BRUNEL UNIVERSITY

Department Of Computer Science

DOCTORAL THESIS

Research Into Illumination Variance In Video Processing

Author:

Seyed Mahdi H S

JAVADI

Supervisors:

Dr. Yongmin LI and

Professor XiaoHui LIU

Year of Submission: 2018

Declaration of Authorship

I, Seyed Mahdi H S JAVADI, declare that this thesis titled, "Research Into Illumination Variance In Video Processing" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

BRUNEL UNIVERSITY

Abstract

CEDPS

Department Of Computer Science

Doctor of Philosophy

Research Into Illumination Variance In Video Processing

by Seyed Mahdi H S JAVADI

In this thesis we focus on the impact of illumination changes in video and we discuss how we can minimize the impact of illumination variance in video processing systems.

Identifying and removing shadows automatically is a very well established and an important topic in image and video processing. Having shadowless image data would benefit many other systems such as video surveillance, tracking and object recognition algorithms.

A novel approach to automatically detect and remove shadows is presented in this paper. This new method is based on the observation that, owing to the relative movement of the sun, the length and position of a shadow changes linearly over a relatively long period of time in outdoor environments, we can conveniently distinguish a shadow from other dark regions in an input video. Then we can identify the Reference Shadow as the one with the highest confidence of the mentioned linear changes. Once one shadow is detected, the rest of the shadow can also be identified and removed. We have provided many experiments and our method is fully capable of detecting and removing the shadows of stationary and moving objects.

Additionally we have explained how reference shadows can be used to detect textures that reflect the light and shiny materials such as metal, glass and water. ...

Author Publications

Some of the research leading to this thesis has appeared previously in the following publications.

Journal Articles

- S. M. Javadi, Y. Li, X. Liu: **Removing Shadows From Video**. – *International Journal of Machine Learning and Computing*, November 2017

Conference Papers

- S. M. Javadi, Y. Li, X. Liu: **Shadow Detection and Removal**. – *6th International Conference on Image Processing*, University of Northumbria, UK, 2017

- S. M. Javadi, Y. Li, X. Liu: **Detecting non-Lambertian material in the video**. – *12th International Conference on Computer Vision, Image and computer Graphics and theory and application* , Porto, Portugal, 2017

Acknowledgements

I would like to express my sincere gratitude to my advisor Dr Yongmin Li for the continuous support of my Ph.D study for the last five years, for his patience, motivation, and immense knowledge of this topic. His guidance helped me in all the time of research and writing of the papers I published.

Additionally, I would like to thank Professor XiaoHui Liu for his insightful comments and encouragement.

Finally, I would like to thank my wife Zahra who was very patient and supportive, my son Amin who was born in the first year of my PhD. My father who supported me spiritually throughout this research and the rest of my family; my mother, brother and sister for their great support and encouragement. ...

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction to image and video processing

Image and Video Processing are important topics in the field of research and development in computer science. Image processing is any form of signal processing for which the input is an image, such as photographs or frames of video; the output of image processing can be either an image or a set of characteristics or parameters related to the input image. Most image processing systems. Video processing is a particular case of computer vision, where the input signals are video files or video streams from a camera.

Video processing has many applications such as surveillance, traffic monitoring, tracking, autonomous cars, activity recognition and many more. In this thesis we analyse the impact of illumination variance in video processing and explain how lighting conditions can be improved.

1.2 Introduction to illumination variance

This thesis discusses the impact of illumination variance in the video processing applications. In particular, it concentrates on outdoor scenarios where the scene is lit prominently by the sun.

These conditions are some of the most common and popular cases in video and image processing systems and there are many applications for both academic and industrial research projects such as video surveillance and tracking. The topic discussed here can be used in many applications, such as traffic monitoring, CCTV surveillance systems, crowd monitoring, activity recognition systems and much more. Our point of interest is to discover how the scenes lighting conditions can be enhanced and in particular we focus on detecting shadows and removing them automatically without user input. Typically there are two reasons for changes to illumination where the natural light is dominant:

- Gradual changes to illumination from sunrise to sunset, which is due to the angle of the sun.
- Rapid changes in illumination due to an object blocking the sun and creating shadows.

The first type of change is gradual and normally has no impact on image and video processing applications since the variance is small during shorter periods. Conversely the second type of illumination change can cause issues. This is because shadows create unwanted edges, which create problems in downstream applications such as object recognition

and tracking. The reason is that most of the applications in image and video processing are using edge detection in one form or another, and they expect to have edges when there is a material change, and this can be used for detecting and distinguishing objects. However, when there are shadows in the scene, their edges can give a false indication of material change and reduce the performance of tracking and object recognition algorithms.

