

Shadow Detection in Complex Environments via An Innovative Information Fusion Approach

Alfredo Cuzzocrea

DIA Department, University of Trieste and ICAR-CNR, Italy
alfredo.cuzzocrea@dia.units.it

Enzo Mumolo

DIA Department, University of Trieste, Italy
mumolo@units.it

Alessandro Moro

Chuo University, Tokyo, Japan
moro@sensor.mech.chuo-u.ac.jp

Kazunori Umeda

Chuo University, Tokyo, Japan
umeda@mech.chuo-u.ac.jp

Gianni Vercelli

DIBRIS Department, University of Genova, Italy
gianni.vercelli@unige.it

Abstract

In this paper a novel moving shadows detecting algorithm is proposed. The algorithm can be used in indoor and outdoor environments. The algorithm we propose fuses together color and stereo disparity information using the Dempster-Shafer combination rule. Some considerations on the nature of the shadow improves the algorithm's ability to candidate the pixels as shadow or foreground. The candidate of both color and disparity information are then weighted by analyzing the effectiveness in the scene.

1. Introduction

Shadow detection plays an important role in many machine vision applications. Correct shadow detection may lead to important performance improvements in scene understanding, object segmentation, tracking, recognition.

It is not difficult for human eyes to distinguish shadows from objects. However, identifying shadows by computer is still a challenging research problem. Shadows occur when objects totally or partially occlude direct light from a light source. Generally speaking, shadows are composed by two parts: self-shadow and cast shadow. The former is the part of the object which is not illuminated by the light source. The last one is the area projected on the scene by the object and is further classified in umbra and penumbra. 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 cast shadow is more properly called moving cast shadow if the object is moving. Moving

shadows cause the erroneous segmentation of objects in the scene. To solve this problem, moving shadows have to be detected explicitly to prevent them being misunderstood as moving objects or their parts.

In order to systematically develop and evaluate various shadow detector, the quality measures to minimize include the following: (i) *Detection rate*, which is the error probability to detect correct shadow pixels; (ii) *Discrimination rate*, which is the probability to identify wrong points as shadow, i.e. the false alarms rate; (iii) *Localization error*, i.e. the average distance between the pixels marked as shadows and the real position of the shadow pixels.

Algorithms to solve the shadow detection problem can be coarsely divided in two groups: *property-based* algorithms and *model-based* algorithms. The most common and flexible are the property-based approaches which use features like geometry, brightness, or color to identify shadowed regions. These techniques do not use any a-priori knowledge as scene geometry, objects disposition and types, or light condition. Instead model based approaches are well suited to particular situations, as car tracking in highway, but have shown less robustness than property-based algorithms when used in a different scene and illumination conditions.

Shadow detection algorithms may be also defined in base of the main property analyzed: geometrical, luminance, color space and difference, texture analysis and edges. These basic components can be combined together to overcome the limitations offered by the methods separately. Moreover the detection of shadow can be used to reconstruct the image, not only to fastener the process. Dark shadows and soft shadows however do not change the phys-

ical dimension of an object. In this paper we propose an approach that takes in consideration the mentioned aspects, together the depth for the shadow.

2. Related Work

Considering the different solutions the authors proposed, we can group the works as described in the introduction.

A familiar direction is to conjecture that shadows reduce the luminance of an image, meanwhile the chrominance stays almost unchanged [20, 14]. However these approach are not valid in outdoor scenes.

Shadow analysis based on color spaces is pretty popular. Cucchiara *et al.* [21] hypothesized that shadows reduce brightness and saturation maintaining hue properties in HSV color space. More recently, Moghimi *et al.* [15], use combination of orthogonal transformation and HSV color space to detect shadow pixels. Yang *et al.* [19] use three components in YUV color space to identify shadow pixels from the candidate foreground. Dong *et al.* [25] analyze the differences between the pixels of object/shadow and that of background according to a RGB color model followed by edge ratios analysis. In [22], the authors compare shadow suppression using RGB and HSV color space and show that HSV color space should be preferred over RGB color space.

Also geometrical information had space on the shadow detection field. Many of the methods in literature normally requires shadows to be on a flat plane. The use of a disparity model has been proposed by Ivanov, *et al.*, [26]. The method described is defined as invariant to an arbitrarily rapid changes in illumination, for modeling background. The negative aspect is that, to overcome rapid changes in illumination, at least three cameras are required. Onoguchi [18] proposed a method based on two cameras to eliminate the shadows of pedestrian based on object height. In order to detect shadow, both object and shadow must be visible on the camera. Salvador, *et al.*, [9] adopt the fact that a shadow darkens the surface, to identify an initial set of shadowed pixels. This particular set is reduced by using color invariance and geometric properties of shadow.

