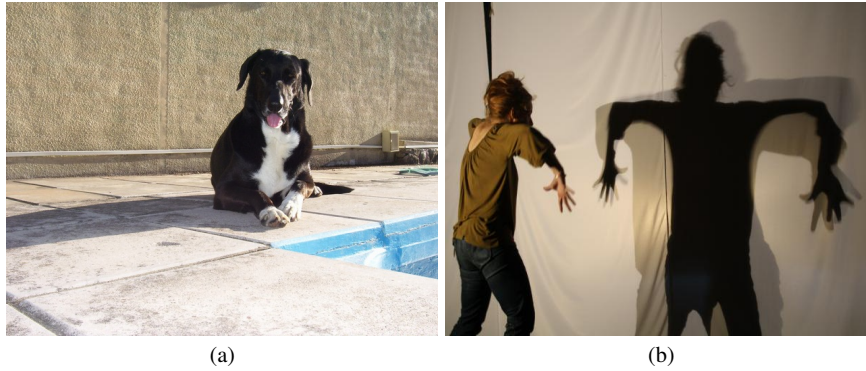
(a)   (b)

Fig. 1.4 Shadows location: (a) shadow is spatially connected to the object; (b) shadow is spatially unconnected to the object.

1.2 Methods for Moving Region Extraction

Detecting regions that correspond to moving objects such as vehicles and people in natural scenes is a significant and difficult problem for many vision-based applications. The extraction of the moving region is the first step to locate where a moving shadow can be detected.

The most used techniques for motion segmentation are: *(i)* background subtraction, *(ii)* frame differencing, (a combination of both), or *(iii)* optical flow. Even though many algorithms have been proposed in the literature [17, 65, 41], the problem of identifying moving objects in complex environment is still far from being completely solved.

Motion segmentation based on optical flow [40, 5] uses characteristics of flow vectors of moving objects over time to detect change regions in an image sequence. These methods can segment moving objects in video sequences even from a moving camera. However, most of these methods are computationally highly expensive and very sensitive to noise.

Temporal differencing technique attempts to extract moving regions by making use of a pixel-by-pixel difference between consecutive frames in a video sequence [55, 57]. It is very adaptive to dynamic scene changes. Nevertheless, it generally fails to extract the entire relevant pixels of moving objects.

Background subtraction is the most commonly used technique for motion segmentation in static scenes [38, 45, 28, 4, 53].

Basically, the methodology behind any background subtraction technique consists in subtracting a model of the static scene 'background' from each frame of a video sequence. (see Fig. 1.6). In general, a background subtraction technique can be divided into three phases: first, the generation of a suitable reference model, normally called background (training phase); second, the measurement procedure or classification (running phase) and finally; the model maintenance (updating phase). For each of these phases, particular challenging exist [25].
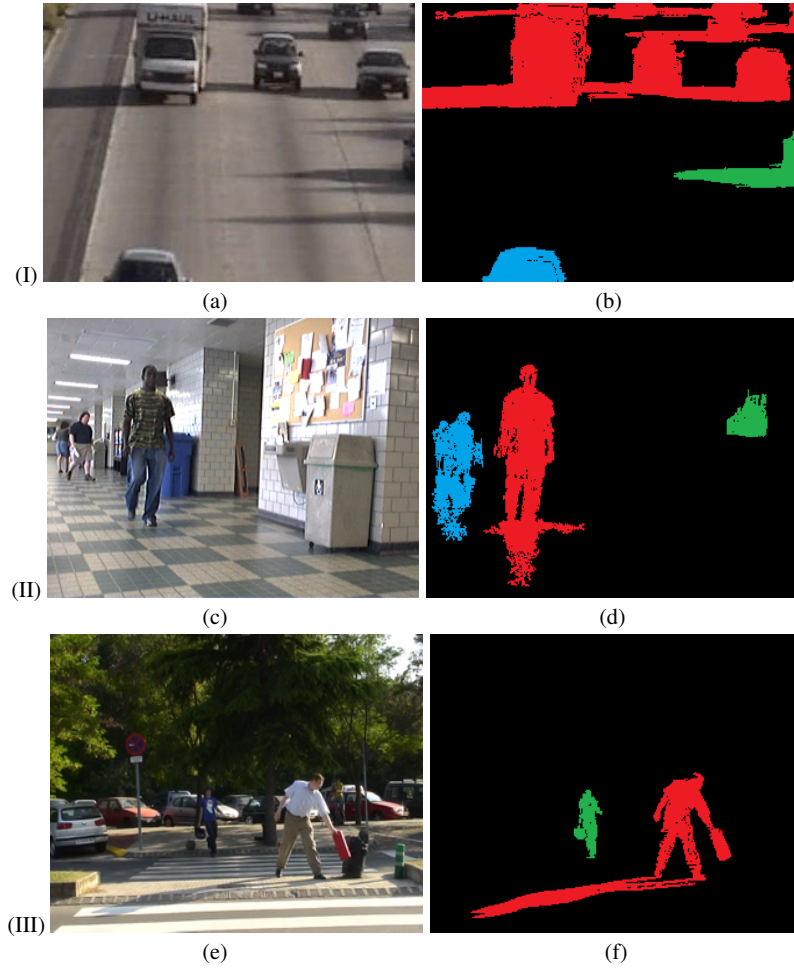
Fig. 1.5 Negative effect of shadow in surveillance scenarios.

There are a large number of different algorithms using this background subtraction scheme. Nonetheless, they differ in: *(i)* the type of cues or structures employed to build the background representation; *(ii)* the procedure used for detecting the foreground region; and *(iii)* the updating criteria of the background model.

A naive version of the background subtraction scheme is employed by Heikkila and Silven [20], which classifies an input pixel as foreground if its value is over a predefined threshold when subtracted from the background model. This approach updates the background model in order to guarantee reliable motion detection using a first order recursive filter. However, this method is extremely sensitive to changes of dynamic scenes such as gradual illumination variation or physical changes such as *ghosts* (i.e., when an object already represented in the background model begins
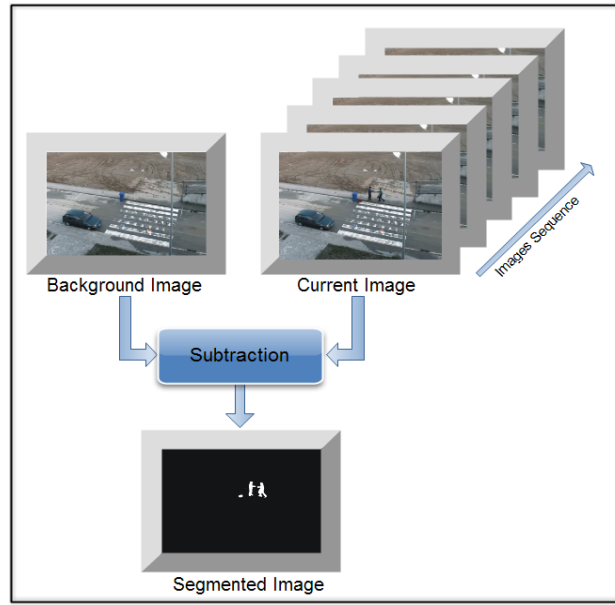
Fig. 1.6 Background Subtraction representation.

to move). In order to overcome these difficulties, statistical approaches have been applied [64]. These approaches make use of statistical properties of each pixel (or regions), which are updated dynamically during all the process in order to construct the background model.

