applied sciences

MDPI

Article

Detection and Removal of Moving Object Shadows Using Geometry and Color Information for Indoor Video Streams

Akmalbek Abdusalomov [1] and Taeg Keun Whangbo [2,*]

[1] Department of IT Convergence Engineering, Gachon University, Sujeong-Gu, Seongnam-Si, Gyeonggi-Do 461-701, Korea; akmalbekabdusalomov@gmail.com

[2] Department of Computer Science, Gachon University, Sujeong-Gu, Seongnam-Si, Gyeonggi-Do 461-701, Korea

* Correspondence: tkwhangbo@gachon.ac.kr

Received: 12 November 2019; Accepted: 26 November 2019; Published: 28 November 2019

Abstract: The detection and removal of moving object shadows is a challenging issue. In this article, we propose a new approach for accurately removing shadows on modern buildings in the presence of a moving object in the scene. Our approach is capable of achieving good performance when addressing multiple shadow problems, by reducing background surface similarity and ghost artifacts. First, a combined contrast enhancement technique is applied to the input frame sequences to produce high-quality output images for indoor surroundings with an artificial light source. After obtaining suitable enhanced images, segmentation and noise removal filtering are applied to create a foreground mask of the possible candidate moving object shadow regions. Subsequently, geometry and color information are utilized to remove detected shadow pixels that incorrectly include the foreground mask. Here, experiments show that our method correctly detects and removes shadowed pixels in object tracking tasks, such as in universities, department stores, or several indoor sports games.

Keywords: color information; median filter; luminance; image enhancement; moving shadow detection; ghosting artifacts

1. Introduction

Moving shadow detection methods have been studied extensively by researchers [1]. They are profitable in various fields of computer vision application, such as in object tracking and foreground recognition. Currently, there are numerous frameworks and hardware systems for obtaining image sequences without any moving shadows or ghosts [2]. However, a variety of problems exist, such as background similarity and multiple shadows, which may make the shadow detection task difficult. We obtained video sequences to experiment with removing shadowed images. Moving foregrounds have become more accessible and can often show up in various positions in neighboring images, causing multiple objects to be detected as a single foreground blob [3].

Shadow pixels are occasionally misclassified as parts of objects, causing mistakes in localization, segmentation, and the tracking of moving objects. Often, the shadows give significant data, for example, the relative position of an object from a source. Shadows can vary from the background surface in luminance and gray level, like the foreground objects in the scene. Furthermore, the utilization of picture shadow detection methods with minimally controlled equipment in real-time frameworks is a difficult issue [4], as the vast majority of the shadow removal techniques are computationally costly. Thus, a shadow removing method on a computer must be customized to fulfill the real-time requirements of video surveillance, with limited computational resources. This current study aims to address moving shadow detection problems, as mentioned above and in the following paragraph.

The main contribution of this manuscript is that we present a simple, reliable, and automatic shadow removal method that is robust against background surface similarity and ghost problems. The method requires a small amount of computational time to detect the presence of moving objects for indoor video surveillance and is based on geometry features. Existing geometry-based methods do not adequately process shadow removal with objects having multiple shadows or where several objects are recognized as a single foreground mask. Our suggested algorithm is comprised of the following steps:

- Enhancing input images based on combined contrast enhancement.
- Extracting moving objects using background subtraction and removing unwanted noises from the enhanced images.
- Detecting and removing shadow pixels that are correctly included in the candidate foreground mask.
- Applying a morphological reconstruction method to eliminate small gaps and holes from the moving object regions.
- Delivering the final result (without shadows) to the object tracking tasks.

The rest of the paper is partitioned into section as follows: Section 2 shows a survey of some related shadow detection strategies. Section 3 provides a brief discussion of the difference between static and dynamic shadows. Section 4 gives a definite discourse on our proposed moving shadow removing method for dealing with moving object tracking issues. Section 5 demonstrates our test results and provides discussion based on the most widely-used databases. Section 6 provides conclusions based on the experimental results and future directions for study.

2. Literature Review

Moving shadow detection has been investigated for years, and many researchers and scientists are working together in the shadow removing domain to reduce processing time and improve the quality of the segmentation result to deliver appropriate object tracking applications. However, shadow detection remains one of the most important and challenging issues in the areas of computer vision, object detection, and machine learning. Detecting shadow regions by the human eye may be a somewhat easy task, but it is a relatively challenging problem for a computer, as shadow pixels also simultaneously move as an object region. For these reasons, most contemporary studies concern detecting and removing shadows.

Yanli Wan et al. [5] introduced a shadow removing technique for moving objects to eliminate ghosting artifacts. In their approach, the ghosting area is rearranged to avoid removing the moving shadow pixels in the scene. However, this method is relatively difficult to use in urban surveillance and in multiple noise environments. Cucchiara et al. [6] utilized shadow features in the hue, saturation, value (HSV) color space to recognize shadow pixels where the object of interest is in motion. These properties demonstrate that cast shadows obscure the background in the luminance component, whereas the saturation and hue spaces change inside specific limits. The HSV color space was utilized because it provides a superior separation of chromaticity and grey level than other color spaces. In [7], the authors reviewed several shadow detection methods, each of which proved their efficiency in detecting and removing shadow pixels in indoor and outdoor environments. Several other research works were proposed for dynamic image sequences [8–10]. The scientists investigated the concept of including a multi-frame differencing system to enhance the division in situations where the shadows may not be effectively removed. Stauder et al. [11] suggested a new physics-based method that used luminance and intensity values to describe physical illumination changes.

Recently published articles are different in terms of productivity and reliability [12,13]. They try to overcome the modern problems in moving shadow detection, such as those in smart city and intelligence-building frameworks. Today, applying computer vision tasks to machine learning and neural networks is becoming a very important research area. Dong Seop Kim et al. [14] employed a convolutional neural network (CNN) in their study, for shadow detection in images using a visible light camera sensor. The researchers presented a shadow detection and removal algorithm that used

a 21 × 21 sliding window-based visual geometry group (VGG) "Net-16" CNN and showed a high accuracy, even in a high-definition surveillance condition. A new method for dynamic object and shadow detection based on motion prediction was proposed by Jong Taek Lee et al. [15], solving the shadow problem by using deep learning. In addition, applying a Markov random field enables a system to refine shadow detection results to improve its performance. In [16], a novel approach is presented for versatile shadow removal by consolidating four distinct filters in a neuro-fuzzy structure. The neuro-fuzzy classifier has the capacity for real-time self-adaptation and training, and its execution has been quantitatively surveyed with both indoor and outdoor video streams.

