Analysis and Detection of Shadows in Video Streams: A Comparative Evaluation

Andrea Prati, Rita Cucchiara
Dip. di Ingegneria dell'Informazione
University of Modena and Reggio Emilia
Modena, Italy - 41100

Ivana Mikić, Mohan M. Trivedi
Dept. of Electrical and Computer Engineering
University of California, San Diego
La Jolla, CA, USA - 92037

Abstract

Robustness to changes in illumination conditions as well as viewing perspectives is an important requirement for many computer vision applications. One of the key factors in enhancing the robustness of dynamic scene analysis is that of accurate and reliable means for shadow detection. Shadow detection is critical for correct object detection in image sequences. Many algorithms have been proposed in the literature that deal with shadows. However, a comparative evaluation of the existing approaches is still lacking. In this paper, the full range of problems underlying the shadow detection are identified and discussed. We classify the proposed solutions to this problem using a taxonomy of four main classes, called deterministic model and non-model based and statistical parametric and non-parametric. Novel quantitative (detection and discrimination accuracy) and qualitative metrics (scene and object independence, flexibility to shadow situations and robustness to noise) are proposed to evaluate these classes of algorithms on a benchmark suite of indoor and outdoor video sequences.

1. Introduction

Moving object segmentation is an essential issue in many computer vision applications dealing with image sequences. Moving shadows do, however, cause serious problems while extracting moving objects, due to the misclassification of shadow points as foreground. Shadows can cause merging of objects, object shape distortion and even object losses (due to the shadow cast over another object). The difficulties associated with shadow detection arise since shadow and objects share two important visual features. Shadows are dark and typically differ significantly from the background and they have the same motion as the objects casting them.

In literature there are many different approaches to moving object segmentation from image sequences, based on inter-frame differencing, background subtraction, optical flow, statistical point classification or feature matching and tracking. However, neither motion segmentation nor change detection methods can distinguish between moving objects and moving shadows. For this reason, the efforts of computer vision community in finding robust shadow detection algorithms have intensified in the recent years.

In this paper we present a survey of shadow detection approaches, providing both a classification and a comparative evaluation of representative algorithms present in literature. This comparison will take into account both the advantages and the drawbacks of each algorithm class and will furnish a quantitative (objective) and qualitative (subjective) evaluation of them. We classify the proposed solutions to this problem using a taxonomy of four main classes, called deterministic model and non-model based and statistical parametric and non-parametric. Novel quantitative (detection and discrimination accuracy) and qualitative metrics (scene and object independence, flexibility to shadow situations and robustness to noise) are proposed to evaluate these classes of algorithms on a benchmark suite of indoor and outdoor video sequences.

In the next Section the reasoning process behind the identification of shadows and the general framework used for shadow detection are described. The approaches that use this process to detect shadows are organized in a taxonomy in Section 3. Each approach is detailed and discussed to emphasize its strengths and its limits. Section 4 presents the evaluation metrics chosen to compare the approaches and outlines their relevance, while Section 5 reports the quantitative and qualitative experimental results. Conclusions end the paper.

2. Shadow perception

2.1. What is a shadow?

Shadows are due to the occlusion of the light source by an object in the scene. The part of an object that is not illuminated is called *self-shadow*, while the area projected on the scene by the object is called *cast shadow* and is further classified into *umbra* and *penumbra* [13]. The umbra corresponds to the area where the direct light is totally blocked by the object, whereas in the penumbra area it is partially blocked. The umbra is easier to be seen and detected, but is more likely to be misclassified as moving object. If the object is moving, the cast shadow is more properly called *moving cast shadow* [22].

It can be demonstrated [8] that, starting from the Kubelka-Munk theory and if we assume white illumination and matte surfaces, we can model the appearance of a point with the following equation:

$$s_k(x,y) = E_k(x,y)\rho_k(x,y) \quad (1)$$

where s_k is the image *luminance* of the point of coordinate (x,y) at time instant k. $\rho_k(x,y)$ is the *reflectance* of the object surface and $E_k(x,y)$ is the *irradiance*, i.e. the amount of luminance energy received by the surface **S** per area unit. This value can be derived by the Phong model with the assumptions that light source is far from the object, the distance between light source and surface **S** is constant, the light source emits parallel light rays and the observation point is fixed. In this case, the irradiance $E_k(x,y)$ can be approximated as:

$$E_k(x,y) = \begin{cases} c_A + c_P \cos\angle(\mathbf{N}(x,y),\mathbf{L}) & illuminate \\ c_A + k(x,y)c_P \cos\angle(\mathbf{N}(x,y),\mathbf{L}) & penumbra \\ c_A & umbra \end{cases} \quad (2)$$

where c_A and c_P are the intensity of the ambient light and of the light source, respectively, $\mathbf{L}$ the direction of the light source, **N**(x,y) the object surface normal and $k(x,y)$ describes the softening due to the penumbra, then $0 \leq k(x,y) \leq 1$.