Haritaoglu et al. in W^4 [19] apply background subtraction by computing for each pixel in the background model, during a training period, three values: its minimum and maximum intensity values, and the maximum intensity difference between consecutive frames. Background model pixels are updated using pixel-based and object-based updating conditions to be adaptive to illumination and physical changes in the scene. However, this approach is rather sensitive to shadows and lighting changes, since the only cue used is intensity.

Alternatively, Wren et al. in Pfinder [67] proposed a framework in which each pixel's value (in YUV space) is represented with a single Gaussian. Then, model parameters are recursively updated. However, a single Gaussian model cannot handle multiple backgrounds, such as waving trees.

Stauffer and Grimson [60, 59] addressed this issue by using a Mixture of Gaussians (MoG) to build a background color model for every pixel.

An improvement of the MoG can be found in Zivkovic et al. [73, 74], where the parameters of a MoG model are constantly updated, while selecting simultaneously the appropriate number of components for each pixel.

Elgammal et al. [14] use a non-parametric Kernel Density Estimation (KDE) to model the background. Their representation samples an intensity values for each pixel to estimate the probability of newly observed intensity values. The back-

ground model is also continuously updated to be adaptive to background changes. In addition to color-based information, their system incorporates region-based scene knowledge for matching nearby pixel locations. This approach can successfully handle the problem of small background motion such as tree branches.

Mittal et al. [40] use adaptive KDE for modeling background in motion, and implement optical flow to detect moving regions. In this way, their approach is able to manage complex background; however, the computational cost of this approach is quite high. Chen et al. [9] combine pixel- and block-based approaches to model complex background. Nevertheless, the method is very sensitive to camouflages and shadows.

Cheng et al. in [10] propose an on-line learning method which is able to work in real-time and can be implemented in GPU, which also gives similar results managing complex background. In [3] Barnich and Droogenbroeck also present a really fast method that can cope with background in motion and bootstrapping problems. The method adopts the idea of sampling the spatial neighborhood for refining the per-pixel estimation. The model updating relies on a random process that substitutes old pixel values with new ones. However, it cannot cope with camouflages and shadows. Another solution to bootstrapping problem is presented by Colombari et al. in [11], where a patch-based technique exploits both spatial and temporal consistency of the static background.

Li et al. [31] and Sheikh et al. [54] use Bayesian networks to cope with dynamic backgrounds. Li et al. uses a Bayesian framework that incorporates spectral, spatial, and temporal features to characterize background appearance. Sheik et al. apply non-parametric density estimation to model the background as a single distribution, thus handling multi-modal spatial uncertainties. Furthermore, they also use temporal information.

The use of layers for image decomposition based on the neighboring pixels is presented in [44]. They state that such approach is robust and efficient to handle dynamic backgrounds. Maddalena et al. [34] use neural networks to overcome the same problem. An improvement of it, using self organizing maps, can be found by Lopez-Rubio et al. [33], which can adapt its color similarity measure to the characteristics of the input video.

Mahadevan et al. in [35] uses a combination of the discriminant center-surround saliency framework with the modeling power of dynamic textures to solve problems with highly dynamic backgrounds and a moving camera. However, this method is not designed for high accurate segmentation.

Toyama et al. [62] in Wallflower use a three-component system to handle many canonical anomalies for background updating. Their work processes input images at various spatial scales, namely pixel, region, and frame levels. Reasonably good foreground detection can be achieved when moving objects or strong illumination changes (for example when turning on/off the light in an indoor scene) are present. However, it fails when modeling small motion in the background or local illumination variations.

Edge cues are also used for motion segmentation. Weiss [66] also extract intrinsic images using edge cues instead of color to obtain the reflectance image. This

process requires several frames to determine the reflectance edges of the scene.
A reflectance edge is an edge that persists throughout the sequence. Given reflectance edges, the approach re-integrates the information to derive a reflectance image. However, the reflectance image also contains scene illuminations because this approach requires prominent changes in the scene, specifically for the position of shadows.

Jabri et al. [26] use a statistical background modeling which combines color (in RGB space) with edges. The background model is computed in two distinct parts: the color model and the edge model. On the one hand, a color model is represented by two images, the mean and the standard deviation images. On the other hand, an edge model is built by applying the Sobel edge operator to each color channel, thereby yielding horizontal and vertical difference images. Subsequently, background subtraction is performed by subtracting the color and edge channels separately using confidence maps, and then combining the results to get the foreground pixels.

Javed et al. [27] present a method that uses multiple cues, based on color and gradient information. The approach tries to handle different difficulties, such as bootstrapping (initialization with moving objects), repositioning of static background objects, ghost and quick illumination changes using three distinct levels: pixel, region and frame level, inspired from [62].

At the pixel level, two statistical models of gradients and color based on mixture of Gaussians are separately used to classify each pixel as background or foreground. At the region level, foreground pixels obtained from the color model are grouped into regions, and the gradient model is then used to eliminate regions corresponding to highlights or ghosts.

Pixel-based models are updated based on decisions made at the region level. Lastly, the frame level ignores the color based subtraction results if more than 50 percent of the results are considered foreground, thereby using only gradient subtraction results to handle global illumination changes. Nevertheless, *ghosts* cannot be eliminated if the background contains a high number of edges.

Some of the aforementioned motion detection approaches generally obtain good segmentation in indoor and outdoor scenarios, thus some of them have been used in real-time surveillance applications for years. However, their performance is highly affected by the moving shadows.

1.3 Taxonomies of Moving Cast Shadow Detection Methods

Moving cast shadow detection algorithms are mainly based on the use of shadow descriptors. They basically model shadows by using properties such as: chromaticity invariant, textural patterns, photometric physical models, or even by analyzing the projected areas in term of size, shape and direction.

The methodology of moving cast shadow detection can further includes geometrical-shadow-information or spatial-shadow-cues as well as a training-shadow-stage, or

a sort of combination of them. In turn, the methods can perform at different levels, considering only the information of a single pixel, using a set of pixels, or even performing with the information of the whole frame.

Diverse information that characterizes moving shadows is exploited and in many cases such information is combined or used in a different way. This fact makes very difficult to classify in a unique manner the moving cast shadow methods.

The main reported taxonomies in the literature were proposed by:

Prati et al. [47], present two layers taxonomy *(algorithm-based taxonomy)*. The first layer classification considers whether the decision process introduces and exploits uncertainty. Deterministic approaches use an on/off decision process, whereas statistical approaches use probabilistic functions to describe the class membership. In turn, both layers are further divided. For statistical approaches the authors include parametric and non-parametric separation. In the case of deterministic methods, algorithms are classified by whether the decision is supported by model-based knowledge or not. Additionally, spectral, spatial and temporal information are also considered.