3. Moving Shadow Detection

Recently, various shadow identification techniques have been announced in science-related literature. They can be divided into two areas. The first area generally concerns static images, while the second concerns image sequences and specifically video contents [17]. Static shadows are shadows cast by immobile objects such as buildings, parked vehicles, and trees. In that regard, moving object identification methods do not suffer from static shadows because these shadows are classified as a piece of background. In contrast, dynamic shadows, the subject of interest for this manuscript, are harmful to moving object recognition algorithms. Shadows can be smoothly stitched to an object in action or can be disconnected from it. In the first case, the shadows usually cause the shape of the original object to look different, making the utilization of subsequent shape recognition strategies less reliable. In the second case, the shadows might be incorrectly categorized as an object in the scene. The work performed in this study concerns the second case, where, in this way, it addresses the issue of the detection and removal of moving shadows cast from objects in video surveillance. This is done to improve the process of moving object identification. Moving shadows are often recognized as foreground objects, and this degrades the expected execution of object tracking and accurate segmentation, as depicted in Figure 1.

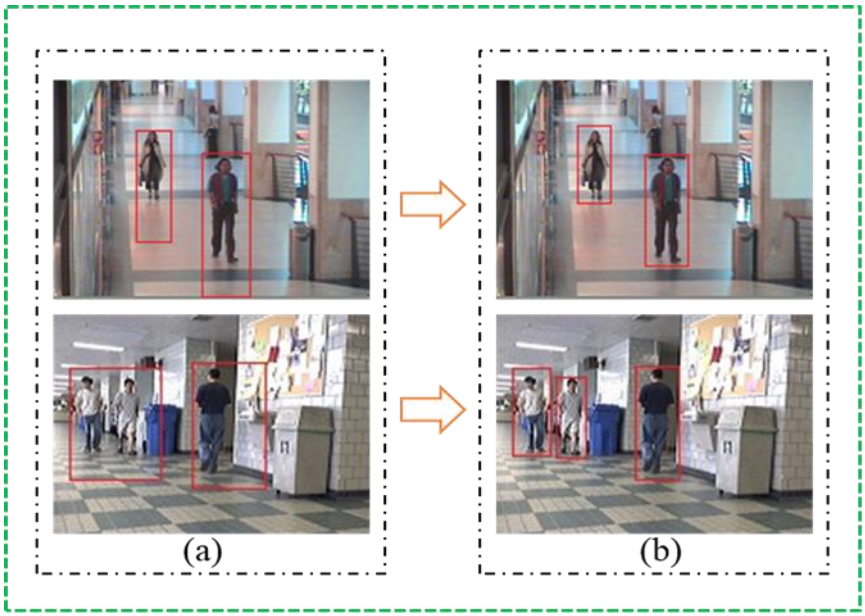

Figure 1. Tracking process on indoor video surveillance: (**a**) Object tracking with shadow region. (**b**) Object tracking without shadow region.

4. Proposed Method

In this section, we discuss our proposed method in detail. Figure 2 shows all steps of the method for detecting and removing a moving shadow.

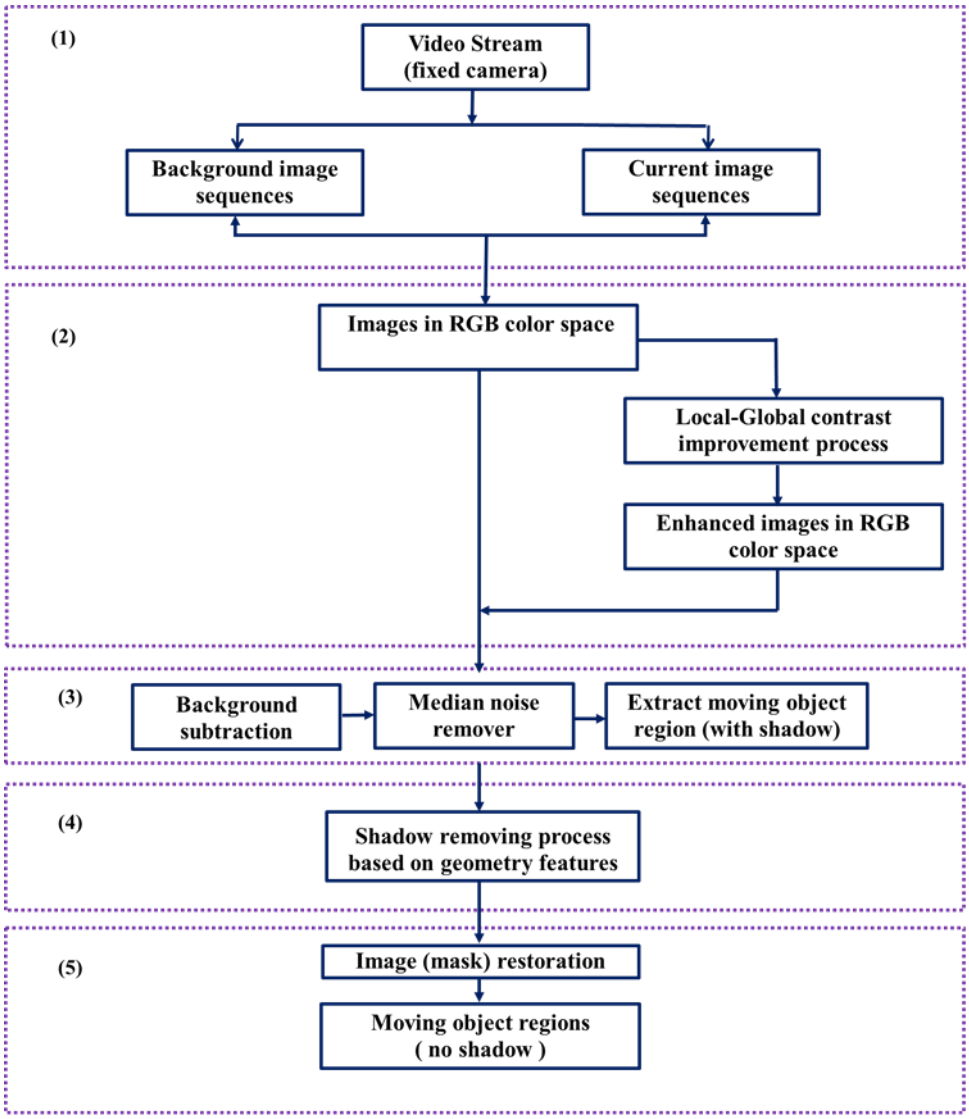

Figure 2. Flowchart of the proposed method: (**1**) Input image sequences, (**2**) image enhancement based on local and global methods, (**3**) object segmentation and noise elimination process, (**4**) shadow removal based on geometry features, (**5**) foreground mask reconstruction for object tracking.