2.2. General framework for shadow detection

These general assumptions are useful to describe the model of the appearance of a point. However, shadow detection algorithms often make additional assumptions [22] that the light source is strong, that the background is static and textured and is assumed to be planar and that the extent of the light source is sufficiently large to cause the formation of penumbra [13].

Typically the first step computed for shadow detection is the difference between the current frame s_k (at time k) and a reference image s_0, that can be the previous frame, as in [22], or a reference frame, typically named *background model* [14][20][3].

With the assumption of a static background, the reflectance $\rho_k(x,y)$ of the background does not change with time, thus we can assume that $\rho_k(x,y) = \rho_0(x,y) = \rho(x,y)$. Therefore, the difference computed for detecting shadow is $D_k(x,y) = \rho(x,y)(E_k(x,y) - E_0(x,y))$ that, if a previously illuminated point is covered by a cast shadow (umbra) at frame k, becomes:

$$D_k(x,y) = \rho(x,y)c_P \cos\angle(\mathbf{N}(x,y),\mathbf{L}) \quad (3)$$

Thus, if the light source is assumed to be strong the difference should be high.

The approaches in literature differ by means of how they distinguish between foreground and shadow points. Most of these works locally exploit pixel appearance change due to cast shadows [13][20][9][5][22]. A possible approach is to compute the ratio $R_k(x,y)$ between the appearance of the pixel in the actual frame and the appearance in a reference frame. Using the above-mentioned constancy of the reflectance, this ratio can be rewritten as:

$$R_k(x,y) = \frac{E_k(x,y)}{E_0(x,y)} \quad (4)$$

The works we will compare start from these assumptions. They can be summarized as the hypotheses (approximated but realistic) that the light source is white, isotropically scattered within the object and sufficiently strong; moreover, the reference image is assumed static, textured and planar; finally, the objects are considered with perfectly matte surfaces (or *Lambertian* surfaces).

3. Taxonomy of shadow detection algorithms

Most of the proposed approaches take into account the described shadow model. To account for their differences, we have organized the efforts present in literature in a taxonomy. The first 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. Introducing uncertainty to the class membership assignment can reduce noise sensitivity by relaxing ill-posed constraints. In the statistical-based methods (as [17][5][11]) the parameter selection is a critical issue. The work reported in [17] is an example of the *parametric* approach, whereas [5][11] are examples of the *non-parametric* approach. However, the correct trade-off between automatic parameter selection and responsiveness to changes has to be found.

Within the deterministic class (see [13][22][3][15]), another sub-classification can be based on whether the on/off decision can be supported by model based knowledge or not. Choosing a *model based* approach achieves undoubtedly the best results, but is, most of the times, too complex and time consuming compared to the *non-model based*. Moreover, the number and the complexity of the models increases rapidly if the aim is to deal with complex and clut-

Statistical parametric				Statistical non-parametric			
Paper	Spectral	Spatial	Temporal	Paper	Spectral	Spatial	Temporal
Friedman and Russell 1997 [6]	C	R	D	**Horprasert et al. 1999 [11]**	C	R	D
Mikić et al. 2000 [17]	C	L	D	Tao et al.[4] 2000 [23]	C	R	D
Deterministic model based				**Deterministic non-model based**			
Paper	Spectral	Spatial	Temporal	Paper	Spectral	Spatial	Temporal
Irvin and McKeown Jr.[1] 1989 [12]	G	L	S	Scanlan et al.[1] 1990 [19]	G	L	S
Wang et al. 1991 [25]	G	R	S	Jiang and Ward[1] 1992 [13]	G	R	S
Kilger 1992 [14]	G	R	S	Charkari and Mori 1993 [2]	G	R	S
Koller et al. 1993 [15]	G	L	S	Sexton and Zhang 1993 [20]	G	L	S
Onoguchi[2] 1998 [18]	G	L	S	Funka-Lea and Bajcsy[1] 1995 [7]	G	R	D
				Sonoda and Ogata 1998 [21]	G	R	S
				Tzomakas and von Seelen 1998 [24]	G	R	S
				Amamoto and Fujii 1999 [1]	G	N/A[3]	D
				Stauder et al. 1999 [22]	G	R	D
				Cucchiara et al. 2001 [3]	C	L	S

Table 1. Classification of the literature on shadow detection. Most of the papers presented in literature are classified according to the four classes proposed. Their authors, the reference work and the year of publication are reported and the spectral, spatial and temporal features used are depicted (G=grey-level, C=color, L=local/pixel-level R=region-level, S=static, D=dynamic).

tered environments with different lighting conditions, object classes and perspective views.