Salvador et al.[50] propose to divide shadow detection methods in *(i)* model-based methods and *(ii)* property-based methods. Model-based approaches work with models that represent a priori knowledge of the geometry of the scene, the object and the illumination. While property-based methods identify shadows by using properties such as the geometry, brightness and/or color of shadows.

Zhang et al.[72] describe moving cast shadow detection methods into: *(i)* color/spectrum-based methods; *(ii)* texture-based methods, and *(iii)* geometry-based methods. The color/spectrum-based methods attempt to describe the color change of a shadowed pixel and find the color feature that is illumination invariant. Texture-based methods consider that the texture of foreground object is different to the texture of background; while the texture of a shadowed area must be the same to the texture of background. Finally the geometric-based methods are focused on the characteristic of the casted shadow area. Usually the characteristics to be analyzed are direction, size and shape of the shadow. Often these methods can hardly be able to avoid the use of some prior knowledge of the scene. In turn, the authors also describe methods that make use of statistical inference of shadow models.

Ullab et al.[63] state that moving shadow removal methods can be partitioned into three categories: *(i)* intensity information, *(ii)* photometric invariant information and *(iii)* color and statistical information. The first classification concentrates in the brightness of the shadowed pixels. Typically a shadowed pixel decreases its brightness compared to the same pixel without shadow. The second classification includes those algorithms that exploit photometric-invariant-shadow property. Normally such photometric invariability can be obtained in normalized color spaces that can separately operate with the brightens and the chroma of the pixels. The last classification stands for methods which usually classify shadow by using statistical model of the pixel's information.

Sanin et al.[52] separate moving cast shadow removal methods into: *(i)* chromaticity-based methods; *(ii)* physical methods; *(iii)* geometrical-based methods and *(iv)* texture-based methods. Additionally, a secondary classification within each cat-

egory is proposed.Chromacity-based methods are divided according their color space, level of granularity and additional spatial or temporal verification. Physical methods are divided according to their physical shadow model, learning algorithm and additional spatial or temporal cues. Geometry-based methods are divided according to their supported object type, whether they support multiple objects per blob, their main geometrical cue and additional cues. Texture-based methods are divided according to their weak shadow detector, texture correlation method and the size of the regions used in the correlation.

1.4 Methods for Moving Cast Shadow Detection

In this section the most classical and well known moving cast shadow detection approaches are presented.

Although, different taxonomies were revised in the previous section, we have organized this methods' review in pixel-level (those methods that perform with the information of a single pixel), region level (those methods that make use of a set of pixels) and frame-level[3] in order to unify as much as possible the different methods' characteristics.

Pixel-Level

Many shadow detection methods assume that a shadowed pixel becomes darker but with a similar chromaticity that the same pixel without shadow. Chromaticity is a measurement of color that is independent of intensity component. The invariability in chroma, between a (non-shadowed) pixel belonging to the background and the same (shadowed) pixel belonging to the current image, together with a brightness decrement, represent a distinctive shadow feature. Often methods that are using this shadow descriptor perform in color spaces where the distinction between brightness and chroma is supported. These common spaces are: HSV, HSI, YUV, C1C2C3, normalized RGB, etc (see Fig.1.7).

For example, Cucchiara et al. [12] use shadow properties in the HSV color space to distinguish shadows from moving objects. These properties show that cast shadows darken the background in the luminance component, while the hue and saturation components change within certain limits.

Horprasert et al. [21] propose a color model that compares intensity to the chromaticity component at each pixel. Each pixel is classified as background, shaded, highlighted or moving foreground through a combination of three threshold values, which are defined over a single Gaussian distribution. An extension of this work based on multiple background pixels organized in a codebook is done by Kim et al. [29].

---

[3] Frame-level methods are included, despite the fact that they are not widely used, in order to obtain a thorough review of the methods.
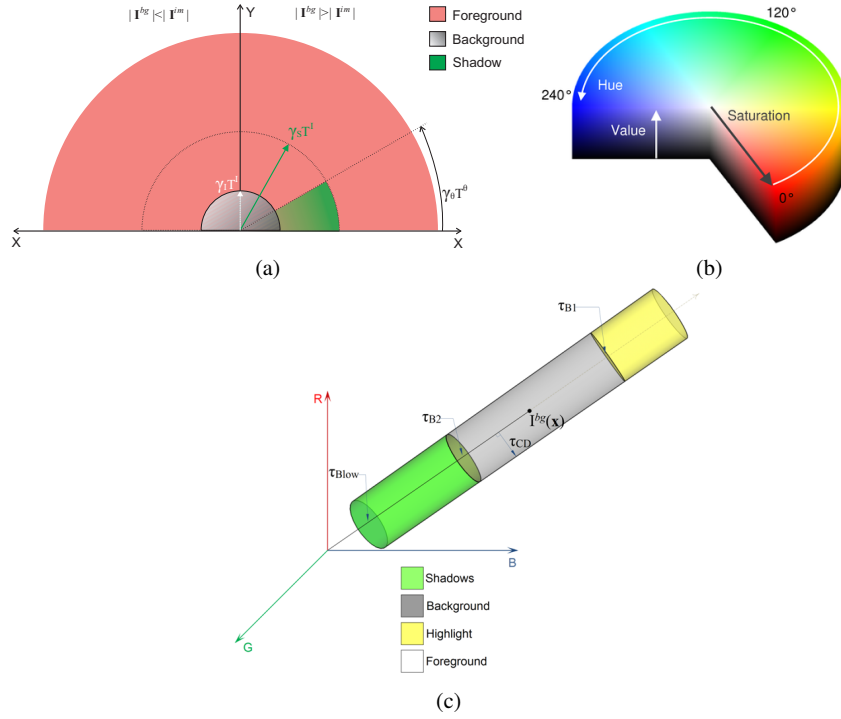
Fig. 1.7 Chromaticity invariant shadow descriptors in different color spaces. (a) Amato et al.[1], (b) Cucchiara et al. [12] and (c) Horprasert et al. [21].

McKenna et al. [39] assume that cast shadows result in significant change in intensity without much change in chromaticity. Pixel's chromaticity is modeled using its mean and variance. In turn, the first-order gradient of each background pixel is also exploited. Moving shadows are then classified as background if the chromaticity or gradient information supports their classification as such.

The advantage of all the mentioned above methods reside in that they are fast (suitable for real-time applications), and easy to implement. However, they are specially restricted to achromatic shadows. Moreover, some of them often require explicit tuning of parameters for each scene.

There are some methods that aim to obtain an inference of the pixel values in the shadowed areas by using some photometric physical model. For such a purpose, a formulation can be achieved by exploiting: a reflectance model, an illumination model or an adaptation of classical color models. In order to obtain the appearance of the shadowed pixels some methods may need a training phase (it could be supervised or unsupervised), or/and some prior knowledge of the scene of interest.