In our proposed shadow removal method, a number of strategies are performed to achieve our goal. To make our method work in real-time, an image frame that can be read from a camera each second is delivered to the system. In the following subsections, we discuss all of the steps of the method.

4.1. Combined Local and Global Contrast Enhancement

As shown in Figure 2, we first utilize an image enhancement technique to improve the image quality and increment the color intensity of significant objects as compared to surrounding regions. In [18], we implemented a global color contrast improvement tool to upgrade shadowy scenery and low-level pixel intensities caused by non-natural illumination sources. In this study, we employed a robust and unique feature enhancement strategy, utilizing both local and global contrast improvement information for reducing the time consumption and generating better outcomes as compared to other image improvement techniques. When applying only one method, it not possible to enhance an input image using its local brightness features, as gray levels having a very high frequency dominate gray levels having a very low frequency.

Therefore, in our proposed method, we utilized a combination of global contrast enhancement and local contrast improvement techniques for such types of data, i.e., those that exist in the less dynamic region of the image, as in [19]. This is done by subtracting two consequent grayscale values of the image.

$$O = I + \frac{M + C_i}{\sigma_i}[m_i - m_{i+1}] \quad (1)$$

in the above, O is a pixel's intensity value in the improved image, corresponding to the I pixel's gray value in the original image. M is the global contrast gain control, with a range of [0–1.5]. C_i is the local contrast control, with a range of 0–1. σ_i is the local standard deviation of the window. m_i is the pixel's grayscale value present in the original image and m_{i+1} is the subsequent pixel's gray scale value in the original image. This method iterates pixel-by-pixel over the entire image.

4.2. Moving Object Extraction

In this step, we extract the foreground moving object out of the enhanced image, based on an existing background subtraction scheme [20]. The background differences method gives a foreground mask in each new frame. However, this approach fails to distinguish between an object and its shadow.

First, the input video is captured as image sequences. An initial frame, which contains no moving object, is derived as a background reference image. The succeeding frames are compared with this frame to detect the moving object. This is calculated as shown in Equation (2).

$$S_n(x, y) = \left| I_n(x, y) - B_n(x, y) \right| \quad (2)$$

where $S_n(x, y)$ is the difference of images, $I_n(x, y)$ is the current frame images, and $B_n(x, y)$ is the adaptive background reference images.

Most object extraction approaches calculate the difference between a reference frame and a current frame in grayscale images [21]. A major problem with this is that pixels are incorrectly classified, and essential data are lost. Consequently, the expected result is not obtained, effectively leading to a loss of information of the moving object. Therefore, we compute the frame differences in the RGB color space and achieve, to a good extent, the sought-after performance in maintaining the original image details. Eventually, the subtracted color image is converted to a grayscale image and a threshold operator is employed to extract the moving object regions from the image sequences, as described in the next step.

After completing the frame difference process, the next step applies a low-value thresholding operation. If the thresholds are lowered, regions of the object that have a similar chromaticity to that of the background will be misclassified as shadows, resulting in a high detection rate.

In the proposed method, the thresholding value is 30.

$$D_n(x, y) = \begin{cases} 255, (S_n(x, y))^2 > Th \\ 0, otherwise \end{cases} \quad (3)$$

where $D_n(x, y)$ is the new binary mask that is applied to segment out objects of interest and contains a value of "255" for each pixel marked as foreground. $S_n(x, y)$ is the difference of images, obtained in the previous section. Additionally, we calculate the square of S_n. Th is a "predefined" threshold. We obtained the threshold value based on several experiments. After checking 20 to 60 values, we obtained the expected results for our method when employing the 30 "threshold" value. However, in [22], the authors utilized the 40 "threshold" value to calculate the overlap regions of the current frame image and its background frame information.

Image registration, fast moving objects, and large illumination changes may impact the precise segmentation of moving object pixels. Prior to calculating the absolute difference of overlap regions, the average of color differences is computed at the optimal seam, and these values are distributed to the overlap regions to minimize the effect of illumination light source changes. Unwanted edges and noises may come into view on the difference of images when the registration mistake is more than a pixel.

4.3. Median Noise Removing

Sometimes, a background subtraction technique fails to yield palatable outcomes for indoor video segmentation. Thus, we adapted background subtraction by enhancing the strategy used, by utilizing median filtering to improve the extracted foreground mask, as in [23]. A median filter is a simple and powerful non-linear digital filter. It is used for reducing the amount of intensity variation between two pixels. Accordingly, we obtain the binary image (kernel) of every conceivably moving object in the overlap region, and the kernel data can be used with a noise removal filter.

Specifically, the median filter replaces a pixel by the median, rather than the normal, for all pixels in a neighborhood w, using Equation (4):

$$y[m,n] = median\{x[i,j], (i,j) \in w\} \tag{4}$$

where w represents a neighborhood characterized by the client that is centered on a location $[m,n]$ in the image.

Moreover, regarding the noise-reducing execution of the median filter, because the median filter is a nonlinear impulse filter, its numerical investigation is moderately complex for an image with irregular noise. For a picture with zero mean noise under an ordinary distribution, the noise difference of the filtering is approximately near the value from the below formula:

$$\sigma_{med}^2 = \frac{1}{4nf^2(\bar{n})} \approx \frac{\sigma_i^2}{n + \frac{\pi}{2} - 1} \cdot \frac{\pi}{2} \tag{5}$$

where σ_i^2 is the input noise power (the variance), n is the size of the median filtering mask, and nf is the function of the noise density.