The shadow is not only detected but it is also possible to remove. Finlayson, *et al.*, [12] utilizes shadow edges along with illumination invariant images to recover full color images. Despite that, a part of the color information is lost in removing the effect of the scene illumination. Weiss [24] uses the reflectance edges of the scene to obtain an intrinsic image without shadows. The approach proposed requires significant changes, and as result the scene illumination is contained in the reflectance image. Matsushita, *et al.*, [27] extend the previous concept. However their method does not consider dynamic cast shadows but only static.

Huerta, *et al.*, [13] apply a multi-stage approach combining color, gradient and textural information with known

shadow properties. This method improve previous models but partial loss of foreground borders due to edge and has weakness to texture-less background and objects.

3. Algorithm Overview

The algorithm is outlined in Fig.1. First, a segmentation of moving objects is realized using background subtraction. On the moving object extracted, a pixel-wise analysis of color consistency is performed. Difference between foreground and background distance is computed with the stereo camera. Finally, a shadow measure is computed using the Dempster-Shafer combination rule.

4. Shadow Properties and Models

In this work both color and distance information obtained by a stereo camera are considered. We consider fundamental radiometric model of the radiance of points in a scene illuminated by a combination of sunlight and, if present, colored direct or diffuse light like sky light (combination of multiple light sources). The model and the assumptions that are made are described in this section. We assume that the camera has a linear relationship between the radiance of a surface and the pixel value assigned to the image point of the surface. That type of camera is defined as linear camera.

For this work it is assumed that the images represent a well illuminated environment in both cases, indoor and outdoor. The material of the objects in the scene are essentially diffuse, which exhibit Lambertian reflectance, constant over time. Instead albedo's (diffuse reflectance) of the surfaces is not necessary to be constant over time. The images is supposed to be properly exposed, i.e. the important area of the image are neither over-exposed (color channel values near 255) not severely under-exposed (value near 0).

4.1. Color Model

It is supposed that the color information ρ at a given pixel p obtained from a recording camera depends on four components: the spectral power distribution (SPD) of the illuminant denoted $E(\lambda)$, the surface reflectance $R(\lambda)$, the sensor spectral sensitivity $Q(\lambda)$ evaluated at each pixel p and a shading factor σ . This assumption is valid for Lambertian surfaces.

$$\rho_p = \sigma \int R(\lambda)E(\lambda)Q_p(\lambda)d(\lambda) \quad (1)$$

The surface reflectance $R(\lambda)$ depends on the material which can have different albedo.

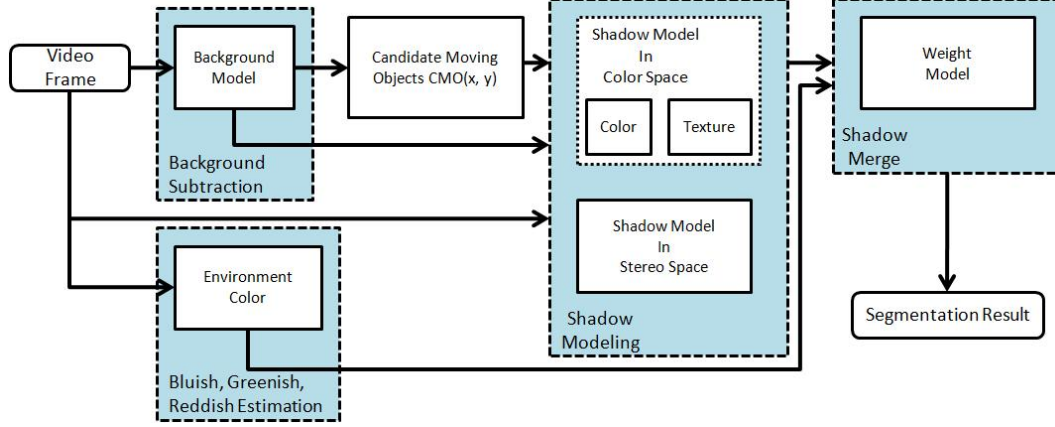


Figure 1. Flow diagram of the proposed method.

In a setting as described above it is possible to formulate the value, ρ , of a pixel as follows, using subscript r to indicate elements related to the red channel (green and blue being similar):

$$\rho_r = \frac{c_r \cdot \varphi_r \cdot E_r}{\pi} \quad (2)$$

where φ_r is the diffuse albedo of the surface point being imaged (ratio of outgoing radiosity to incoming radiance), and E_r is the incoming radiance in the red channel. Thus, $\varphi_r \cdot E_r$ is the reflected radiosity. Dividing this by π [sr] yields the reflected radiance of the surface (since the radiosity from a diffuse surface is π times the radiance of the surface). Finally, c_r is the (typically unknown) scaling factor translating the measured radiance into pixel value (0 to 255 range for an 8 bit camera) for a linear camera. This scaling value depends on the aperture of the lens, the shutter speed, the gain, the white-balancing etc. of the camera.