We have elaborated on these topics and explained how the outdoor shadows could be effectively detected and removed. Moreover, we have provided novel solutions for this established problem in image and video processing science. Shadow detection is the biggest part of this research but not all of it. Additionally it has been discussed how to detect the presence of cloud in the scene and explained how some other materials can be detected such as glass, water and metal or any other material that does not fit into the Lambertian reflectance model i.e. any glossy and shiny materials.

1.3 Types of shadows

There are two types shadows in video and image processing:

- Cast shadow.
- Self shadows.

Cast shadows are generated when an object blocks the main source of light and casts a shadow over the background. However self shadow is produced when one side of the object is not facing towards the light source; hence, its illumination differs from another part of the object.

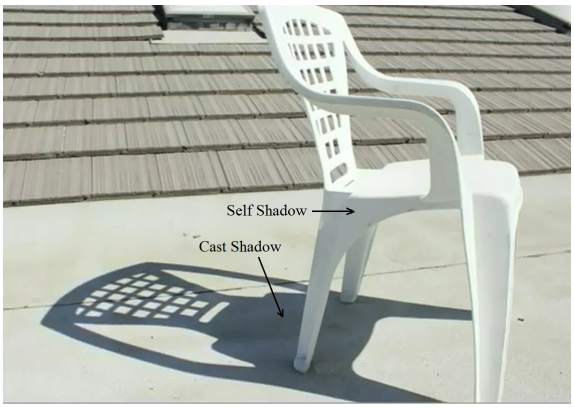


FIGURE 1.1: Self and Cast Shadows

Additionally, cast shadows can be categorised into the following forms:

- Umbra.
- Penumbra.

Umbra is the darkest part of the shadow where the source light is completely blocked. Penumbra is the region where only part of the source light is blocked. The effect of these two types of shadows depends on the distance between the source of light, an object blocking light and position of the camera. In normal outdoor conditions where is scene illuminated by the sun and the camera is too focused (as seen in normal CCTV footage), the effect of penumbra can be discarded. This research has focused on detecting and removing umbra. Penumbra shadow becomes

important and substantial where the source of light is very close to the blocking object (such as artificial lab environment) or where the camera sensor is too close or focused on the shadow region. The following figures illustrate the difference between these types of shadows.

Umbra cast shadows have the greatest impact on the verity of image and video processing applications due to their strong edges. These shadow edges may cause segmentation problems for object and shape recognition methods as well as tracking algorithms. This research project focuses on these shadows and provides a novel solution to detect and remove cast shadows from stationary and moving objects automatically.

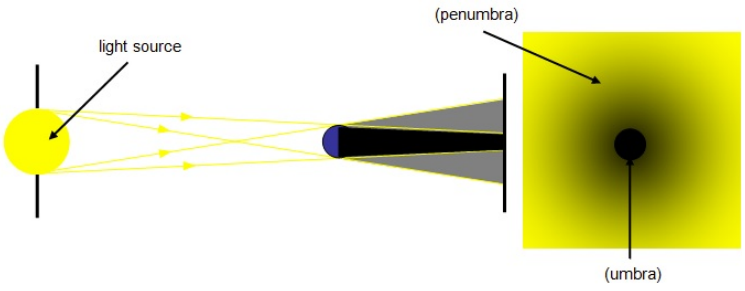


FIGURE 1.2: Umbra and Penumbra shadows

1.4 Introduction to reference shadows

Shadow clocks or sundials have been used by humans for thousands of years to identify the time of the day based on the length of the shadows. A sundial is a device that tells the time by the position of the sun in the sky. As the sun appears to move across the sky, the shadow moves and aligns with different hour lines. The oldest sundial was found in Egypt

dating from 1500 BC. We have used the same concept to distinguish shadows from dark objects in a method we call the reference shadows. Once some of the shadows are detected, the reminder will be identified and removed as, explained in other chapters.

This is the first time an ancient and simple method has been utilised to produce a robust and effective shadow detection method in video processing. Many experiments and samples from real life scenarios and complex video footage from busy scenes have been used to test and verify the performance of the algorithms mentioned in this document.

FIGURE 1.3: The change in length of shadow have been used for thousands of years to tell the time of day.

1.5 Why this topic was chosen

This PhD research project was a true journey for the author. The project started as video surveillance and activity recognition research topic but very quickly the impact of cast shadows in segmentation and shape identification was identified.

The decision was made to remove shadows from objects and then use shadowless output for tracking. However the impact and complexity of shadow detection became clear; hence, the author decided to concentrate on this remarkable topic.