4.4. Shadow Detection and Removal

The next step of our algorithm is the detection and removal of shadow regions from the foreground mask based on geometry feature information. The fundamental benefit of geometry features is that they deal specifically with the current input image. For example, they do not depend on a specific estimation of a background reference image. In our previous academic work, we applied a gradient-based texture correlation method in this step to remove shadow pixels and obtain improved results as compared to other state-of-the-art methods [23,24].

In general, in geometry-based applications, special attention is paid to the orientation, size, illumination source, and shape of shadows. For these reasons, a geometry-based method is used to distinguish shadow pixels from moving foreground objects [25]. The main hypothesis here is based on the fact that geometry properties impose scene limitations, for example, on particular target types, commonly walkers or animals (e.g., in an inside environment). For example, geometry properties might require objects and shadows to have an extraordinary orientation and one might expect a unique light source or background surface similarity.

Initially, we investigate the vertical peaks on each blob to recognize possible heads and subsequently utilize these data to split the blobs into set of individual-shadow sets. Given an individual-shadow area, R, its focal point of gravity, $(\overline{x}, \overline{y})$, and orientation, θ, these are found as follows:

$$(\overline{x}, \overline{y}) = \left(\frac{1}{|R|} \sum_{(x,y) \in R} x, \frac{1}{|R|} \sum_{(x,y) \in R} y \right) \tag{6}$$

$$\theta = \frac{1}{2} \arctan\left(\frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}} \right) \tag{7}$$

where $|R|$ is the area of the region in pixels and $\mu_{p,q}$ are the corresponding central moments. The point below the center of gravity with the maximum vertical change is considered to be the pixel where the shadow begins, and a segment oriented according to θ is used to roughly split a candidate shadow region R_2. Then, the following Gaussian model is built from the pixels in R_2:

$$G(s,t,g) = \exp\left[-\left(\frac{w_s s^2}{\sigma_s^2} + \frac{w_t t^2}{\sigma_t^2} + \frac{w_g\left(g-\mu_g\right)^2}{\sigma_g^2}\right)\right] \quad (8)$$

where s and t are the curved coordinates of each pixel, g is its gray level ($g = I(s,t)$), and w and σ^2 are the weight and difference of every segment in the Gaussian model, respectively. This model outlines the brightness of the shadow regions and contains the coordinates where the shadow is estimated to be. When the model is assembled, every pixel in the original region R is classified as an object or shadow, as indicated by whether it concurs with the Gaussian model or not. We additionally enhanced the shadow removal task using a combination of local and global contrast image enhancement and median noise filtering applied to the foreground binary mask, where noise can be removed as described in the above chapters.

4.5. Morphological Restoration

The final step of our suggested method is the process of filling small gaps or holes that arise from the distorted foreground mask (image) caused by shadow removal. It is generally accepted that shadow removal is a destructive process. This happens because of the similarity of the pixel intensities in the foreground object and corresponding background. Therefore, it has some negative effects, such as removing key parts of an object's shape after considering them as shadows. As a result, the original shapes of moving objects are distorted. To handle this issue, we employed a morphological operator to the foreground binary mask for restoration after the distorting shadow removal process.

Using appropriate thresholds is necessary for the foregoing geometry-based removal process to assure that all shadow pixels are removed [26]. As a consequence, only the regions not affected by noise and those clearly free of shadows are subject to the shape reconstruction process performed using Equation (9), which is represented as follows:

$$R = M_s \cap (\widetilde{M} \oplus SE) \quad (9)$$

where M_s is the mask image, $\widetilde{M}$ is the restored image, and SE is a structure element, whose size usually depends on the size of the objects of interest, although a 3×3 square element proved to work well in all of our tests. Furthermore, we use a manual image superimposing tool to superimpose the extracted object on the foreground image for the performance testing of our proposal.

5. Experiment Results and Analysis

In this section, we show our experimental results to assess the performance of our method in terms of its quality and speed. In addition, we provide comparisons between our strategy and the current well-known approaches. To eliminate wasting time in the image segmentation and shadow removal process, we have used techniques that are the quickest and are appropriate for low-powered equipment, such as local-global contrast enhancement, foreground segmentation with background subtraction, and a morphological restoration of the final shadow removed mask data. We implemented the proposed method in Visual Studio 2015 C++ and performed all experiments on a PC with a 3.60 GHz CPU and 4 GB of RAM. To evaluate the performance of the shadow removal method, the method has been tested in the widely used shadow detection CAVIAR and ATON datasets (hallway, lab, room and corridor) [27,28]. Moreover, we employ indoor video sequences at a size of 320×240, obtained from Gachon University and the Tashkent University of Information Technologies (TUIT). To check the robustness and effectiveness of the shadow remover, we compare our method with popular commercial

shadow detecting [23] software and state-of-the-art methods, such as the physical-based method (PBM) [29], gradient-based method (GBM) [24], and color-based method (CBM) [30]. Figures 3–6 show the results of visual comparison.

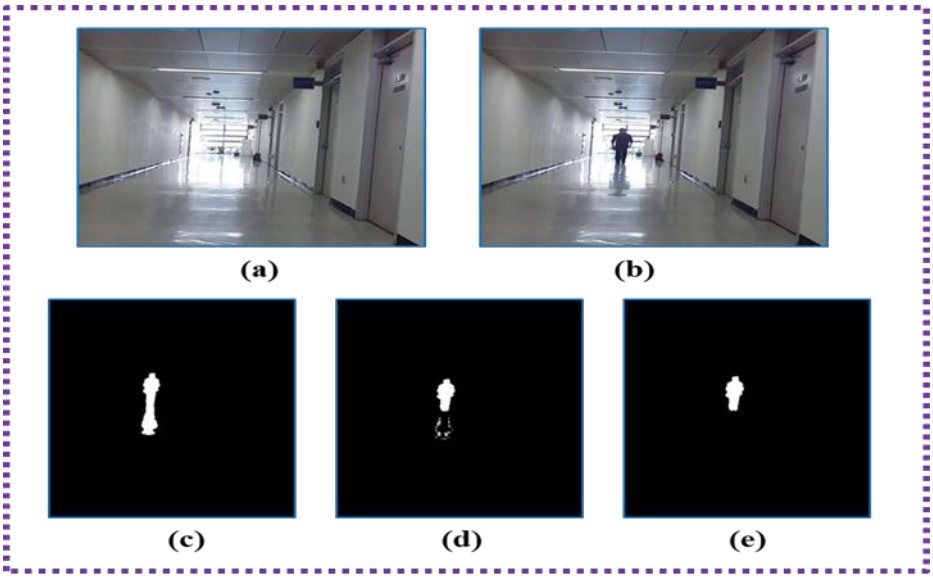

Figure 3. Shadow removing process: (**a**) Background image, (**b**) current image, (**c**) foreground mask (with shadow), (**d**) shadow detection and removal process, (**e**) applied morphological restoration and final foreground mask without shadow pixels.