Finally, we can describe each approach in terms of its use of *spectral*, *spatial* and *temporal* features. Approaches can exploit differently spectral features, i.e. using gray level or color information. Some approaches improve results by using spatial information working at a region level, instead of pixel level. Finally, some methods exploit temporal redundancy to integrate and improve results.

In Table 1 we report most of the papers dealing with shadow detection. Their spectral, spatial and temporal features are outlined. In this paper, we focus our attention on four algorithms (reported in bold in Table 1) representative of three of the above-mentioned classes. The deterministic model-based class (as, for instance, [15]) has not been considered due to its complexity and to its lack of generality.

3.1. Statistical non-parametric (SNP)

As an example of statistical non-parametric (SNP) approach we choose the one described in [9] and detailed in [11]. This work considers the *color constancy* ability of human eyes and exploits the Lambertian hypothesis to consider color as a product of irradiance and reflectance. The distortion of the brightness α_i and the distortion of the chrominance CD_i of the difference between expected color of a pixel and its value in the current image are computed as:

$$\alpha_i = \frac{\left(\frac{I_R(i)\mu_R(i)}{\sigma_R^2(i)} + \frac{I_G(i)\mu_G(i)}{\sigma_G^2(i)} + \frac{I_B(i)\mu_B(i)}{\sigma_B^2(i)}\right)}{\left(\left[\frac{\mu_R(i)}{\sigma_R(i)}\right]^2 + \left[\frac{\mu_G(i)}{\sigma_G(i)}\right]^2 + \left[\frac{\mu_B(i)}{\sigma_B(i)}\right]^2\right)} \tag{5}$$

$$CD_i = \sqrt{\left(\frac{I_R(i) - \alpha_i\mu_R(i)}{\sigma_R(i)}\right)^2 + \left(\frac{I_G(i) - \alpha_i\mu_G(i)}{\sigma_G(i)}\right)^2 + \left(\frac{I_B(i) - \alpha_i\mu_B(i)}{\sigma_B(i)}\right)^2} \tag{6}$$

and normalized w.r.t. their root mean square of pixel i. The values $\widehat{\alpha}_i$ and $\widehat{CD}_i$ obtained are used to classify a pixel in four categories:

$$C(i) = \begin{cases} Foreg.: & \widehat{CD}_i > \tau_{CD} \quad or \quad \widehat{\alpha}_i < \tau_{\alpha lo}, \quad else \\ Backg.: & \widehat{\alpha}_i < \tau_{\alpha 1} \quad and \quad \widehat{\alpha}_i > \tau_{\alpha 2}, \quad else \\ Shad.: & \widehat{\alpha}_i < 0, \quad else \\ Highl.: & otherwise \end{cases} \tag{7}$$

The rationale used is that shadows have similar chromaticity but lower brightness than the background model. A statistical learning procedure is used to automatically determine the appropriate thresholds.

3.2. Statistical parametric (SP)

The algorithm described in [17] for traffic scene shadow detection is an example of statistical parametric (SP) approach and it takes into account most of the assump-

[1]This paper addresses still images

[2]This paper uses an innovative approach based on *inverse perspective mapping* in which the assumption is that the shadow and the object that casts it are overlapped if projected on the ground plane.

[3]This paper has the unique characteristic to use the DCT to remove shadow. The rationale used by the authors is that a shadow has, in the frequency domain, a large DC component, whereas the moving object has a large AC component.

[4]Since this paper uses a fuzzy neural network to classify points as belonging or not to a shadow, it can be considered a statistical approach.

tions described in Section 2. This algorithm claims to use two sources of information: *local* (based on the appearance of the pixel) and *spatial* (based on the assumption that the objects and the shadows are compact regions). The a-posteriori probabilities of belonging to background, foreground and shadow classes are maximized. The a-priori probabilities of a pixel belonging to shadow are computed by assuming that $\mathbf{v} = [R, G, B]^T$ is the value of the pixel not shadowed and by using an approximated linear transformation $\bar{\mathbf{v}} = \mathbf{D}\mathbf{v}$ (where $\mathbf{D} = diag(d_R, d_G, d_B)$ is a diagonal matrix obtained by experimental evaluation) to estimate the color of the point covered by a shadow. The $\mathbf{D}$ matrix is assumed approximately constant over flat surfaces. If the background is not flat over the entire image, different $\mathbf{D}$ matrices must be computed for each flat subregion. The spatial information is exploited by performing an iterative probabilistic relaxation to propagate neighborhood information.