In the kind of outdoor daylight setting we are addressing in this paper the total incoming radiance at a point is a sum of two contributions, $E_r = E_r^{SUN} + E_r^{SKY}$, again using subscript r for red color channel as example. The amount of radiance received from the sun, E_r^{SUN} , depends on several factors: the radiance of the sun, how large a fraction of the sun's disk is visible from the point in interest (if the sun's disk is completely occluded the point is full shadow, also called umbra), and on the angle between the surface normal at the point and the direction vector to the sun from the point. If the sun's disk is only partially occluded the point is in the penumbra (soft shadow).

Two kinds of shadows can appear in an image: the penumbra and the umbra. The difference between them can be modelled by the following equation

$$\rho(x, y) = E(x, y)\varphi(x, y) \quad (3)$$

4.2. Reddish, Greenish and Bluish

Many works took in consideration the diffuse sources. In particular in outdoor scenes, the diffuse source it is obviously the sky. Diffuse source has different values of total incoming radiance at a point $E(a)$. A no white diffuse source can have effect to the cast shadow. If we consider an outdoor environment, beside a reduction in the intensity, an outdoor cast shadow will result in a change of chrominance. Considering again the outdoor scene, the illumination of the sky has higher power components in the lower wavelengths λ (450-495nm) of the visible spectrum, and it is therefore assumed bluish as argued in [17]. When the direct illumination of the sun is blocked and a region is only illuminated by the diffuse ambient light of the sky, materials appears to be more bluish. This "bluish effect" and the chrominance distortion can be exploited for shadow detection and grouping of potential shadow pixel. Colour balance, or in the specific case, white balance can apply an adjustment of the intensities of colours. Even with this compensation, the effect of coloured diffuse sources remain on the objects, and it is possible to take an advantage on it, when possible. Because, in order to detect the shadow, the variation between a background image and current image is estimated, coloured effect can be detected. Bluish however is just one large but particular condition. In fact it is possible that, due to the particular scenario, the diffuse light has a different colouration. For instance, lights in coloured environment and also coloured lights which cause a chrominance distortion. If it exists, then it is possible to consider a general chrominance distortion in one of the three channels. Thus, objects which suffer more environment color will have an intensity variation bigger on the component which does not represent the color. For example, in outdoor scenario, the shadow will suffer more the effect of the sky and the intensity changes will be bigger in red and green channels than in blue chan-

nel.

5. An Innovative Information Fusion Approach for Supporting Shadow Detection in Complex Environments

As described in Section III we initially segment each frame into background, foreground, and shadow. This is performed by combining the results from background subtraction process and shadow detection. The methods are described below. The advantage of these methods is that they do not require a training phase. Nevertheless they give generally good performances as it will be shown in the experimental Section.

5.1. Background Subtraction

In this work we use a subtraction stereo method described in [11]. This method consist in a threshold subtraction .

For each pixel p , a background model is learned, from which the foreground probability can be estimated. Potential moving objects can be extracted by simply thresholding this density distribution, and within the segmentation the cast shadows can be evaluate over both the color and stereo domain.

Another drawback of using background subtraction is that for long sequence (day s, week s, months), it can be difficult to maintain the background model due to high varying illumination, precipitation, season changes, etc.

5.2. Color Shadow Detection

The method we are going to describe is a modification of the works described in [16] and [10]. It offers interesting performance in shadow detection without any training phase.

Considering that the radiance influences the cast shadow not linearly than the intensity ratios between neighboring shadow pixels depends on the source direction. The variations in background and segmented image will be used to calculate the error score within a small region, used for discriminating a pixel as shadow. The error score is computed as reported in eq.(4).

$$\Psi(x, y) = \sum_{c \in R, G, B} \sum_{i, j \in \omega(x, y)} |d_c(i, j) - d'_c(i, j)| \quad (4)$$

Even if colors cannot be used singularly to extract shadows, they represent an important source of information. The color difference which is used to estimate if a pixel belongs

to a cone shadow is calculated comparing the color information between the background and detected foreground image. The color space is modelized as follows

$$\begin{cases} C_1(x, y) = \arctan\left(\frac{I_r(x, y)}{I_b(x, y)}\right) \\ C_2(x, y) = \arctan\left(\frac{I_g(x, y)}{I_b(x, y)}\right) \end{cases} \quad (5)$$

The color score error due to the variation of color is computed as

$$\Lambda(x, y) = |C_1(x, y) - C'_1(x, y) + C_2(x, y) - C'_2(x, y)| \quad (6)$$

5.3. Stereo Shadow Detection

As previously mentioned, in ideal condition, it would be possible to separate an object from the shadow by the estimated distance. Once obtained the distance value of each point of the image from the background, a pixel is considered object if the difference is higher than zero (ideal condition), or ϵ (real condition).