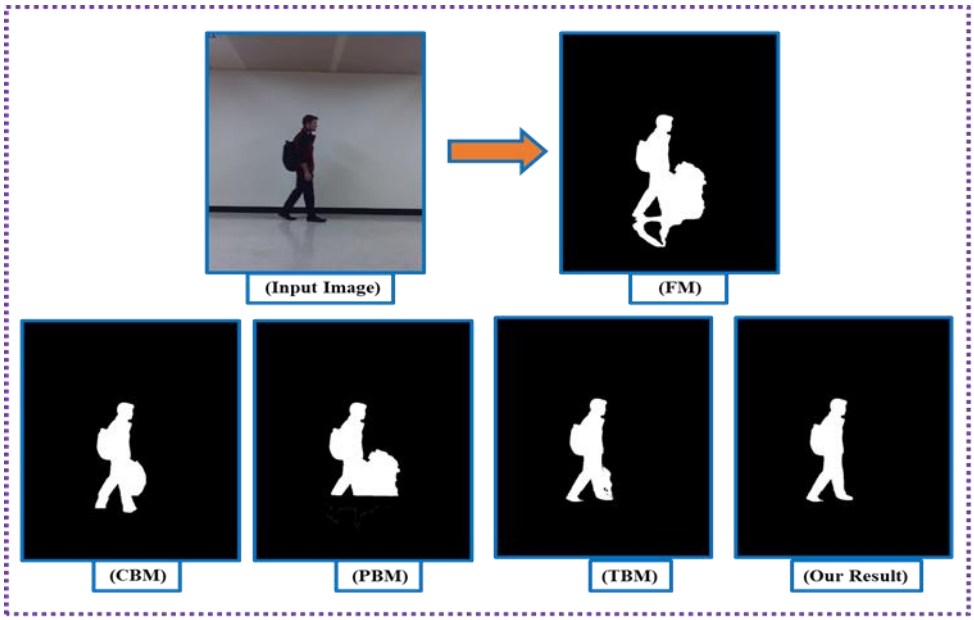

Figure 4. Comparisons between reliable existing approaches and our method's result: Input image (II); foreground mask for all shadow removal methods (FM); color-based method [30] (CBM); physical-based [29] method (PBM); texture-based [31] method (TBM); and our proposed method.

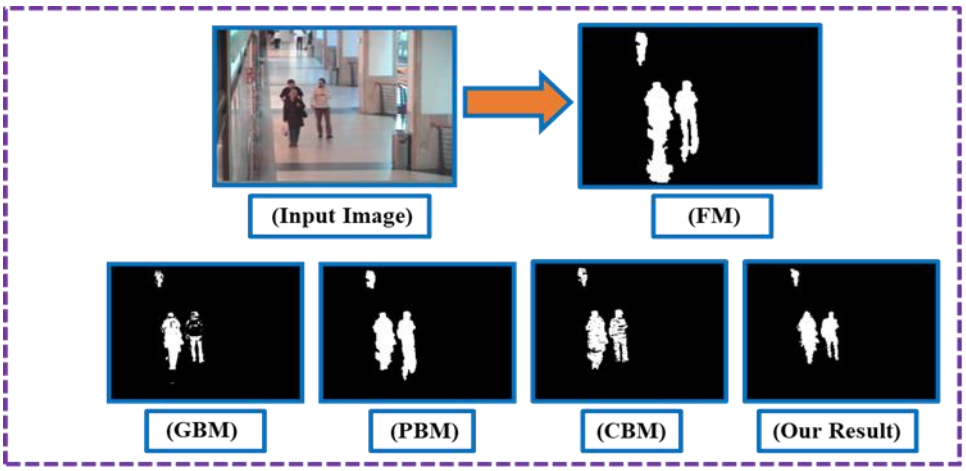

Figure 5. Visual comparison between previous approaches and proposed method: input image (II); foreground mask for all shadow removal methods (FM); gradient-based method (GBM) [24]; physical-based method (PBM) [29]; color-based method (CBM) [30]; and our proposed method.

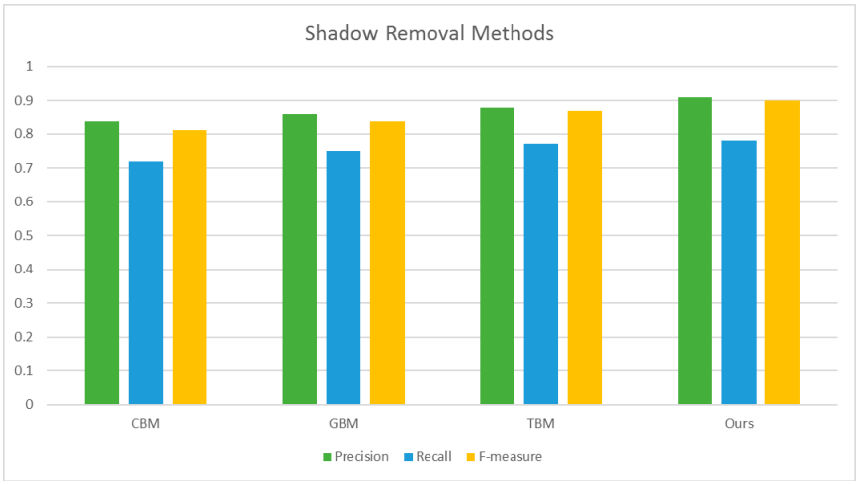

Figure 6. Quantitative comparisons of shadow removal approaches using the CAVIAR 10k dataset.