In this statistical *parametric* approach the main drawback is the difficult process necessary to select the parameters. Manual segmentation of a certain number of frames has to be done to collect statistics and to compute the values of matrix $\mathbf{D}$.

3.3. Deterministic non-model based with color exploitation (DNM1)

The system described in [3] is an example of deterministic non-model based approach (and we called it DNM1). This algorithm works in the HSV color space. The main reasons are that HSV color space corresponds closely to the human perception of color [10] and it has revealed more accuracy in distinguishing shadows. In fact, a shadow cast on a background does not change significantly its hue [4]. Moreover, the authors exploit saturation information since they note that shadows often lower the saturation of the points. The resulting decision process is reported in the following equation:

$$SP_k(x,y) = \begin{cases} 1 & if \quad \alpha \leq \frac{I_k^V(x,y)}{B_k^V(x,y)} \leq \beta \\ & \wedge \, (I_k^S(x,y) - B_k^S(x,y)) \leq \tau_S \\ & \wedge \, |I_k^H(x,y) - B_k^H(x,y)| \leq \tau_H \\ 0 & otherwise \end{cases} \tag{8}$$

where $I_k(x,y)$ and $B_k(x,y)$ are the pixel values at coordinate (x,y) in the input image (frame k) and in the background model (computed at frame k), respectively. The use of β prevents the identification as shadows of those points where the background was slightly changed by noise, whereas α takes into account the "power" of the shadow, i.e. how strong the light source is w.r.t. the reflectance and irradiance of the objects. Thus, stronger and higher the sun (in the outdoor scenes), the lower α should be chosen. The choice of the parameters τ_H and τ_S is less straightforward and, for now, is done empirically.

3.4. Deterministic non-model based with spatial redundancy exploitation (DNM2)

Finally, we compare the approach presented in [22]. This is also a deterministic non-model based approach, but we have included it because of its completeness (is the only work in literature that deals with penumbra in moving cast shadows).

The shadow detection is provided by verifying three criteria: the presence of a "darker" uniform region, by assuming that the ratio of equation **??** is locally constant in presence of cast shadows; the presence of a high difference in luminance w.r.t reference frame (using equation **??**); and the presence of static and moving edges. Static edges hint a static background and can be exploited to detect nonmoving regions inside the frame difference. Moreover, to detect penumbra the authors propose to compute the width of each edge in the difference image. Since penumbra causes a soft luminance step at the contour of a shadow, they claim that the edge width is the more reliable way to distinguish between objects contours and shadows contours (characterized by a width greater than a threshold).

This approach is one of the most complete and robust proposed in literature. Nevertheless, in this case the assumptions and the corresponding approximations introduced are strong and they could lack in generality. Also, the penumbra criterion is not explicitly exploited to *add* penumbra points as shadow points, but it is only used to *remove* the points that do not fit this criterion. Moreover, the proposed algorithm uses the previous frame (instead of the background) as reference frame. This choice exhibits some limitations in moving region detection since it is influenced by object speed and it is too noise sensitive. Thus, to make the comparison of these four approaches as fair as possible, limited to the shadow detection part of the system, we implement the DNM2 approach using a background image as reference image, as the other three approaches do.

4. Performance evaluation metrics

In this section, the methodology used to compare the four approaches is outlined. In order to systematically evaluate various shadow detectors, it is useful to identify the following two important quality measures: *good detection* (low error probability to detect correct shadow points should occur) and *good discrimination* (the probability to identify wrong points as shadow should be low, i.e. low false alarm rate). The first one can be achieved by minimizing the *false negatives (FN)*, i.e. the shadow points classified as background/foreground, while to obtain

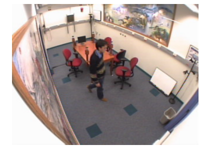
	Highway I	**Highway II**	**Campus**	**Laboratory**	**Intelligent room**
					
Sequence Type	outdoor	outdoor	outdoor	indoor	indoor
Shadow strength	medium	high	low	very low	low
Shadow size	large	small	very large	medium	large
Object class	vehicles	vehicles	vehicle/people	people/other	people
Object size	large	small	medium	medium	medium
Object speed	medium	high	low	low	low
Noise level	medium	medium	high	low	medium

Table 2. The sequence benchmark used. The benchmark should be complete and not trivial to stress the shadow detection capabilities of the approach under comparison. The sequence set chosen has both indoor and outdoor scenes, including large and smoothed shadows as well as small and dark ones. It contains different object classes with various size and speed. An evaluation of noise in the images of each sequence is reported too.

a good discrimination, the *false positives (FP)*, i.e. the foreground/background points detected as shadows, should be minimized.