Photometric physical model methods that implement statistical learning-based methodology have been developed to learn and remove cast shadows [36, 46, 37, 23]. For example, in the work of [37] a nonparametric framework to model surface
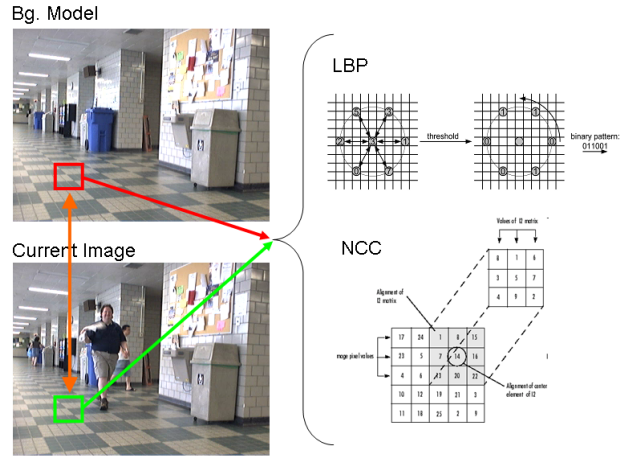
Fig. 1.8 Texture-oriented methodology.

behavior when shadows are cast on them is introduced. Physical properties of light sources and surfaces are employed in order to identify a direction in RGB space at which background surface values under cast shadows are found. However, these approaches are particularly affected by the training phase. These methods require a long training period.

In the work of Siala et al. [56] a statistical non-parametric shadow detection method is presented. First, in the learning phase, an image containing foreground, background and moving shadow is selected. The moving shadow regions are manually annotated. The information obtained from this annotation is then used to create a diagonal model that describes the shadow appearance in the RGB ratio color space. The shadow detection is obtained by performing a one class classification based on a support vector domain description (SVDD).

Region-Level

Although most of the methods that perform with a set of pixels typically make use of texture information, there are few methods that exploit other shadow descriptor such as: chormaticity invariant or photometric physical models.

Methods that use texture as shadow descriptor basically are based on the idea that a shadow is a semi-transparent region in the image. Thus, they assume that a strong correlation between two regions, one affected by shadow and the same region without the shadows effect, must exist. These methods try to obtain such a corre-

lation using for example: local binary patterns (LBP), normalized cross-correlation (NCC), color cross covariant (CCC), Markov random field, etc. (see Fig. 1.8). For instance, Grest et al.[18] propose to tackle the detection of moving cast shadows using two similarity measurements, one is based on the normalized cross correlation (NCC) and the other is the color cross covariant (CCC). Basically the authors are interested in comparing pixel values at the same position in two images, (the current image and a reference image) and then inferring if there is a correlation between the information of these pixels. The computation of these measurements are done over a given window size. The NCC is calculated using the brightness of the pixel, while the CCC is obtained in the biconic HSL (Hue, Saturation and Lightness) color space. The authors assume that: *(i)* a shadowed pixel is darker than the corresponding pixel in the background image; *(ii)* the texture of the shadowed region is correlated with the corresponding texture of the background image. Despite the fact that CCC is used to solve the limitation of the method to distinguish shadow from object over homogeneous areas, still the success of the approach under shadows-camouflage areas is far to be achieved.

Other approach based on NCC is proposed by Yuan et al. [71]. The authors proposed to include a multi-frame differencing strategy to improve the segmentation in those cases where the shadows cannot successfully be removed. This strategy is based on that shadowed regions differ a little in two consecutive frames. Therefore the biggest part of the shadows can be eliminated by frame difference, but only remain some shadow edges. These shadow edges are removed by using a new frame differencing step.

Jacques et al.[7] propose to detect shadows regions by using intensity measurement of a set of pixels. This measurement are computed by ratio-pixels (image/background) in a fixed 3×3 windows and the decision is based on a statistical non-parametric inference.

In the work of Yao et al. [70], textures are computed using the LBP combined with a RGB color model. The authors state that LBP can work robustly to detect moving shadows on rich texture regions. However, it fails when both the background image and the foreground objects share the same texture information. Therefore, to handle these situations, in this work the authors make use of a shadow invariant color distance in the RGB color space. They claims that pixel values changed due to shadows are mostly distributed along in the axis going toward the RGB origin point. Thus, they propose to compare the color difference between an observed color pixel and a background color pixel using their relative angle in RGB color space with respect to the origin and the changing range of the background color pixel up to last time instant.

Leone et al. [30] use a textural shadow descriptor by projecting the neighborhood of pixels onto a set of Gabor functions, extracted by applying a generalized scheme of the Matching Pursuit strategy. The methodology for shadow detection is based on the observation that shadows are half-transparent regions which retain the representation of the underlying background surface pattern. This approach assumes that shadow-regions contain same textural information, both in the current and in the background images.

In the work of Amato et al.[1] a method that introduces two discriminative features to detect moving cast shadow is presented. These features are computed based on angular and modular patterns, which are formed by similarity measurement between two sets of RGB color vectors. Unlike the most texture-based methods that often exploit spatial information, the patterns used in this approach are only photometric. This method could also be categorized as chormaticity invariant since it make uses of chroma and intensity information of a set of pixels to form a textural pattern.

Salvador et al. [50, 49] introduce a two stage method for segmenting moving shadows. The first stage segments the moving shadows in each frame of the sequence. In this stage the property that shadows casted on a surface reduce the surface intensities is exploited by using the photometric invariant C1C2C3 color space. In addition, to obtain a more robust result, the authors propose two schemes: *(i)* analyze a set of pixels (neighborhood) instead of a single pixel, and *(ii)* include geometrical verification based on boundary analysis of the shadow-candidate regions and testing the position of shadows with respect to objects. The second stage is used to obtain a coherent description of the segmented shadows over time. Therefore, the authors introduce a tracking shadow algorithm. An extension of this work was presented in [51] where the algorithm can segment cast shadow for both still and moving images.

Yang et al [69] propose a moving cast shadow detection algorithm that combines shading, color, texture, neighborhoods and temporal consistency in the scene.

In comparison with methods that perform at the pixel-level, the aforementioned methods normally exploit texture information or use information from a set of pixels, making the detection more robust against noise and more efficient in those cases where ambiguity in the pixel's information occurs. However, the main drawback of these methods reside in the choice of the region's size that will be used. In other words, a strong dependency between the size of the region and the success of the method exists.

Many factors are involved in the choice of the region's size, for example: size of the object, textural composition of the background as well as of the object, etc. Consequently, an optimal region's size highly depends on the scene; moreover, the optimal size can change for different frames, even the optimal region size can changes within the frame. Furthermore, in many cases the computational time will vary with the size of such a region.

On the other hands, there are other region-based methods that perform with local adaptive regions. Basically they attempt to segment the moving area and then analyze and classify each segment based on shadow properties. These methods take advantage from pixel-level methods since they can make use of the information of a set of pixels. Additionally, they have also an advantage with respect to the fixed region-level methods since they can automatically adapt the area of analysis. A summary of this kind of methods is given below.

Toth et al. [61] propose a shadow detection algorithm based on color and shading information. They segment an image into several regions based on color information

and the mean shift algorithm. They consider that the intensity values of a shadow pixel divided by the same pixel in the background image should be constant over a small segment.