Figure 3 shows the overall process of our proposed method in images (the video streams were taken at Gachon University in Korea). In these steps, we utilized a number of algorithms to apply to real practice. First, we apply a local-global contrast enhancement algorithm to the input background reference image and the current image in which the moving object exists. Then, we segment the foreground moving object out of the enhanced image by calculating pixel differences. The foreground subtraction algorithm provides us with a moving object mask, but it contains both the object and shadow pixels. A foreground subtraction approach cannot distinguish between a moving object and its shadow region. Such an error is caused by the fact that a moving object and its moving shadow share similar motional characteristics as demonstrated in Figure 3c. According to the above section, shadow removal is destructive procedure, in that it generates unwanted gaps and holes in the foreground mask. By utilizing the morphological operator, the consequences of the shadow removal process are eliminated (see Figure 3c). In the final step, we detect the original shape and edges of the moving object. The suggested method can be employed for smart city object tracking and security system applications.

5.1. Qualitative Results

To prove the accuracy of our framework in eliminating ghost artifacts (shadows existing on both the wall and floor) and multiple shadows, more experiments are necessary. Figure 4 illustrates

comparisons between the proposed method and other state-of-the-art shadow removal algorithms, using similar foreground masks from the background subtraction section. The shadow is caused by the indoor light source and the results show that most of the existing shadow removal methods fail to generate a foreground mask without shadow pixels or significantly distort the object's shape. Our experimental results show that, in many cases, such as indoor scenes, a physical-based method faces problems when objects have spread-out shadow regions. In contrast, the scheme easily determines and removes monodirectional shadow pixels in moving objects. Moreover, ghost effects also make removal a difficult task. In addition, in the area where shadow pixels removed, salt and pepper noise occurs in PBMs. We addressed that problem by utilizing by the morphological approach in the last step of our method.

Texture-based methods are potentially powerful methods for detecting shadows, as textures are highly distinctive, do not depend on colors, and are robust to illumination changes [30,31]. A color-based approach fails when objects are darker and have colors similar to that of the background surface. A texture-based shadow detector better emphasizes the moving object, because this method also uses global contrast enhancement in the preprocessing step, as described in [23]. Furthermore, it effectively solves the problem occurring when multiple objects have multidirectional shadow regions. However, it has not yet been tested regarding ghost artifacts and background similarity issues. The results of our method appear to be attractive and accurate, as compared to the results of others. The moving shadow problems are addressed, and a seamless result is achieved using our method.

In the above experiments, we have mainly tested the performance of shadow remover frameworks on a single object. In Figure 5, we implement all of the selected methods to examine how they work on multiple objects. It is known that removing shadow pixels from multiple objects is a more difficult task than removing them from one object. For example, ghost effects cause a doubling of walking people on the foreground mask and it can appear unclear, or several objects can appear as a single moving blob on the foreground mask (kernel). From the results, we could see that our method works effectively, even when there are many objects in an image sequence, as shown in Figure 5. Our improved shadow removal method delineates foreground objects more accurately in indoor (e.g., corridor) video surveillance. In addition, in many cases, both our method and a gradient-based method (GBM) perform well on handling multiple dynamic object problems. In contrast, physical-based method (PBM) and color-based method (CBM) strategies fail most of the time and misclassify the moving shadow region as an object region, or as a part of them.

5.2. Quantitative Results

The quantitative results are represented in this section. Quantitative analysis shows the average precision and recall rates, along with F-measures. Precision and recall rates can be obtained by using Equations (10) and (11). The results show that the proposed shadow remover accomplished the highest precision rate, at 0.92. The CBM [30], GBM [24], and texture-based method (TBM) [31] yielded precision rates of 0.85, 0.87 and 0.89, respectively. In addition, we computed the F-measure value which balanced measurements between the means of the precision and recall rates. The results of the quantitative comparison of the three algorithms are shown in Figure 6 and Table 1.

Table 1. Quantitative analysis of four shadow remover methods performed using the CAVIAR 10k dataset.

Methods	Color-Based Method (CBM)	Gradient-Based Method (GBM)	Texture-Based Method (TBM)	Ours
Precision	0.85	0.87	0.89	0.92
Recall	0.73	0.76	0.78	0.79
F-Measure	0.82	0.85	0.88	0.90

Precision and recall rates can be obtained as follows:

$$precision = \frac{TDO \cap GT}{ADO} \tag{10}$$

$$recall = \frac{TDO \cap GT}{GT} \tag{11}$$

where TDO denotes truly detected shadow regions, ADO denotes all detected objects, and GT denotes a manually-labelled ground trust value. In addition, we computed the F-measure value which balanced measurements between the mean of precision and recall rates. A higher F-measure means a higher performance, and it is defined as follows:

$$F_\beta = \frac{(1+\beta^2)precision \times recall}{\beta^2 \times precision + recall} \tag{12}$$

In that regard, β^2 = 0.3 was suggested by many shadow detection approaches, to raise the importance of the precision value. The point for weighting precision more than recall is that recall can be easily obtained by setting the entire region to the foreground. We used a fixed threshold that changed from 0 to 255. On each threshold, pair of precision and recall scores were computed and were lastly combined to form a precision-recall (PR) curve. The resulting PR curve can be scored by its maximal F-measure, which is a good summary of the shadow region detection performance when using a fixed threshold.

In addition to the PR and F-measure, we can also describe the false positive rate (FPR) and true positive rate (TPR) when binarizing the foreground mask with a series of fixed thresholds. The receiver operating characteristics (ROC) curve is a plot of TPR against FPR and is obtained by sliding the threshold. TPR and FPR can be defined as follows:

$$TPR = \frac{|BM \cap GT|}{|GT|} \tag{13}$$

$$FPR = \frac{\left|\overline{BM} \cap \overline{GT}\right|}{\left|\overline{GT}\right|} \tag{14}$$

where $\overline{BM}$ and $\overline{GT}$ denote the opposites of the binary mask, BM, and ground-truth, GT, respectively. Figure 7 shows the PR and ROC curves of the proposed method, along with those of alternative shadow removal algorithms.