A reliable and objective way to evaluate this type of visual-based detection is still lacking in literature. In [16], the authors proposed two metrics for moving object detection evaluation: the *Detection Rate (DR)* and the *False Alarm Rate (FAR)*. These figures are not selective enough for shadow detection evaluation, since they do not take into account whether a point detected as shadow belongs to a foreground object or to the background. If shadow detection is used to improve moving object detection, only the first case is problematic, since false positives belonging to the background do not affect neither the object detection nor the object shape.

For this reason, we choose as metrics the *shadow detection accuracy* η and the *shadow discrimination accuracy* ξ computed as follows:

$$\eta = \frac{TP_S}{TP_S + FN_S} \tag{9}$$

$$\xi = \frac{\overline{TP_F}}{TP_F + FN_F} \tag{10}$$

where the subscript S stays for shadow and F for foreground. TP is the number of *true positives* (i.e. the number of points correctly classified). The $\overline{TP_F}$ is the number of ground-truth points of the foreground objects minus the number of points detected as shadows but belonging to foreground objects.

These quantitative measures do not complete the evaluation. Other important features of a shadow detection algorithm should be: *robustness to noise*, *flexibility to shadow strength, size and shape*, *object independence*, *scene independence*, *computational load* and *detection of indirect cast shadow and penumbra*. Indirect cast shadows are the shadows cast by a moving object over another moving object and their effect is to decrease the intensity of the moving object covered, probably affecting the object detection, but not the shadow detection. However, how the algorithm deals with them and with the penumbra problem is an evaluation parameter.

5. Quantitative and qualitative comparison

In this section, the experimental results and the quantitative and qualitative comparison of the four approaches are presented. First, a set of sequences to test the algorithms was chosen to form a complete and non trivial benchmark suite. We select the sequences reported in Table 2, where both indoor and outdoor sequences are present, where shadows range from dark and tight to light and large and where the object type, size and speed vary considerably. The *Highway I* and the *Highway II* sequences show a traffic environment (at two different light conditions) where the shadow suppression is very important to avoid misclassification and erroneous counting of vehicles on the road. The *Campus* sequence is a noisy sequence from outdoor campus site where cars approach to an entrance barrier and students are walking around. The two indoor sequences report two laboratory rooms in two different perspectives and lighting conditions. In the *Laboratory* sequence, besides walking people, a chair is moved in order to detect its shadow.

	Highway I		Highway II		Campus		Laboratory		Intelligent Room	
	$\eta\%$	$\xi\%$	$\eta\%$	$\xi\%$	$\eta\%$	$\xi\%$	$\eta\%$	$\xi\%$	$\eta\%$	$\xi\%$
SNP	81.59%	63.76%	51.20%	78.92%	80.58%	69.37%	84.03%	92.35%	78.63%	89.92%
SP	59.59%	84.70%	46.93%	91.49%	72.43%	74.08%	64.85%	95.39%	78.50%	91.99%
DNM1	69.72%	76.93%	54.07%	78.93%	82.87%	86.65%	76.26%	89.87%	76.52%	92.32%
DNM2	75.49%	62.38%	60.24%	72.50%	69.10%	62.96%	60.34%	81.57%	71.68%	86.02%

Table 3. Experimental results. Each approach has been tested on the benchmark. For each experiment the shadow detection accuracy η and the shadow discrimination accuracy ξ in percentage are reported.

5.1. Quantitative comparison

To compute the evaluation metrics described in Section 4, the ground-truth for each frame is necessary. We obtained it by segmenting the images with a long and accurate manual classification of points in foreground, background and shadow. We prepared ground truth on tens of frames for each video sequence representative of different situations (dark/light objects, multiple objects or single object, occlusions or not).

Results are reported in Table 3. To establish a fair comparison, algorithms do not implement any background updating process, but compute the reference image and other parameters from the first N frames (with N varying with the sequence considered). Moreover, Fig. 1 reports the ranking in shadow detection and shadow discrimination accuracy derived from Table 3. Finally, two AVI files have been included with this paper, reporting visual results of the four approaches in a subset of the *Intelligent Room* and of the *Highway I* sequences.