In [16] an algorithm for outdoor scenarios is presented. Luminance, chrominance and gradient density information are exploited to create a shadow confidence score. Such a shadow score is based on three rules. The first rule claims that the luminance of the cast shadow is lower than the background. The second rule claims that the chrominance of the cast shadow is identical or slightly shifted when compared with background. And the last rule claims that the difference in gradient density between the cast shadow and background is lower than the difference in the distance of gradient between the object and background. The final classification combines the shadow score with a geometrical supporter. The geometrical cue used is based on the fact that the cast shadow is at the boundary region of moving foreground mask. That is, the cast shadow can be formed in any direction of the object, but not inside the object. However, the method is restricted to: *(i)* the areas where the shadows are casted on are not textured and *(ii)* the object shape is a convex hull which makes inappropriate to detect non-rigid object.

Rosin et al.[48] present a method based on the notion of a shadow as a semi-transparent region in the image which retains a (reduced contrast) representation of the underlying surface pattern, texture or gray value. The method uses a region growing algorithm which apply a growing criterion based on a fixed attenuation of the photometric gain over the shadow region, in comparison to the reference image. The problem with this approach is that region growing algorithm cannot perform accurately in the penumbra part of the shadow due to the intensity's variations inside of the shadow region.

Xu et al. [68] detect shadow region in indoor environment. The proposed method assumes that the shadow often appears around the foreground object. A number of techniques are used including initial change detection masks, Canny edge maps, multi-frame integration, edge matching, and conditional dilation. The method tries to detect shadow regions by extracting moving edges.

Chang et al. [8] propose a parametric Gaussian shadow model to detect and suppress pedestrian's shadow. The model makes uses of several features including the orientation, mean intensity, and center position being estimated from the properties of object movements.

In the work of Hsieh et al. [22] a line-based shadow modeling process is proposed to detect moving shadows in traffic surveillance. When a vehicle moves along a lane, it will have several boundary lines parallel or vertical to this lane. Then, the lane can provide useful information for shadow elimination and do not destroy vehicle shapes. In the method first all lanes dividing lines are detected. These lanes dividing lines from video sequences are detected by vehicle's histogram. This histogram is obtained by accumulating different vehicles' positions in a training period. According to these lines and their directions, two kinds of lines are used to eliminate shadows. The first one is the lines that are parallel to the dividing lines and the second one is the lines vertical to the dividing lines.

In [42] the authors propose an outdoor shadows removal method. It is based on a spatio-temporal-reflection test and a dichromatic reflection model. The approach is divided in several sequential steps. The step one starts with the motion mask, which is computed based on mixture of Gaussians. The intensity test takes the second step. This is in charge to discard all the foreground pixels that are more brightness than their corresponding background pixels. The third step so-called blue ratio test exploits the observation that shadows pixels falling on neutral surfaces, tend to be more blueish (this step can only be performed in neutral surfaces the authors propose to define a neutral surface based on the saturation level). The fourth step so-called albedo ratio segmentation performs a segmentation based on a spatio-temporal albedo ratio. Basically, this step attempts to obtain segmented regions with uniform reflectance. Step five removes the effect of the sky illumination. The authors claim that the reflection due to sky illumination (ambient reflection) is considered as an additive component; therefore they subtract the foreground pixels from the background. The regions that belong to the foreground will result with a very different color vectors that it is the contrary of the pixels belonging to the shadow regions. The last step aim to classify those regions that could not be labeled in previous stages. This stage computes the dominant color of the unclassified-regions (body color estimation) and compare with the body colors of material surfaces pre-stored as a background model (using a supervised-learning phase).

Similar to [42] Huerta et al. [24] use a multi-stage approach, however they use multiple cues: color and gradient information, together with known shadow properties. In this way, regions corresponding to potential shadows are grouped by considering the "bluish effect" and an edge partitioning. Additionally, temporal similarities between textures and spatial similarities between chrominance angle and brightness distortions are analyses for all potential shadows regions. Furthermore, geometrical shadow position is taken to account to avoid a misclassification of moving shadows.

In the method of Amato et al. [2], first an initial change detection mask containing moving objects and moving shadows is obtained using a background subtraction technique. Then, objects masks are computed by using connected component analysis. Based on the shadow luminance model, the authors state that in the luminance ratio space, a low gradient constancy exists in all shadowed regions, as opposed to foreground regions which, in most cases, exhibit higher gradients. To exploit these foreground-shadow characteristics, the authors designed a novel gradient-based segmentation algorithm to partition each object area into a set of low gradient segments (objects sub-segments). Then, objects sub-segments are classified as shadow or foreground, following three criteria: (i) luminance difference criterion; (ii) segment size criterion; and (iii) extrinsic terminal point weight criterion.

The challenge in these methods is not only in being able to properly analyze the segments, but also in the segmentation process. Nevertheless, this adaptive methodology is a promising way to detect moving cast shadows since the analysis is done with the context of the shadowed area having all the shadow information.

Frame-Level

There are a very few moving cast shadow detection methods that perform at the frame-level. Normally, these methods are not used in a background subtraction context. Hence, some of the closest methods related to the research line proposed in this work are briefly describe.

Liu et al. [32] detect shadows using pixel-level information, region-level information, and global-level information. Pixel-level information is extracted using GMM in the HSV color space. Local-level information is used in two ways. First, if a pixel gets a sample that is likely to be a shadow, then not only the GMM of that pixel is updated but the GMM of neighbor pixels is also updated. Second, Markov random fields are employed to represent the dependencies among neighboring pixels. For global-level information, statistical feature is exploited for whole scene over several consecutive frames.

Stauder et al. [58] use a physics-based luminance model to describe illumination changes. They assume a plain textured background and a cast shadow is determined by combining the results of change detection, static edge detection, shading change detection and penumbra detection.

1.5 Open Issues and Difficulties to Overcome

Shadows are normally considered as a local illumination problem. Obviously, areas affected by cast shadow experience a change of illumination. Often this illumination change is considered only as a decrease in brightness, without significant variation in chromaticity. However, the assumption that pixel's chromaticity is invariant to cast shadows is not always correct. It is correct, in fact, only when the chromatic components of the light sources are similar between them and there is no color blending among objects. This type of shadow is often called an achromatic shadow, while those that are not achromatic are referred to as chromatic shadows [2]. Removing chromatic shadows is a particularly challenging task due to the fact that they are extremely difficult to distinguish from the foreground because they have not a clearly defined photometric pattern. The interplay between color and texture in the background and shadows is highly variable and difficult to characterize. Another non trivial problem occurs when there is no difference in chromaticity between foreground object and background (e.g. black car is moving in highway), hence inducing a strong similarity between shadow-foreground pixels. Such effect is called as shadow camouflage. Despite of the fact that many articles of moving cast shadow detection have been published during the las years, only few works in the literature address these two major problems: chromatic shadow identification, and shadow detection in camouflaged areas. Although, methods that aim to extract certain patterns (region-based) make the detection more robust against noise and more efficient in