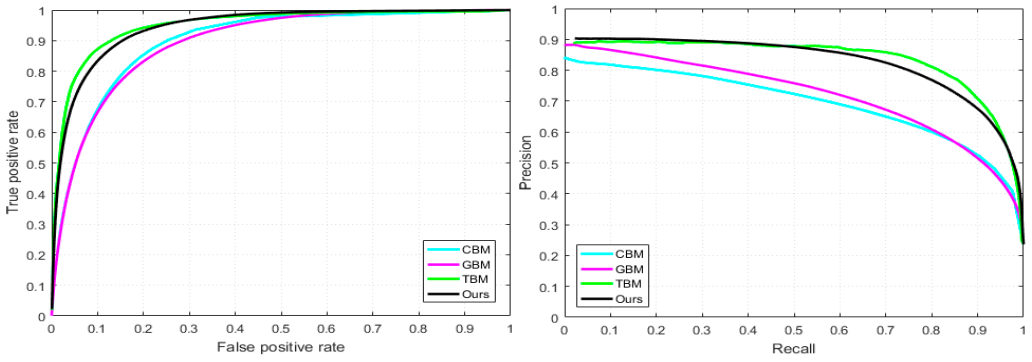

Figure 7. Quantitative comparisons using precision-recall (PR), (left column) and operating characteristic (ROC), (right column) curves of the proposed method, and those produced by various approaches on the CAVIAR 10k dataset.

Moreover, we contrasted our strategy and other shadow expulsion frameworks (physical-based and gradient-based), to determine the speed of the execution of the method. In most cases, our methodology

and the GB method indicate relatively similar times for handling of results. In some cases, our shadow detector takes somewhat more time with different calculations, as shown in Table 2. The reason for this is that our method requires additional stages to improve the input current frame and compute the geometry property points for every pixel.

Table 2. Average frame processing time (in seconds) per sequence for various shadow removal methods.

	Processing Time (In Seconds)		
	Gradient Based Method	**Physical Based Method**	**Our Result**
Gachon University	0.30	0.36	0.32
TUIT University	0.25	0.29	0.26
Corridor (dataset)	0.20	0.27	0.22
Average	0.25	0.30	0.27

In that regard, we successfully solved moving object and background surface similarity problems in inside environments. However, when our method was implemented in an outdoor environment, it sometimes faced difficulties in detecting and removing shadowed regions. For this reason, original object shapes might be distorted, or foreground regions might be removed as shadow pixels, as demonstrated in Figure 8b. In contrast, if the moving object and background surface have pixel intensities with large differences, the suggested method could remove shadow pixels without distortion, as demonstrated in Figure 8a. In our future research, we are going to address that problem by applying neural network and deep learning approaches. Clearly, the proposed method still works successfully when tracking moving objects in all types of indoor environments (e.g., conference halls, offices, or indoor sports game locations).

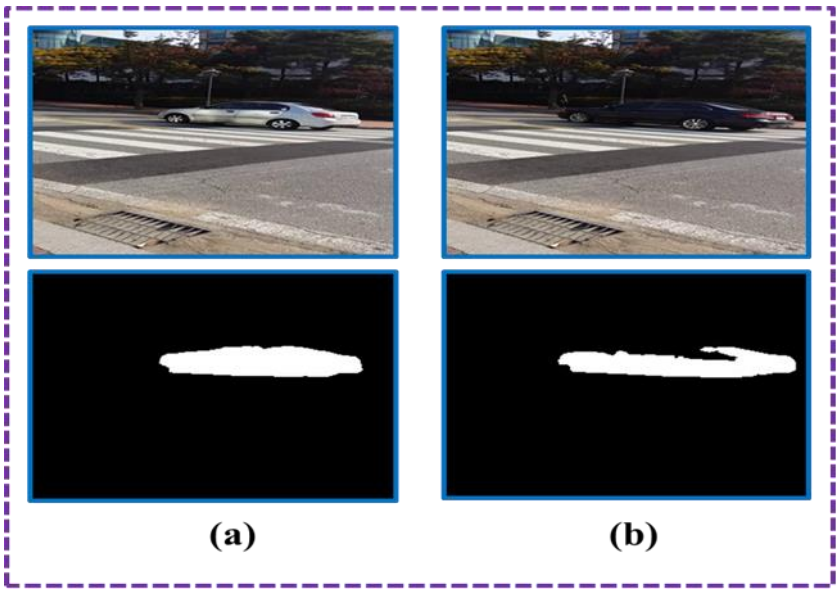

Figure 8. Experimental results of shadow removal: (**a**) input image and obtained foreground mask by our method (without distortion), (**b**) input image and obtained foreground mask by our method (with distortion).

6. Conclusions and Future Work

This paper proposed a new method for detecting and extracting shadow pixels from moving objects. We used a robust image enhancement technique to increase the contrast of image sequences after the capturing process and a background subtraction strategy to calculate the difference between two input images. Then, shadow regions were removed by applying a unique and fast geometry-based

method. Following that, a median filtering approach was utilized to eliminate salt and pepper noise in the foreground mask. Finally, gaps and holes were filled using morphological reconstruction to generate a clear moving target mask. From the experiments, it can be seen that our method addresses the ghosting artifacts caused by multiple objects and the background surface similarity caused by the current image and reference image having the same intensity values.

Potential future work includes improving the quality of our method by employing a convolution neural network (CNN) to handle strong shadows in outdoor environments. We are going to implement our strategy for smart and safety cities, to easily realize the natural shape of moving objects and track them.

Author Contributions: This manuscript was designed, written and performed the experiments by A.A. Theory and experiments analyzed and commented by T.K.W.

Funding: This work was supported by the Gachon University research fund of 2018 (GCU-2018-0681).