The SNP algorithm is very effective in most of the cases, but with very variable performances. It achieves the best detection performance and high discrimination accuracy in the indoor sequences (*Laboratory* and *Intelligent Room*), with percentages up to 92%. However, the discrimination accuracy is quite low in the *Highway I* and *Campus* sequences. This can be justified by the dark aspect of the objects in the *Highway I* scene and by the strong noise of the *Campus* sequence.

The SP approach achieves good discrimination accuracy in most of the cases. Nevertheless, its detection accuracy is poor in all the cases but the *Intelligent room* sequence. This is mainly due to the approximation of constant $\mathbf{D}$ matrix on the entire image. Since the background can be rarely assumed as flat on the entire image, this approach lacks in generality. Nevertheless, good accuracy in the case of *Intelligent room* test shows how this approach can deal with indoor sequences once the constancy of the $\mathbf{D}$ matrix is almost guaranteed.

The DNM1 algorithm is the one with the most stable performance, even with totally different video sequences. It achieves good accuracy in almost all the sequences, but it outperforms the other algorithms only in the *Campus* sequence.

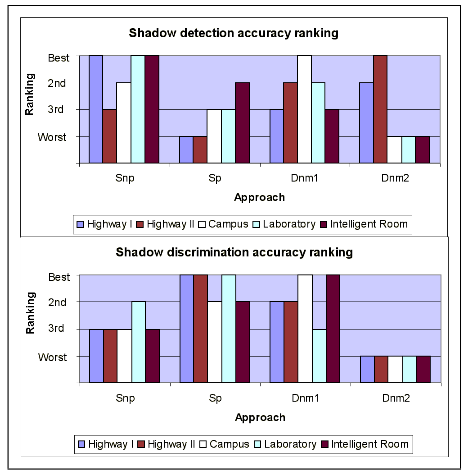
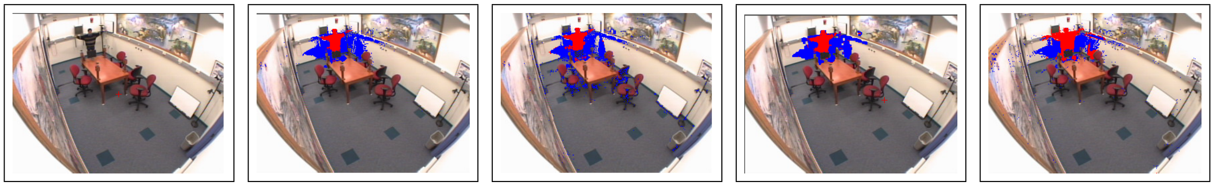


Figure 1. Ranking of the four approaches with regard to the shadow detection accuracy and the shadow discrimination accuracy. For each approach the ranking for each sequence is reported.

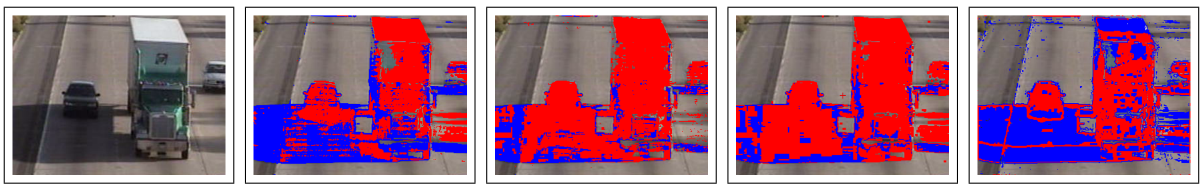
The DNM2 algorithm suffers from the assumption of planar background. This assumption fails in the case of the *Laboratory* sequence where the shadows are cast both on the floor and on the cabinet. The low detection performance in the *Campus* sequence is mainly due to noise and this algorithm has proven low robustness to strong noise. Finally, this algorithm achieves the worst discrimination result in all the cases. This is due to its assumption of textured objects: if the object is not textured (or seems not textured due to the distance and the quality of the acquisition system), the probability that parts of the object are classified as shadow arises. Nevertheless, this approach is very promising and complete and outperforms the others in the more difficult sequence (*Highway II*).