METHODS	Chromatic Shadows	Shadow Camouflage	Surface Topology
Cucchiara et al. [12]	*High*	*High*	*Low*
Horprasert et al. [21]	*High*	*High*	*Low*
McKena et al. [39]	*High*	*High*	*High*
Kim et al. [29]	*High*	*High*	*Low*
Siala et al. [56]	*Low*	*High*	*Low*
M.-Brisson et al. [37]	*High*	*High*	*Medium*
Huang et al. [23]	*High*	*High*	*Medium*
Fung et al. [16]	*High*	*High*	*High*
Huerta et al. [24]	*Low*	*High*	*High*
Toth et al. [61]	*Low*	*Medium*	*High*
Nadimi et al. [42]	*Low*	*High*	*Medium*
Amato et al. [1]	*High*	*Medium*	*High*
Yuan et al. [71]	*Low*	*Medium*	*High*
Grest et al. [18]	*High*	*Medium*	*High*
Yao et al. [70]	*High*	*Medium*	*High*
Leone et al. [30]	*Low*	*High*	*High*
Jacques et al. [7]	*Low*	*High*	*High*
Yang et al. [69]	*Low*	*High*	*Medium*
Amato et al. [2]	*Low*	*Low*	*Medium*

Table 1.1 Qualitative evaluation for different methods. The table valuates the negative effect degree with: *Low, Medium and High*.

those cases where ambiguity in the pixel's information occurs; these methods may also suffer from the chromatic shadow effect. Furthermore, an intrinsic limitation of some of these methods resides in the textural composition of the background as well as of the object. The surfaces' topology in terms of texture or texture-less plays a significant role in the patters extraction task.

Table 1.1 presents a qualitative comparison among several moving cast shadow detection algorithms. It reports the negative impact that chromatic shadow and shadow camouflage might cause over the performance of the methods. The table valuates the negative effect degree with: *Low, Medium and High*. In turn, the table also shows the dependency of algorithms' performance respect to surface topology (namely texture or texture-less). The degree of this dependency is similarly classified.

1.6 Evaluation of Moving Cast Shadow Detection Methods

This section explains the essential tools to evaluate the performance of moving cast shadow detection approaches. Commonly, moving cast shadow detection methods are evaluated using the metrics and the sequences described below.

Metrics

The quantitative comparison normally is based on two standard metrics for evaluating the performance of cast shadow detection algorithm introduced by Prati et. al [47]: shadow detection rate (η) and shadow discrimination rate (ξ). These two metrics are as follow:

$$\eta = \frac{TP_S}{TP_S + FN_S}; \quad \xi = \frac{\bar{\bar{TP}}_F}{TP_F + FN_F}. \tag{1.1}$$

where TP and FN stand for true positive and false negative pixels detected respect to both shadows S and foreground F. $\bar{\bar{TP}}_F$ is the number of true positive foreground pixels detected minus the number of points detected as shadows but belonging to the foreground.

The shadow detection rate η is related to the percentage of shadow pixels correctly classified, while the shadow discrimination rate ξ is concerned with foreground pixels correctly classified.

Sequences

The typical sequences normally used for evaluating moving cast shadow detection methods are[4]:

1. (Highway II, Campus, Laboratory and Intelligent Room)[5]

2. (Hallway, HWI, HWIII)[6]

3. (CVC-Outdoor, Football Match, Pets-2009 View 7)[7]

The characteristics of the sequences are summarized in Table 1.2. The characteristics' description is organized in terms of: *(i)* Frames, *(ii)* Scene, *(iii)* Object and *(iv)* Shadows. Where *(i)* includes number of frames, hand-labeled (ground truth) frames and image size. The second category *(ii)* specifies the type, background and noise of the scene. The third category *(iii)* reports the classes and the sizes of the foreground objects. Finally, the last category *(iv)* describes the casted shadows on the scene in terms of size, visibility (referred to the perception of the human eye), direction, camouflage and chromatic effect.

---

[4] Note that for a quantitative evaluation a ground truth is necessary, the sequences as well as their ground truth are publicly accessible in the listed links.

[5] http://cvrr.ucsd.edu/aton/shadow/

[6] http://vision.gel.ulaval.ca/~CastShadows/

[7] http://www.cvc.uab.es/~aamato/ShadowsDetection/;
http://www.cvg.rdg.ac.uk/PETS2009/a.html

Sequences				
		HighwayII	Campus	Laboratory	Intelligent Room

		HighwayII	Campus	Laboratory	Intelligent Room
Frames	Number	500	1179	887	300
	Hand-labeled	5	6	7	113
	Size	320x240	352x288	320x240	320x240
Scene	Type	Outdoor	Outdoor	Indoor	Indoor
	Background	Textured-less	Variable	Variable	Variable
	Noise	High	High	High	High
Object	Class	Vehicles	Vehicle and People	People	People
	Size	Variable	Large	Large	Large
Shadows	Size	Small	Large	Variable	Large
	Visibility	High	Low	Low	Low
	Direction	Single horizontal	Single horizontal	Multiple	Multiple
	Camouflage	High	Low	Low	Low
	Chromatic effect	High	Low	Low	Low

Sequences			
		Hallway	HWI	HWIII

		Hallway	HWI	HWIII
Frames	Number	1800	440	2227
	Hand-labeled	13	8	7
	Size	320x240	320x240	320x240
Scene	Type	Indoor	Outdoor	Outdoor
	Background	Textured	Textured-less	Texture-less
	Noise	Medium	Medium	Medium
Object	Class	People	Vehicles	Vehicles
	Size	Variable	Large	Variable (small)
Shadows	Size	Variable	Large	Variable
	Visibility	Low	High	High
	Direction	Multiple	Single horizontal	Single horizontal
	Camouflage	Low	High	High
	Chromatic effect	Low	Low	High

Sequences			
		CVC Outdoor	Football Match	Pets 2009 V7

		CVC Outdoor	Football Match	Pets 2009 V7
Frames	Number	800	2699	795
	Hand-labeled	12	13	16
	Size	320x240	320x240	720x576
Scene	Type	Outdoor	Outdoor	Outdoor
	Background	Textured	Textured-less	Variable
	Noise	Low	Medium	Low
Object	Class	People	People	People
	Size	Large	Small	Variable
Shadows	Size	Large	Small	Variable
	Visibility	High	Low	Low
	Direction	Single horizontal	Multiple horizontal	Single horizontal
	Camouflage	Low	Low	Low
	Chromatic effect	Medium	Low	Low

Table 1.2 Description of typical sequences normally used for evaluating moving cast shadow detection methods.

1.7 Conclusion

Firstly, the problematic of moving cast shadows in video surveillance applications has been introduced in this chapter. Additionally, classical and well-known background subtraction methods were also described. Later on, a comprehensive survey of the most significant moving cast shadow approaches as well as their taxonomies has been presented. We have observed that moving cast shadow detection methods that only exploit chromaticity invariant property are not intrinsically prepared to cope with 'chromatic shadows'. In turn, methods that perform at the 'pixel level' highly decrease their performance in those cases where 'shadow camouflage' and 'chromatic shadows' occur, since the information of a single pixel is not enough to discriminate between shadow and foreground due to the ambiguity in their pixels values. In comparison with methods that perform at the pixel-level, the region-based method make the detection more robust against noise and more efficient in those cases where ambiguity in the pixel's information occurs. However, these methods may also suffer from the chromatic shadow effect. Furthermore, an intrinsic difficulty of some of these methods (fixed region-based) resides in the criterion of the region's size that is used. Thus a strong dependency between the size of the region and the success of the method exists. Several factors are involved in the choice of the region's size, for example: size of the object, textural composition of the background as well as of the object, etc. Consequently, an optimal region's size is highly depending on the scene; moreover, an optimal size can change in different frames of the same scene or even the optimal region's size can change within the frame. Finally, the metric and the most employed surveillance data sets to evaluate the performance of moving cast shadow algorithms were reported.