Because the calculation of distance generally is problematic in saturation regions, only the points with an intensity less than a threshold are used.

Stereo cameras have a range within the error measures calculated which is lower than the measured distance. The working range changes from camera to camera and it depends on several factors, software (matching algorithms) and hardware (sensors, camera displacement and numbers, focal length).

The estimated errors in Fig. 2 shows a quadratic relation between distance and error. According to our tests, accurate values can be obtained within 5 meters.

We took in consideration the Bumblebee2 stereo camera and estimated the errors due to the range measuring the real distance and estimated distance. The error is shown in Fig. 2 as vertical segments.

If the measure is taken after the best distance, the error shows a quadratic behaviour. Otherwise it has a quasi-linear behaviour.

If the distance information has been calculated both for background and foreground and it is within the best range, then if $d_{x,y} \neq \emptyset \wedge d'_{x,y} \neq \emptyset \wedge d_{x,y} \leq \epsilon$

$$S_{x,y} = \begin{cases} 0 & \text{if } |d_{x,y} - d'_{x,y}| \leq m(d_{x,y} - \epsilon) + q \\ 1 & \text{otherwise.} \end{cases} \quad (7)$$

If the distance information has been calculated both for background and foreground but it is over the best range, then if $d_{x,y} \neq \emptyset \wedge d'_{x,y} \neq \emptyset \wedge d_{x,y} > \epsilon$

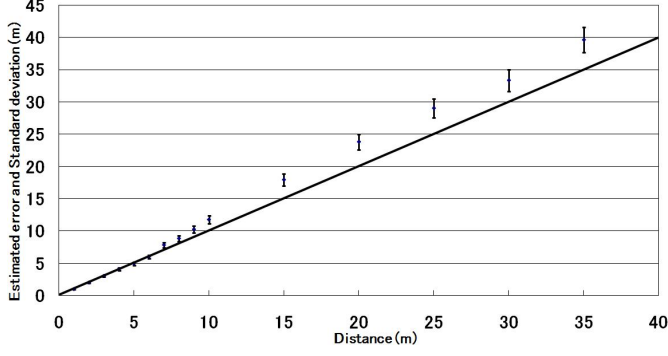


Figure 2. Error calculated for the stereo camera used. The error increases with the distance.

$$S_{x,y} = \begin{cases} 0 & \text{if } |d_{x,y} - d'_{x,y}| \leq \frac{1}{m} (d_{x,y} - \epsilon)^2 + q \\ 1 & \text{otherwise.} \end{cases} \quad (8)$$

From the hypothesis that the camera is able to estimate a correct distance within a certain range r , we propose a method to automatically adjust m and q . Given the distance difference between background and current images, it is reasonable to suppose that within a certain confidence range r' , the pixels have to be labeled as shadow. Instead out of r' , pixels are expected to belong to a moving object. A set of points measured around r are collected and labeled based on the previous consideration. If a pixel is recognized as *shadow* and the distance difference is lower than r' then the point is marked as *success* otherwise as *failure*. If a pixel is recognized as *no shadow* and the distance difference is equal or greater than r' , then the point is marked as *success* otherwise as *failure*. An example can be seen in fig.3.

Once obtained the graph, we modify the parameters only if the percentage of success inside and outside the range r' is lower than an accuracy value.

At each iteration, the parameters are adjusted using a random function as follow:

$$m = m + \text{sign} \cdot \text{rand} \quad (9)$$

$$q = q + \text{sign} \cdot \frac{\text{rand}}{100} \quad (10)$$

where *sign* is positive if the percentage of failures inside the range is higher or equal to the percentage of failures outside. Negative in opposite case.