Summarizing, the statistical approaches show a good ro-



(a) Raw image (b) SNP result (c) SP result (d) DNM1 result (e) DNM2 result

Figure 2. Results of shadow detection in the indoor sequence *Intelligent room*. **These snapshots have been taken from the video included with the paper. In all these figures and in the video, red pixels identify foreground point and blue pixels indicate shadow points.**



(a) Raw image (b) SNP result (c) SP result (d) DNM1 result (e) DNM2 result

Figure 3. Results on indirect cast shadow. In the case of moving shadow casts on a moving object, the four algorithms behave in a different way. The SP algorihtm detects correctly most of the indirect cast shadow (Fig. (b)), where the SNP (Fig. (c)) and the DNM1 (Fig. (d)) algorithms do not. The best detection is achieved by the DNM2 method (Fig. (e)). In all these figures, red pixels identify foreground point and blue pixels indicate shadow points.

bustness to noise, due to statistical modeling of noise. On the other hand, deterministic approaches (in particular if pixel-based and almost unconstrained as DNM1) exhibit a good flexibility to different situations. Difficult sequences, like *Highway II*, require, however, a more specialized and complete approach to achieve good accuracy. To help evaluating the approaches the results on the *Highway I* outdoor sequence and on the *Intelligent room* indoor sequence are included with this paper and a snapshot for the last one is reported in Fig. 2.

5.2. Qualitative comparison

To evaluate the behaviour of the four algorithms with respect to the qualitative issues presented in Section 4, we vote them ranging from ”very low” to ”very high” (see Table 4). The DNM1 method is the most robust to noise, thanks to its pre- and post-processing algorithms [3]. The capacity to deal with different shadow size and strength is high in both the SNP and the DNM1. However, the higher flexibility is achieved by the DNM2 algorithm which is able to detect even the penumbra in an effective way. Nevertheless, this algorithm is very object-dependent, in the sense that, as already stated, the assumption on textured objects affects strongly the results. Moreover, the two frame difference approach proposed in [22] is weak as soon as the object speeds increase.

The planar background hypothesis makes the DNM2 and especially the SP approaches more scene-dependent than the other two. Although we can not claim to have implemented these algorithms in the most efficient way, the DNM2 seems the more time consuming, due to the amount of processing necessary. On the other hand, the SNP is very fast.

Finally, we try to evaluate the behaviour of the algorithms in presence of indirect cast shadows (see Section 4). In Fig. 3 the results in the case of indirect cast shadow are shown. The DNM2 approach is able to detect both the penumbra and the indirect cast shadow in a very effective way. Note that the penumbra in Fig. 3(e) has been detected as foreground because, accordingly to [22], we use penumbra information only to assert the classification process and not to detect penumbra points. The SP and the DNM1 methods lack in detecting indirect cast shadows. The

	Robustness to noise	Flexibility to shadow	Object independence	Scene independence	Computational load	Indirect shadow & penumbra detection
SNP	high	high	high	high	very low	high
SP	high	medium	high	low	low	low
DNM1	very high	high	high	high	low	very low
DNM2	low	very high	low	medium	high	very high

Table 4. Qualitative evaluation. Six parameters have been chosen to overview the features of the approaches presented. A vote from "very low" to "very high" to each approach has been given for each parameter.

pixel-based decision made can not distinguish correctly between this type of moving shadows and those shadows cast on the background. However, the SP approach is able to detect penumbra, at least if sufficiently narrow.

6. Conclusions

We can conclude that if a *general-purpose system*, able to detect shadow in many different situations, is needed, less assumptions should be made. For this reason, a *pixel-based deterministic non-model based approach*, as DNM1, assures best results. If the main goal is to detect efficiently *every kind of shadows or special shadows in one specific environment*, more assumptions yield better results. For this reason, even if it is not in this evaluation, the best choice is the *deterministic model-based approach*. In this situation, if *the object classes are too numerous* to allow modeling of every class, a *complete deterministic approach*, like the DNM2, should be selected. If *the environment is indoor*, the *statistical approaches* are the more reliable, since the scene is stable and constant and a statistical description is very effective. If there are different planes onto which the shadows can be cast, an approach like SNP is the best choice. *If the shadows are scattered, narrow, or particularly "blended" to the environment*, a region-based dynamic approach, typically deterministic, is the best choice (as DNM2 in the *Highway II* scene reported in this paper). Finally, if *the scene is noisy*, a *statistical approach or a deterministic approach with effective pre- and post-processing steps* should be used.

References

[1] N. Amamoto and A. Fujii. Detecting obstructions and tracking moving objects by image processing technique. *Electronics and Communications in Japan, Part 3*, 82(11):28–37, 1999.

[2] N. Charkari and H. Mori. A new approach for real time moving vehicle detection. In *Proceeding of IEEE/RSJ Int'l Conference on Intelligent Robots and Systems*, pages 273–278, 1993.

[3] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati. Detecting objects, sahdows and ghosts in video streams by exploiting color and motion information. In *Proceedings of the IEEE Int'l Conference on Image Analysis and Processing, to appear*, 2001.