References

1. A. Amato, M. Mozerov, X. Roca, and J. Gonzàlez. Robust real-time background subtraction based on local neighborhood patterns. *EURASIP Journal on Advances in Signal Processing*, pages 1–7, June 2010.
2. A. Amato, M.G. Mozerov, A.D. Bagdanov, and J. Gonzàlez. Accurate moving cast shadow suppression based on local color constancy detection. *Image Processing, IEEE Transactions on*, 20(10):2954–2966, oct. 2011.
3. O. Barnich and M. Van Droogenbroeck. Vibe: A universal background subtraction algorithm for video sequences. *IEEE TIP*, 20(6):1709–1724, June 2011.
4. S. Brutzer, B. Hoferlin, and G. Heidemann. Evaluation of background subtraction techniques for video surveillance. In *IEEE CVPR'11*, pages 1937–1944, June 2011.
5. A. Bugeau and P. Perez. Detection and segmentation of moving objects in highly dynamic scenes. In *IEEE CVPR'07*, pages 1–6, June 2008.
6. Anthony Caputo. *Digital Video Surveillance and Security*. Butterworth-Heinemann, 2010.
7. J. Cezar Silveira Jacques, C. Rosito Jung, and S.R. Musse. A background subtraction model adapted to illumination changes. In *Image Processing, 2006 IEEE International Conference on*, pages 1817–1820, October 2006.
8. Chia-Jung Chang, Wen-Fong Hu, Jun-Wei Hsieh, and Yung-Sheng Chen. Shadow elimination for effective moving object detection with gaussian models. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 2, pages 540–543, 2002.
9. Y. Chen, C. Chen, C. Huang, and Y. Hung. Efficient hierarchical method for background subtraction. *Pattern Recognition*, 40(10):2706–2715, October 2007.
10. Li Cheng, M. Gong, D. Schuurmans, and T. Caelli. Real-time discriminative background subtraction. *IEEE TIP*, 20(5):1401–1414, 2011.
11. A. Colombari, A. Fusiello, and V. Murino. Patch-based background initialization in heavily cluttered video. *IEEE TIP*, 19(4):926–933, April 2010.
12. R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirotti. Improving shadow suppression in moving object detection with hsv color information. In *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE*, pages 334–339, 2001.
13. Weiyao Lin (Ed.). *Video Surveillance*. InTech, 2011.
14. A. Elgammal, D. Harwood, and L. S. Davis. Nonparametric background model for background subtraction. In *ECCV'00*, pages 751–767, Dublin, 2000.
15. D. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Prentice Hall, August 2002.
16. G.S.K. Fung, N.H.C. Yung, G.K.H. Pang, and A.H.S. Lai. Effective moving cast shadow detection for monocular color image sequences. In *Image Analysis and Processing, 2001. Proceedings. 11th International Conference on*, pages 404–409, sep 2001.
17. D. M. Gavrila. The visual analysis of human movement: A survey. *Computer Vision and Image Understanding*, 73:82–98, 1999.
18. Daniel Grest, Jan michael Frahm, and Reinhard Koch. A color similarity measure for robust shadow removal in real time. In *In Vision, Modeling and Visualization*, pages 253–260, 2003.
19. I. Haritaoglu, D. Harwood, and L.S. Davis. W4: Real-time surveillance of people and their activities. *IEEE TPAMI*, 22(8):809–830, 2000.
20. J. Heikkila and O. Silven. A real-time system for monitoring of cyclists and pedestrians. In *Proceedings of the Second IEEE Workshop on Visual Surveillance*, pages 74–81, Washington, DC, USA, 1999. IEEE Computer Society.
21. Thanarat Horprasert, David Harwood, and Larry S. Davis. A statistical approach for real-time robust background subtraction and shadow detection. In *ICCV Frame-Rate WS*. IEEE, 1999.
22. Jun-Wei Hsieh, Shih-Hao Yu, Yung-Sheng Chen, and Wen-Fong Hu. A shadow elimination method for vehicle analysis. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 4, pages 372–375, aug. 2004.
23. Jia-Bin Huang and Chu-Song Chen. Moving cast shadow detection using physics-based features. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 0:2310–2317, 2009.

24. I. Huerta, M. Holte, T.B. Moeslund, and J. Gonzàlez. Detection and removal of chromatic moving shadows in surveillance scenarios. In *ICCV2009*, Kyoto, Japan, 2009.

25. Ivan Huerta, Ariel Amato, F. Xavier Roca, and Jordi Gonzàlez. Multiple cues fusion for robust motion segmentation using background subtraction. *Neurocomputing, Elsevier,in press*, 2011.

26. H.W.S. Jabri, Z.Duric, and A.Rosenfeld. Detection and location of people in video images using adaptive fusion of color and edge information. In *15th ICPR*, volume 4, pages 627–630, Barcelona, Spain, September 2000.

27. O. Javed, K. Shafique, and M. Shah. A hierarchical approach to robust background subtraction using color and gradient information. In *Proc. of the Workshop on Motion and Video Computing (MOTION'02)*, page 22, Orlando, 2002.

28. M. Karaman, L. Goldmann, D. Yu, and T. Sikora. Comparison of static background segmentation methods. In *VCIP '05*, Beijing, China, July 2005.

29. K. Kim, T.H. Chalidabhongse, D. Harwood, and L.S. Davis. Real-time foreground-background segmentation using codebook model. *Real-Time Imaging*, 11(3):172–185, June 2005.

30. A. Leone and C. Distante. Shadow detection for moving objects based on texture analysis. *Pattern Recognition*, 40(4):1222–1233, April 2007.

31. L. Li, W. Huang, I. Yu-Hua Gu, and Qi Tian. Statistical modeling of complex backgrounds for foreground object detection. *IEEE TIP*, 13(11):1459–1472, November 2004.

32. Zhou Liu, Kaiqi Huang, Tieniu Tan, and Liangsheng Wang. Cast shadow removal combining local and global features. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pages 1–8, june 2007.

33. E. Lopez-Rubio, R.M. Luque-Baena, and E. Dominguez. Foreground detection in video sequences with probabilistic self-organizing maps. *International Journal of Neural Systems*, 21(3):225–246, 2011.

34. L. Maddalena and A. Petrosino. A self-organizing approach to background subtraction for visual surveillance applications. *IEEE TIP*, 17(7):1168–1177, July 2008.

35. V. Mahadevan and N. Vasconcelos. Spatiotemporal saliency in dynamic scenes. *IEEE TPAMI*, 32(1):171–177, 2010.

36. N. Martel-Brisson and A. Zaccarin. Learning and removing cast shadows through a multidistribution approach. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(7):1133–1146, 2007.

37. Nicolas Martel-Brisson and André Zaccarin. Kernel-based learning of cast shadows from a physical model of light sources and surfaces for low-level segmentation. In *CVPR08*, pages 1–8, 2008.