5.4. Dempster-Shafer Fusion

The Basic Belief Assignment can be viewed as a generalization of a probability density function. More precisely,

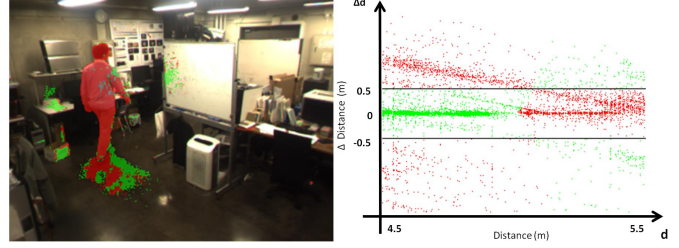


Figure 3. Example of success and failure in stereo shadow detection. In the left picture, the pixels are depicted in red or green to put in evidence if it is shadow (green) or foreground (red). On the right, the graph shows the estimation success. Each colored pixel represent a case of success or failure. In red, a pixel is labeled as estimated. In green, a failure.

a Basic Belief Assignment $m(\cdot)$ is a function that assigns a value in $[0, 1]$ to every subset \mathcal{A} of θ that satisfies the following:

$$\sum_{\mathcal{A} \subseteq \theta} m(\mathcal{A}) = 1, \quad m(\emptyset) = 0$$

It is worth noting that $m(\mathcal{A})$ is the belief that supports the subset \mathcal{A} of θ , not the elements of \mathcal{A} . This reflects some ignorance because this means that we can assign belief only to subsets of θ , not to the individual hypothesis as in classical probability theory.

Consider two Basic Belief Assignments $m_1(\cdot)$ and $m_2(\cdot)$ and the corresponding belief functions $bel_1(\cdot)$ and $bel_2(\cdot)$. Let \mathcal{A}_j and \mathcal{B}_k be subsets of θ . Then $m_1(\cdot)$ and $m_2(\cdot)$ can be combined to obtain the belief mass assigned to $\mathcal{C} \subset \theta$ according to the following formula [23]:

$$m(\mathcal{C}) = m_1 \oplus m_2 = \frac{\sum_{j,k, \mathcal{A}_j \cap \mathcal{B}_k = \mathcal{C}} m_1(\mathcal{A}_j) m_2(\mathcal{B}_k)}{1 - \sum_{j,k, \mathcal{A}_j \cap \mathcal{B}_k = \emptyset} m_1(\mathcal{A}_j) m_2(\mathcal{B}_k)} \quad (11)$$

The denominator is a normalizing factor, which measures how much $m_1(\cdot)$ and $m_2(\cdot)$ are conflicting.

5.5. Basic Belief Assignment for Shadow Estimation

The quantities above computed, i.e. $\Psi(x, y)$, $\Lambda(x, y)$ and $S(x, y)$ can be considered as the outputs of three experts that, from different knowledge, represent the possibility that the pixel (x, y) is a shadow. More precisely, the more each of these parameters is close to zero, the more likely the relative pixel represents a shadow. From such quantities, three other quantities are computed as follows: $\Psi'(x, y) = \max((1 - \Psi(x, y)), 0)$, $\Lambda(x, y) = \max((1 -$

$\Lambda(x, y), 0), S'(x, y) = (1 - S(x, y))$. These three quantities are then normalized such that their sum is equal to one. The next step we performed is to divide, with empirical thresholds, each quantity in three sections, corresponding to Umbra, Penumbra and Luminance respectively. Since the $S'(x, y)$ parameter given that the $S'(x, y)$ parameter distinguishes only the shadow by the light, the probability that the pixel represents umbra or penumbra is the same. Thus, we considered the following set of possible hypotheses, $\theta = \{U, P, L\}$ that is Umbra, Penumbra, Light. This give rise to the following power set: $\{U\}, \{P\}, \{L\}, \{UP\}, \{UL\}, \{PL\}, \{UPL\}$. Each subset is assigned a belief m on the basis of the three knowledge sources. The final shadow index Θ is obtained by combination:

$$\Theta = \bigoplus_{i=1}^n m_i = ((m_1 \oplus m_2) \oplus m_3)$$

6. Confidence Estimation

In order to estimate the diffuse chrominance of the scene, we propose a method which analyzes the variation in a sequence of images. If I_c is the intensity of the current frame and I'_c is the intensity of a given background image, the diffuse chrominance is considered the highest variation in a sequence. c can assume red, green, or blue value, and *gray* is equal to gray value. We consider that each channel has 8 bit resolution. Even if it is obvious that colored objects influence the estimation of the color, considering all the points of the image in a sequence will reduce that effect.

Given the background image and the current image, the histogram of the differences is computed as

$$\forall_p, H_{c, |I_p - I'_p|} = H_{c, |I_p - I'_p|} + 1 \quad (12)$$

The color value is then

$$cv_c = \frac{\sum_{i=0}^{256} H_{c,i} \cdot i}{s} \quad (13)$$

The difference in intensity between the gray scale image in each channel is computed as reported in eq.(14).

$$DI_c = \frac{1 - |cv_c - cv_{gray}|}{256} \quad (14)$$

and the proportional variation of each channel respect the gray scale image is computed as

$$PVI_c = \begin{cases} - \left(1 - \min \left(\frac{cv_c}{cv_{gray}}, \frac{cv_{gray}}{cv_c} \right) \right) & \text{if } cv_c > cv_{gray} \\ 1 - \min \left(\frac{cv_c}{cv_{gray}}, \frac{cv_{gray}}{cv_c} \right) & \text{otherwise.} \end{cases} \quad (15)$$

Differences and proportional variations are gathered for the number of frames necessary to estimate the average and variance of the sequence analyzed. Empirically we estimated that 100 frames are sufficient for that analysis.