[4] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirotti. Improving shadow suppression in moving object detection with hsv color information. In *Proceedings of the IEEE Int'l Conference on Intelligent Transportation Systems, to appear*, 2001.

[5] A. Elgammal, D. Harwood, and L. Davis. Non-parametric model for background subtraction. In *Proceedings of IEEE ICCV'99 FRAME-RATE Workshop*, 1999.

[6] N. Friedman and S. Russell. Image segmentation in video sequences: a probabilistic approach. In *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence*, 1997.

[7] G. Funka-Lea and R. Bajcsy. Combining color and geometry for the active, visual recognition of shadows. In *Proceedings of IEEE Int'l Conference on Computer Vision*, pages 203–209, 1995.

[8] J.-M. Geusebroek, A. Smeulders, and R. van den Boomgaard. Measurement of color invariants. In *Proceedings of IEEE Int'l Conference on Computer Vision and Pattern Recognition*, volume 1, pages 50–57, 2000.

[9] I. Haritaoglu, D. Harwood, and L. Davis. W4: real-time surveillance of people and their activities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):809–830, Aug. 2000.

[10] N. Herodotou, K. Plataniotis, and A. Venetsanopoulos. A color segmentation scheme for object-based video coding. In *Proceedings of the IEEE Symposium on Advances in Digital Filtering and Signal Processing*, pages 25–29, 1998.

[11] T. Horprasert, D. Harwood, and L. Davis. A statistical approach for real-time robust background subtraction and shadow detection. In *Proceedings of IEEE ICCV'99 FRAME-RATE Workshop*, 1999.

[12] R. Irvin and D. McKeown Jr. Methods for exploiting the relationship between buildings and their shadows in aerial imagery. *IEEE Transactions on Systems, Man, and Cybernetics*, 19:1564–1575, 1989.

[13] C. Jiang and M. Ward. Shadow identification. *Proceedings of IEEE Int'l Conference on Computer Vision and Pattern Recognition*, pages 606–612, 1992.

[14] M. Kilger. A shadow handler in a video-based real-time traffic monitoring system. *Proceedings of IEEE Workshop on Applications of Computer Vision*, pages 11–18, 1992.

[15] D. Koller, K. Daniilidis, and H. Nagel. Model-based object tracking in monocular image sequences of road traffic scenes. *International Journal of Computer Vision*, 10:257–281, 1993.

[16] G. Medioni. Detecting and tracking moving objects for video surveillance. In *Proceedings of IEEE Int'l Conference on Computer Vision and Pattern Recognition*, volume 2, pages 319–325, 1999.

[17] I. Mikic, P. Cosman, G. Kogut, and M. Trivedi. Moving shadow and object detection in traffic scenes. In *Proceedings of Int'l Conference on Pattern Recognition*, Sept. 2000.

[18] K. Onoguchi. Shadow elimination method for moving object detection. In *Proceedings of Int'l Conference on Pattern Recognition*, volume 1, pages 583–587, 1998.

[19] J. Scanlan, D. Chabries, and R. Christiansen. A shadow detection and removal algorithm for 2-d images. In *Proceedings of Int'l Conference on Acoustics, Speech and Signal Processing*, volume 4, pages 2057–2060, 1990.

[20] G. Sexton and X. Zhang. Suppression of shadows for improved object discrimination. In *Proc. IEE Colloq. Image Processing for Transport Applications*, pages 9/1–9/6, Dec. 1993.

[21] Y. Sonoda and T. Ogata. Separation of moving objects and their shadows, and application to tracking of loci in the monitoring images. In *Proceedings of Int'l Conference on Signal Processing*, pages 1261–1264, 1998.

[22] J. Stauder, R. Mech, and J. Ostermann. Detection of moving cast shadows for object segmentation. *IEEE Transactions on Multimedia*, 1(1):65–76, Mar. 1999.

[23] X. Tao, M. Guo, and B. Zhang. A neural network approach to the elimination of road shadow for outdoor mobile robot. In *Proceedings of IEEE Int'l Conference on Intelligent Processing Systems*, volume 2, pages 1302–1306, 1997.

[24] C. Tzomakas and W. von Seelen. Vehicle detection in traffic scenes using shadows. Technical Report 98-06, IR-INI, Institut fur Nueroinformatik, Ruhr-Universitat Bochum, FRG, Germany, Aug. 1998.

[25] C. Wang, L. Huang, and A. Rosenfeld. Detecting clouds and cloud shadows on aerial photographs. *Pattern Recognition Letters*, 12(1):55–64, Jan. 1991.