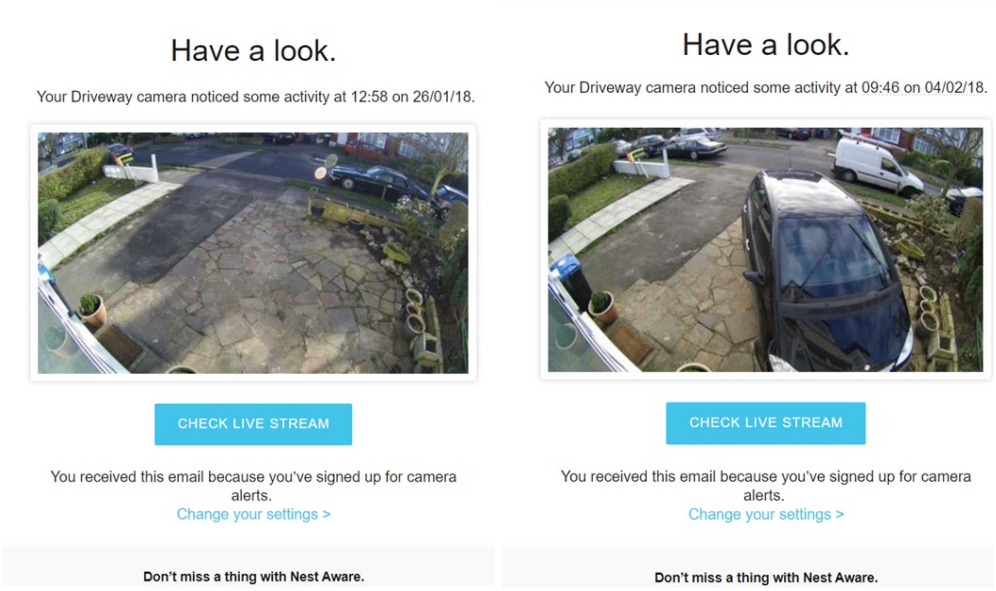
1.6 How this research can be used in video processing systems

In addition to shadow detection and removal, we have demonstrated how the illumination invariances can be used to detect some events and features in the video. It is explained how shiny objects which reflect the sun can be identified in video automatically and we have also highlighted the potential for detecting the presence of cloud in the scene by analysing all the reference shadows. When clouds block the sun, the shadows are removed, and a natural shadow-less image is produced which can be used as a reference to produce high quality coloured shadow-less images in other frames of the video. This process is explained in Chapter 5.

1.7 Real-life Example

Smart home cameras are getting very popular and many users around the world purchase these products to provide security for their home or work place. One of the market leaders is a product called Nest and according

to industry standards, it has the best activity recognition feature. Having said that, this product constantly mistaken shadows with foreground objects and also incorrectly detects glare and reflectance as an abnormal activity which is illustrated in Figure 1.4



(a) (b)

FIGURE 1.4: (a) Shadow is detected as a person and in correct notification is sent. (b) Reflectance on the car is detected as activity and incorrect notification is sent.

The research in this thesis can be used to address issues like this and in this research we have provided novel methods to detect and remove shadows as well as reducing impact of glare and reflectance.

1.8 Experiment data

Most of the data used for experiments has been collected by the author using normal video recording devices (smart phone and Panasonic camcorder V270). Video footage recorded from inside and around Brunel University has been used as test data; additionally, other publicly available

images and video have been also used. In some sections of the document, data used by other authors have been utilised to compare the results of our methods with existing publications. Where is this the case, it has been clearly explained and referenced.

1.9 Main contributions

In this thesis we focused on the special case of illumination variance in outdoor scenarios where the camera is static and the sun is the dominant source of light. The main contributions of this thesis are:

1. A new method to detect shadows automatically is introduced.
2. How shadows can be removed automatically.
3. A novel solution to detect shiny materials in the video.
4. How the presence of clouds can be detected automatically.

In the following chapters these novel contributions are presented and explained in details.

1.10 Dissertation content

Chapter 2 is the literature review where we analyse the history of shadow and cloud detection from early publications to state-of-art and summarise the weakness and strength of each method. Chapter 3 explains our shadow

detection method in details, and Chapter 4 elaborates the shadow removal methods analysed in this research. Chapter 5 outlines the possibility of detecting some features in video such as shiny materials and also discusses identifying events such as the presence of cloud by evaluating the illumination characteristics of the scene. Finally Chapter 6 summarises the topics discussed in this document and explains how researchers can continue this work in future.