38. A. McIvor. Background subtraction techniques. In *In Proc. of Image and Vision Computing*, Auckland, New Zealand, 2000.

39. Stephen J. McKenna, Sumer Jabri, Zoran Duric, Azriel Rosenfeld, and Harry Wechsler. Tracking groups of people. *Computer Vision and Image Understanding: CVIU*, 80(1):42–56, 2000.

40. A. Mittal and N. Paragios. Motion-based background subtraction using adaptive kernel density estimation. In *Proc. CVPR'04*, volume 2, pages 302–309, Washington DC, USA, July 2004.

41. Thomas B. Moeslund and Erik Granum. A survey of computer vision-based human motion capture. *Computer Vision and Image Understanding*, 81(3):231 – 268, 2001.

42. S. Nadimi and B. Bhanu. Physical models for moving shadow and object detection in video. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(8):1079–1087, aug. 2004.

43. Goro Obinata and Ashish Dutta. *Vision Systems: Segmentation and Pattern Recognition*. I-Tech Education and Publishing, 2007.

44. K. A. Patwardhan, G. Sapiro, and V. Morellas. Robust foreground detection in video using pixel layers. *IEEE TPAMI*, 30(4):746–751, April 2008.

45. M. Piccardi. Background subtraction techniques: a review. In *IEEE International Conference on Systems, Man and Cybernetics*, volume 4, pages 3099 – 3104, The Hague, Netherlands, 2004.

46. F. Porikli and J. Thornton. Shadow flow: a recursive method to learn moving cast shadows. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 1, pages 891–898 Vol. 1, 2005.
47. Andrea Prati, Ivana Mikic, Mohan M. Trivedi, and Rita Cucchiara. Detecting moving shadows: Algorithms and evaluation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(7):918–923, 2003.
48. Paul Rosin and Tim Ellis. Image difference threshold strategies and shadow detection. In *in Proc. British Machine Vision Conf*, pages 347–356. BMVA Press, 1995.
49. Elena Salvador, Andrea Cavallaro, and Touradj Ebrahimi. Shadow identification and classification using invariant color models. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, volume 3, pages 1545–1548, 2001.
50. Elena Salvador, Andrea Cavallaro, and Touradj Ebrahimi. Spatio-temporal shadow segmentation and tracking. In *Proc. of Visual Communications and Image Processing*, pages 389–400, 2003.
51. Elena Salvador, Andrea Cavallaro, and Touradj Ebrahimi. Cast shadow segmentation using invariant color features. *Computer Vision and Image Understanding*, 95(2):238–259, 2004.
52. Andres Sanin, Conrad Sanderson, and Brian C. Lovell. Shadow detection: A survey and comparative evaluation of recent methods. *Pattern Recogn.*, 45(4):1684–1695, April 2012.
53. J.C. SanMiguel and J.M. Martinez. On the evaluation of background subtraction algorithms without ground-truth. In *Advanced Video and Signal Based Surveillance (AVSS), 2010 Seventh IEEE International Conference on*, pages 180 –187, sept. 2010.
54. Y. Sheikh and M. Shah. Bayesian modeling of dynamic scenes for object detection. *IEEE TPAMI*, 27(11):1778–1792, November 2005.
55. J. Shen. Motion detection in color image sequence and shadow elimination. *Visual Communications and Image Processing*, 5308:731–740, January 2004.
56. K. Siala, M. Chakchouk, F. Chaieb, and O. Besbes. Moving shadow detection with support vector domain description in the color ratios space. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 4, pages 384–387, aug. 2004.
57. P. Spagnolo, T.D Orazio, M. Leo, and A. Distante. Moving object segmentation by background subtraction and temporal analysis. *Image and Vision Computing*, 24(5):411–423, May 2006.
58. Jürgen Stauder, Roland Mech, and Jörn Ostermann. Detection of moving cast shadows for object segmentation. *IEEE Transactions on Multimedia*, 1(1):65–76, 1999.
59. C. Stauffer, W. Eric, and L. Grimson. Learning patterns of activity using real-time tracking. *IEEE TPAMI*, 22(8):747–757, 2000.
60. C. Stauffer and W.E.L. Grimson. Adaptive background mixture models for real-time tracking. In *IEEE CVPR'99*, volume 1, pages 22–29, Ft. Collins, CO, USA, 1999.
61. Daniel Toth, Ingo Stuke, Andreas Wagner, and Til Aach. Detection of moving shadows using mean shift clustering and a significance test. In *International Conference on Pattern Recognition (ICPR 2004)*, volume 4, pages 260–263, 2004.
62. K. Toyama, J.Krumm, B.Brumitt, and B.Meyers. Wallflower: Principles and practice of background maintenance. In *Proc. ICCV'99*, volume 1, pages 255–261, Kerkyra, Greece, 1999.
63. Habib Ullah, Mohib Ullah, Muhammad Uzair, and Fasih ur Rehman. Comparative study: The evaluation of shadow detection methods. *INTERNATIONAL JOURNAL OF VIDEO & IMAGE PROCESSING AND NETWORK SECURITY (IJVIPNS)*, 10(2):1–7, April 2010.
64. L. Wang, W. Hu, and T. Tan. Recent developments in human motion analysis. *Pattern Recognition*, 36(3):585–601, 2003.
65. Liang Wang, Weiming Hu, and Tieniu Tan. Recent developments in human motion analysis. *Pattern Recognition*, 36(3):585 – 601, 2003.
66. Yair Weiss. Deriving intrinsic images from image sequences. In *Proc. ICCV'01*, volume 02, pages 68–75, Vancouver, Canada, 2001.
67. C.R. Wren, A. Azarbayejani, T. Darrell, and A.P. Pentland. Pfinder: Real-time tracking of the human body. *IEEE TPAMI*, 19(7):780–785, 1997.
68. Dong Xu, Xuelong Li, Zhengkai Liu, and Yuan Yuan. Cast shadow detection in video segmentation. *Pattern Recognition Letters*, 26(1):91–99, 2005.

69. Yang, Lo, Chinag, and Tai. Moving cast shadow detection by exploiting multiple cues. *Image Processing, IET*, 2(2):95–104, 2008.
70. J. Yao and J.M Odobez. Multi-layer background subtraction based on color and texture. In *IEEE CVPR'07*, pages 17–22, Minneapolis, Minnesota, USA, June 2007.
71. Chao Yuan, Chenhui Yang, and Zhiming Xu. Simple vehicle detection with shadow removal at intersection. In *Proceedings of the 2010 Second International Conference on Multi-Media and Information Technology*, volume 02 of *MMIT '10*, pages 188–191. IEEE Computer Society, 2010.
72. Wei Zhang, Q.M. Jonathan Wu, and Xiangzhong Fang. *Vision Systems: Segmentation and Pattern Recognition*. Moving Cast Shadow Detection. Goro Obinata and Ashish Dutta, InTech.
73. Z. Zivkovic. Improved adaptive gaussian mixture model for background subtraction. In *Proc. ICPR'04*, volume 2, pages 23–26, August 2004.
74. 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, May 2006.