First the DI are normalize respect the maximum and minimum value of all the color DI .

$$DI_c = \frac{DI_c - \min(DI)}{\max(DI) - \min(DI)} \quad (16)$$

We compute the average and variance of the differences and proportional variations. The sequence average difference and variation are estimated as shown in eq.(17) and eq.(18).

$$SAD_c = \ln \left(\frac{\mu DI_c}{1 - \delta^2 DI_c} \right) \quad (17)$$

$$SAPV_c = \left| \ln \left(\frac{\mu PVI_c}{1 - \delta^2 PVI_c} \right) \right| \quad (18)$$

For each combination of colors the difference of sequence average difference and variation is computed in order to estimate the prevailing color.

$$\begin{aligned} \Delta_{rg} &= SAD_r - SAD_g \\ \Delta_{rb} &= SAD_r - SAD_b \\ \Delta_{gb} &= SAD_g - SAD_b \end{aligned} \quad (19)$$

and finally the color strenght is estimated as described in eq.(20).

$$\begin{aligned} CS_r &= (\Delta_{rg} \geq 0) \cdot |\Delta_{rg}| + (\Delta_{rb} \geq 0) \cdot |\Delta_{rb}| \\ CS_g &= (\Delta_{rg} < 0) \cdot |\Delta_{rg}| + (\Delta_{gb} \geq 0) \cdot |\Delta_{gb}| \\ CS_b &= (\Delta_{rb} < 0) \cdot |\Delta_{rb}| + (\Delta_{gb} < 0) \cdot |\Delta_{gb}| \end{aligned} \quad (20)$$

The suggested best chrominance will be the highest value.

Obviously, one color will prevails, unless the image is not completely gray scale. Because we want to avoid errors due to noisy or particular configurations, we consider colored diffuse light only the best color value which satisfies the following equation.

$$color = \begin{cases} white & \text{if } CS_{best} < \min(SAPV) \\ best & \text{otherwise.} \end{cases} \quad (21)$$

We consider that a pixel segmented as foreground cannot be a shadowed pixel if its intensity is higher than background. Thus, a pixel is candidate as shadow if

$$sp_a = (I_a^R < \mu^R) \wedge (I_a^G < \mu^G) \wedge (I_a^B < \mu^G) \quad (22)$$

Moreover, in the case of bluish effect (similar to greenish and reddish), the changes on the intensity component of the red and blue channels are bigger than the blue channel. To

be more flexible, the increment will be proportional. This fact can be used to reduce the shadow region as follows:

If bluish

$$cs_a = (k(I_a^R - \mu^R) > (I_a^B - \mu^B) \wedge k(I_a^G - \mu^G) > (I_a^B - \mu^B)) \wedge sp_a \quad (23)$$

If reddish

$$cs_a = (k(I_a^G - \mu^G) > (I_a^R - \mu^R) \wedge k(I_a^B - \mu^B) > (I_a^R - \mu^R)) \wedge sp_a \quad (24)$$

If greenish

$$cs_a = (k(I_a^R - \mu^R) > (I_a^G - \mu^G) \wedge k(I_a^B - \mu^B) > (I_a^G - \mu^G)) \wedge sp_a \quad (25)$$

The mask so obtained is then used to mark the pixels that should be analyzed.

If the color of the environment is not white the pixels which are not candidate to be shadow are labeled as moving objects.

6.1. Merge Methods

Once the shadow values are obtained with the previously described methods, the difficult task is to find a relationship between result obtained from color information and shadow information. Because the measures are not directly relationable, we estimate the strength of each detection method.

The shadow detection method described in section V.A return a value we called shadow parameter. It is possible to combine the shadow parameter with the distance in order to obtain a *confidence* value. Likewise, it is possible to obtain a confidence value from the stereo shadow detection. The confidence values are higher for the stereo information if the distance of a point from the camera, is near the focal point. On the opposite side, confidence value is higher for the points which have a distance lower or higher than the focal point.