Acknowledgments: I would like to express my sincere gratitude and appreciation to my supervisor, Taeg Keun Whangbo under the Department of Computer Science (Gachon University) for his support, comments, remarks and engagement over the period in which this manuscript was written.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Khan, S.H.; Bennamoun, M. Automatic shadow detection and removal from a single image. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *6*, 431–446. [CrossRef]
2. Cucchiara, R.; Grana Piccardi, C.M.; Prati, A. Detecting moving objects, ghosts, and shadows in video streams. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *25*, 1337–1342. [CrossRef]
3. Sanin, C.; Sanderson, C.; Lovell, B.C. Shadow detection: A survey and comparative evaluation of recent methods. *Pattern Recognit.* **2012**, *45*, 1684–1695. [CrossRef]
4. Leone, A.; Distante, C.; Buccolieri, F. A texture-based approach for shadow detection. In Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, Como, Italy, 15–16 September 2005; pp. 371–376.
5. Wan, Y.; Miao, Z. Automatic panorama image mosaic and ghost eliminating. In Proceedings of the International Conference on Multimedia and Expo, Hannover, Germany, 26 April–23 June 2008; pp. 945–948.
6. Cucchiara, R.; Grana, C.; Piccardi, M.; Prati, A.; Sirotti, S. Improving shadow suppression in moving object detection with hsv color information. In Proceedings of the IEEE Intelligent Transportation Systems, Oakland, CA, USA, 25–29 August 2001; pp. 334–339.
7. Abdusalomov, A.; Whangbo, T.K.; Djuraev, O. A Review on various widely used shadow detection methods to identify a shadow from images. *Int. J. Sci. Res. Publ.* **2016**, *6*, 2250–3153.
8. Horprasert, T.; Harwood, D.; Davis, L.S. A statistical approach for real-time robust background subtraction and shadow detection. *In ICCV Frame-Rate WS. IEEE* **1999**, *99*, 1–19.
9. Kim, K.; Chalidabhongse, T.H.; Harwood, D.; Davis, L.S. Real-time foregroundbackground segmentation using codebook model. *Real Time Imaging* **2005**, *11*, 172–185. [CrossRef]
10. Yuan, C.; Yang, C.; Xu, Z. Simple vehicle detection with shadow removal at intersection. In Proceedings of the 2nd International Conference on Multimedia and Information Technology, Hong Kong, China, 28–30 December 2010; pp. 188–191.
11. Stauder, J.; Mech, R.; Ostermann, J. Detection of moving cast shadows for object segmentation. *IEEE Trans. Multimed.* **1999**, *1*, 65–76. [CrossRef]
12. Amato, A.; Huerta, I.; Mozerov, M.; Gonsalez, J. Moving Cast Shadows Detection Methods for Video Surveillance Applications. In *Wide Area Surveillance*; Vijayan Asari, K., Ed.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 23–47. [CrossRef]
13. Katharavayan, R.S.; Nagarathinam, K. A Survey of Moving Cast Shadow Detection Methods. *Int. J. Sci. Eng. Res.* **2014**, *5*, 752–764.
14. Kim, D.; Arsalan, M.; Park, K. Convolutional Neural Network-Based Shadow Detection in Images Using Visible Light Camera Sensor. *Sensors* **2018**, *18*, 960. [CrossRef] [PubMed]

15. Lee, J.T.; Kang, H.; Lim, K.T. Moving Shadow Detection using Deep Learning and Markov Random Field. *J. Korea Multimed. Soc.* **2015**, *18*, 1432–1438. [CrossRef]
16. Lo, B.P.L.; Yang, G.-Z. *Neuro-Fuzzy Shadow Filter*; Imperial College of Science, Technology and Medicine: London, UK, 2014.
17. Amato, I.; Huerta, M.; Mozerov, M.G.; Roca, X.; Gonzalez, J. *Moving Cast Shadow Detection Method for Video Surveillance Applications*; Springer: Berlin/Heidelberg, Germany, 2012.
18. Usmanov, R.; Abdusalomov, A.; Kuchkorov, T.; Mukhiddinov, M. Image enhancement based on histogram equalization for indoor environment objects. In Proceedings of the International Scientific-Practical and Spiritual-Educational Conference Dedicated to the 1235th Anniversary of Muhammad al-Khwarizmi, Tashkent, Uzbekistan, 5–6 April 2018.
19. Singh, A.; Kumar, N. A Global-Local Contrast based Image Enhancement Technique based on Local Standard Deviation. *Int. J. Comput. Appl.* **2014**, *93*, 975–8887. [CrossRef]
20. Bradski, G.; Kaehler, A. *Learning OpenCV: Computer Vision with the OpenCV Library*; O'Reilly Media: Newton, MA, USA, 2008.
21. Jacques, J.C.S.; Jung, C.R.; Musse, S.R. Background subtraction and shadow detection in grayscale video sequences. In Proceedings of the 18th Brazilian Symposium on Computer Graphics and Image Processing, SIBGRAPI '05, Washington, DC, USA, 9–12 October 2005.
22. Abdukholikov, M.; Whangbo, T. Fast image stitching method for handling dynamic object problems in Panoramic Images. *Ksii Trans. Internet Inf. Syst.* **2017**, *11*. [CrossRef]
23. Abdusalomov, A.; Whangbo, T.K. An Improvement for the Foreground Recognition Method using Shadow Removal Technique for Indoor Environments. *Int. J. Wavelets Multiresolution Inf. Process.* **2017**, *15*, 1750039. [CrossRef]
24. Sanin, A.; Sanderson, C.; Lovell, B. Improved shadow removal for robust person tracking in surveillance scenarios. In Proceedings of the 10th International Conference on Pattern Recognition Systems, Tours, France, 8–10 July 2019; pp. 141–144.
25. Hsieh, J.-W.; Hu, W.-F.; Chang, C.-J.; Chen, Y.-S. Shadow elimination for effective moving object detection by Gaussian shadow modeling. *Image Vis. Comput.* **2003**, *21*, 505–516. [CrossRef]
26. Wang, C.; Zhang, W. A Robust Algorithm for Shadow Removal of Foreground Detection in Video Surveillance. In Proceedings of the Asia-Pacific Conference on Information Processing, Shenzhen, China, 18–19 July 2009.
27. CAVIAR Test Case Scenarios. Available online: http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/ (accessed on 15 May 2019).
28. Shadow Detection. Available online: http://cvrr.ucsd.edu/aton/shadow/ (accessed on 15 May 2019).
29. Huang, J.-B.; Chen, C.-S. Moving cast shadow detection using physics-based features. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 2310–2317.
30. Shan, Y.; Yang, F.; Wang, R. Color space selection for moving shadow elimination. In Proceedings of the 4th International Conference on Image and Graphics, Sichuan, China, 22–24 August 2007; pp. 496–501.
31. Leone, A.; Distante, C. Shadow detection for moving objects based on texture analysis. *Pattern Recognit.* **2007**, *40*, 1222–1233. [CrossRef]