Chapter 2

Literature review

Automatic shadow detection and producing invariant illumination images is a deep-rooted topic and probably it is as old as computer vision science if not older. For example, in the late 1960s, scientists such as Barrow and Tenenbaum began working on this topic. Since then, it has been a very popular subject for computer vision research as well as industry, due to its impact on other systems, such as object tracking, shape recognition and so on.

Overall and in high level, there are a few different approaches to shadow detection:

- HSV, HSI and YUV based models for shadow detection.
- Methods that produce illumination invariant images.
- Semi-automatic and neural network based methods.
- Texture and characteristic estimation methods.

Most of the major work in this area uses one or some of the above approaches to detect shadows, and these are explained in the next few sections. Mosleh and Sharfraz [19] have also summarised these approaches in their research

Additionally there are many research studies where shadows are detected in special conditions. An example is the work by Yanfeng [22] which only detects shadows from moving objects. Wang and Deng [3] research is focused only shadows in road. Yank and Sui is also a similar study and their proposed method removes the shadows from vehicles [24]. Similarity Ji and Zhao [23] also focus only on removing shadows from moving objects and Russell focuses on human shadow detection [16]. This is a very popular approach in shadow removal from objects and there are many other studies similar to this but they fail to detect shadows from static objects [4] [15] and [13]. All of these methods only detect shadows from foreground objects but the novel solution presented in this thesis can detect shadows from both foreground and background in out door.

In addition to shadow detection we have provided solutions for the following two problems in this thesis:

1. Detecting source of light reflection in the video.

2. Detecting the presence of cloud in the video.

These two topics are not reviewed and researched as much as shadow detection which is a very well-established research subject in image and video processing. However, overall, the existing approaches can be categorised as [78]:

- Analysing the position of the sun.
- Outdoor illumination modelling.

2.1 Outdoor illumination modelling and analysing the position of the sun

Outdoor natural illumination model is vital to understand and it has been researched and analysed extensively within the image and video processing community. For example Narasimhan et al [92] used a high quality set of photos taken on an hourly basis from a stationary camera over one year and used this information to analyse the impact of illumination and weather on video frames. This method has been used by several other researchers such as [76], [66] and [109] which produced high quality shadow-less image. The main limitation of this solution is if the position or direction of the camera changes, it can no longer detect shadows.

In [78] Lalonde et. al have explained how the camera parameters can be obtained automatically by analysing the source of illumination from the sun. The position of the sun has been researched by Cozman and Krotkov [44] and they used this to calculate the altitude and longitude of the camera. Furthermore, Terbi et al have designed a system in [114] that can estimate the orientation of the camera in addition to altitude and longitude.

The appearance of the sky has been analysed extensively in physics and one of the most prominent works was conducted by Perez et al [96] which is the result of measuring sky luminance. This model has been used frequently in relighting models such as the work from Yu and Malik [125]. As an alternative approach Stummpfel [108] mapped the sky into an HDR (High Dynamic Range) format and used that information to relight the scene. In [78] two independent sources of information are analysed: the sun's position and the sun's appearance and used these information to calculate the camera focal length. The shadow detection solution which is presented in this study does not require any details about camera characteristic.

2.2 HSV, HSI, YUV and colour consistency based methods for shadow detection

An important work in this area is presented by Victor Tsai [115] who analysed various models such as HSI, HSV. HCV, YIQ and YC_b C_r for shadow detection. He demonstrated that is (H_e+1)/(I_e+1) ratio can be used to distinguish shadow pixels from non-shadow regions. H_e and I_e are Intensity-equivalent and hue-equivalent and they are the H and I component of HSI model. The key finding of this research is pixels in shadowed regions will have higher values the (H_e+1)/(I_e+1) ratio than those pixels in nonshadowed regions. The Otsu's method [95] was then applied to find the correct threshold. Although this a very simple method and performs very well

in many cases, it fails when distinguishing between shadows and dark objects, which have a similar colour to shadows. This is one of the biggest problems with many shadow detection system, especially those based on simple HSI thresholding algorithms. The strength of method presented in this thesis is it can successfully differentiate between shadows and dark materials.

Another approach is using YUV model which has more similarities to human vision than RGB. In [59]a YUV based system is proposed for object detection from the shadow. The algorithm utilises all components in YUV colour space to identify shadow pixels from the candidate foreground regions. An adaptive threshold is designed to enhance shadow detection accuracy and adaptive capacity in various lighting conditions. This estimator uses edge detection method to obtain global texture, as well statistical calculations to obtain the required thresholds. An important advantage of this method to previously mentioned system is it works both indoor with artificial lighting and outdoor. On the other hand the weakness of this system is it only detects the shadow of moving objects.