The curve that define the probability to be shadow is not centered around zero, but differs if a pixel is detected as shadow or no shadow. The equations can be so resumed. If the distance from the camera is less or equal to the focal point:

$$W_{c,x,y} = \begin{cases} pShadow \cdot \left(\frac{d}{m} + q \right) & \text{if shadow} \\ \frac{Fp}{pShadow} \cdot \left(\frac{d}{m} + q \right) & \text{otherwise.} \end{cases} \quad (26)$$

$$W_{s,x,y} = \begin{cases} \frac{1}{\Delta_d} \cdot \frac{1}{\frac{d}{m} + q} & \text{if shadow} \\ 2\Delta_d \cdot \frac{1}{\frac{d}{m} + q} & \text{otherwise.} \end{cases} \quad (27)$$

However, if the distance from the camera increases:

$$W_{c,x,y} = \begin{cases} pShadow \cdot \left(\frac{1}{m} \cdot (d - Fp)^2 + q \right) & \text{if shadow} \\ \frac{Fp}{pShadow} \cdot \left(\frac{1}{m} \cdot (d - Fp)^2 + q \right) & \text{otherwise.} \end{cases} \quad (28)$$

$$W_{s,x,y} = \begin{cases} \frac{1}{\Delta_d} \cdot \frac{1}{\frac{1}{m} \cdot (d - Fp)^2 + q} & \text{if shadow} \\ 2\Delta_d \cdot \frac{1}{\frac{1}{m} \cdot (d - Fp)^2 + q} & \text{otherwise.} \end{cases} \quad (29)$$

Stereo information may be not always available in all the points of the image. This is due to several factors: light conditions, distance of the object from the camera, visibility of an object from both the cameras. If the stereo information is not available, we choose to use only the color information. In the case a pixel is labeled as shadow (or no shadow) with both the methods, the pixel will be labeled with the detected value. If the detection value is different, the pixel will be labeled with the value of the method which have higher weight.

7. Qualitative Evaluation

The proposed algorithm has been implemented using a Bumblebee2 stereo camera from Point Grey Research on an Intel Quad CPU at 2.83 GHz and 4GB Ram. Unfortunately we did not find a 3D data set suitable for us because we need results obtained with that stereo camera. Thus we acquired several video sequences in different conditions of illumination and camera orientation. Both in-door and outdoor environments have been considered. The results presented hereafter are obtained from our captured sequences. The results presented here are only qualitative; in fact we did not performed actual comparisons with ground truths yet but we want only to evaluate the results from a qualitative point of view, for now. The results appear to us quite good, as shown in Fig.4) and Fig.5).

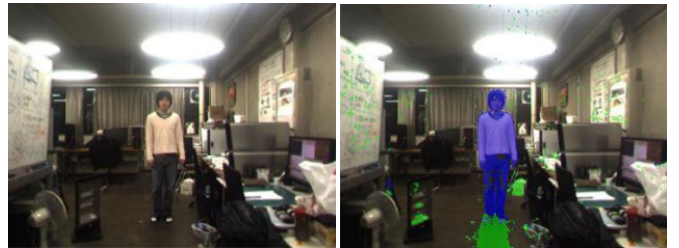


Figure 4. In this figure an example of shadow detection performed with the proposed algorithm for an in-door environment is reported. Shadow pixels are represented in green.

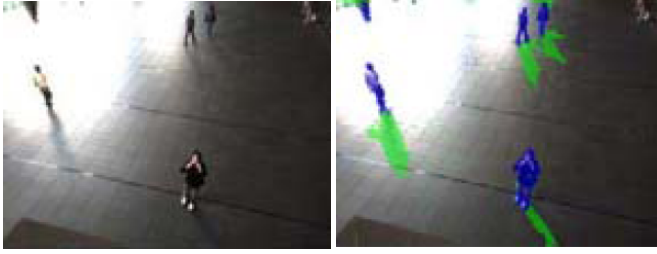


Figure 5. In this figure an example of shadow detection performed with the proposed algorithm for an out-door environment is reported. Shadow pixels are represented in green.

8. Conclusions and Future Work

In this paper, a new moving cast shadow detection algorithm that requires a stereo camera is proposed. The algorithm exploits the color, texture, temporal and depth information. Although many shadow detections have been performed, in this paper only the theory and some examples are reported. Accurate ground-truth based performances and comparisons with state of the art algorithms will be presented in future papers. Also a GPU implementation of the algorithm will be explored. Future work is mainly focused in extending our proposed framework by means of several characteristics, as to enhance it significantly. For instance, some interesting properties to be investigated in future are: *fragmentation* (e.g., [1, 3]), *approximation* (e.g., [4, 5]), *privacy preservation* (e.g., [6, 7]), *big data* (e.g., [8, 2]).

References

- [1] A. Bonifati and A. Cuzzocrea. Efficient fragmentation of large XML documents. In *DEXA 2007, Regensburg, Germany, September 3-7, 2007*, pages 539–550, 2007.
- [2] A. Cuzzocrea, L. Bellatreche, and I. Song. Data warehousing and OLAP over big data: current challenges and future research directions. In *DOLAP 2013, San Francisco, CA, USA, October 28, 2013*, pages 67–70, 2013.
- [3] A. Cuzzocrea, J. Darmon, and H. Mahboubi. Fragmenting very large XML data warehouses via k-means clustering algorithm. *IJBIDM*, 4(3/4):301–328, 2009.
- [4] A. Cuzzocrea, F. Furfaro, S. Greco, E. Masciari, G. M. Mazzeo, and D. Saccà. A distributed system for answering range queries on sensor network data. In *3rd IEEE (PerCom 2005 Workshops), 8-12 March 2005, Kauai Island, HI, USA*, pages 369–373, 2005.
- [5] A. Cuzzocrea and U. Matrangolo. Analytical synopses for approximate query answering in OLAP environments. In *DEXA 2004 Zaragoza, Spain, August 30-September 3, 2004*, pages 359–370, 2004.
- [6] A. Cuzzocrea, V. Russo, and D. Saccà. A robust sampling-based framework for privacy preserving OLAP. In *DaWaK 2008, Turin, Italy, September 2-5, 2008*, pages 97–114, 2008.
- [7] A. Cuzzocrea and D. Saccà. Balancing accuracy and privacy of OLAP aggregations on data cubes. In *DOLAP 2010, Toronto, Ontario, Canada, October 30, 2010*, pages 93–98, 2010.
- [8] A. Cuzzocrea, D. Saccà, and J. D. Ullman. Big data: a research agenda. In *IDEAS '13, Barcelona, Spain - October 09 - 11, 2013*, pages 198–203, 2013.
- [9] A. C. E. Salvador and T. Ebrahimi. Cast shadow segmentation using invariant color features. *CVIU*, 95(2):238–259, 2004.
- [10] A. M. et al. Auto-adaptive threshold and shadow detection approaches for pedestrian detection. *AWSVCI*, pages 9–12, 2009.
- [11] K. U. et al. Subtraction stereo. a stereo camera system that focuses on moving regions. *Proc. Of SPIE-IS T Electronic Imaging.*, 7239 Three Dimensional Imaging Metrology, 2009.
- [12] C. L. G. Finlayson, S. Hordley and M. Drew. On the removal of shadows from images. *IEEE TPAMI*, 28(1):59–68, 2006.
- [13] T. M. I. Huerta, M. Holte and J. González. Detection and removal of chromatic moving shadows in surveillance scenarios. *ICCV*, pages 1499–1506, 2009.
- [14] D. H. K. Kim, T. Chalidabhongse and L. Davis. Real-time foreground-background segmentation using codebook model. *Real-Time Imaging*, 11(3):172–185, 2005.
- [15] H. P. M.K. Moghimi. Shadow detection based on combinations of hsv color space and orthogonal transformation in surveillance videos. In *Proceedings of Iranian Conference on Intelligent Systems*.
- [16] C. C. M.T. Yang, K.H. Lo and W. Tai. Moving cast shadow detection by exploiting multiple cues. *Image Processing.*, 2:95–104, 2008.
- [17] S. Nadimi and B. Bhanu. Physical models for moving shadow and object detection in video. *IEEE TPAMI*, 26(8):1079–1087, 2007.
- [18] K. Onoguchi. Shadow elimination method for moving object detection. *ICPR*, 1:583–587, 1998.
- [19] N. A. Q. Yang, K.H. Tan. Shadow removal using bilateral filtering. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 2012.
- [20] M. P. R. Cucchiara, C. Grana and A. Prati. Detecting moving objects, ghosts, and shadows in video streams. *IEEE TPAMI*, 25(10):1337–1342, 2003.
- [21] M. P. A. P. S. S. R. Cucchiara, C. Grana. Improving shadow suppression in moving object detection with hsv color information. *C. A.F. editor, Proc. IEEE ITSC*, pages 334–339, 2001.
- [22] R. K. S. Surkutlawar. Shadow suppression using rgb and hsv color space in moving object detection. *International Journal of Advanced Computer Science and Applications*, 2013.
- [23] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976.
- [24] Y. Weiss. Deriving intrinsic images from image. *ICCV*, 2:68–75, 2001.
- [25] G. J. X. Dong, K. Wang. Moving object and shadow detection based on rgb color space and edge ratio. In *Proceedings of CISIP'09*.
- [26] A. B. Y. Ivanov and J. Liu. Fast lighting independent background subtraction. *IJCV*, 37(2):199–207, 2000.

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