Suny and Deb [45] proposed a solution based on the YCbCr model. In this system, an approach based on statistics of intensity in the YCbCr colour space is proposed for detecting shadows. When the shadow candidates are detected, a shadow density model is applied and the image is then segmented into several regions that have the same density. Finally, the shadows are removed by relighting each pixel in the YCbCr colour space and correcting the colour of the shadowed regions in the

RGB colour space. Similar to other studies in this category,[45] struggles to differentiate between shadows and black materials.

Methods based on the colour consistency of image regions and pixels are also very popular in this topic. They work by analysing the changes in the colour of edges and then try to guess if the boundary is generated by a change in the material or illumination. in [126] authors have used surface descriptor and colour-shade, which allows them to include the physical considerations derived from the image formation model capturing gradual colour surface variations. Their solution is designed to work in greyscale colour space and they utilised illumination nation pairs to detect shadows. In this category of research estimating the colour of prevailing scene illumination is a related problem which has received much attention [46],[56],[42] and [57]. In these works, calculating the colour constancy is a major problem and often they perform well only in restrictive conditions or assumptions. Detecting and removing chromatic shadows is also a particularly difficult task because they are extremely difficult to distinguish from moving objects.

A recent study by Mo and Zhu [17] proposes an object-oriented shadow detection method without manual intervention and a shadow compensation method by regional matching. In the proposed method, pixel-based soft shadow detection, which uses Gaussian mixture model to simulate the gray distribution and refines soft shadow map with guiled filtering, is combined with image segmentation result to obtain accurate shadow regions with complete shape and no hole. Then shadow regions are compensated, with less loss of details and brightness imbalance, referring to their optimal homogeneous nonshadow region obtained by regional

matching based on Bag-of-Words. Although this method provide good results for moving objects, it fails to detect shadows from background objects. Other works such as [25], [8] and [21]. Unlike any of these methods, the reference shadow system proposed in this thesis can detect shadows from both stationary and moving objects.

2.3 Illumination invariant based methods

There has been a significant development in shadow detection and removal. Introduced in one of the early work by Barrow and Tenenbaum [36], was the concept of "Intrinsic images". Intrinsic images are a mid level transform of the observed images. They are viewpoint dependent, and the physical causes of changes in illumination at different points are not made explicit. Barrow and Tenenbaum explained that this mid-level transform could be very useful for supporting a range of visual inferences.

Multiple frames have been used to compute the intrinsic images [119]. They approached the subject of shadow detection is by formulating this problem as a maximum-likelihood estimation problem based on the assumption that derivative-like filter outputs applied to illumination will tend to be sparse.They derived the ML estimator under this assumption and showed that it provides a suitable method for recovering reflectance. Furthermore, they assumed that filter outputs are independent of space and time. In this method, Weiss used 35 images from morning until evening and we have compared our results with this method. The advantage of our method is we were able to identify shadow in a similar

noisy environment in approximately 10 minutes.

Another important study on this topic was published by Finlayson et al [51], who introduced the concept of greyscale invariant image and presented a computational model to estimate the invariant image. The same authors published another study and created Invariant Image by Entropy Minimization [52]. In our research, we are inspired by the work of Finlayson et al but we used reference shadows instead of camera calibration to create invariant image automatically. Finlayson defined greyscale invariant image as:

$$gs = c_1 R'(R) - c_2 R'(B) \tag{2.1}$$

Where in RGB space, $R'(R) = log(R/G)$ and $R'(B) = Log(B/G)$. c_1 and c_2 are constants such that the vector [c_1 c_2] is in the direction is orthogonal to the lighting direction.

Experiments have confirmed images with a different level of illumination will map to the same greyscale invariant image. The most important benefit of this image is shadowless (which occur when there is a change in luminance) will disappear. Now if we divide (2) by c_1, we get the following:

$$gs' = R'(R) - CR'(B) \tag{2.2}$$

where $C = c_2/c_1$

Others such as Finlayson [51] have used manual camera calibration to identify c_1 and c_2 but we used the reference shadow to automatically create grey-scale invariant image by estimating the value of C.

A generic problem with illumination invariant based methods is many other important details in the image is also removed together with shadows and this may create some issues in segmentation.

[87] and [88] used multiple frames but it is not clear how these methods will perform when there are other changes in the scene; for example, multiple object movement. Moreover, this method can be useful in a controlled environment but the performance of this system in natural outdoor scenarios is unknown. The reference shadow solution which is presented in this thesis can successfully detect shadows in noisy environments with many moving objects

There are also other methods that detect shadows based on the previous detected shadows positions and the use of the sun [64], or about calibration methods using the sun, and the sky [79] or the position of shadows [31]. These methods require a significant amount of training data and also need additional information about the scene such as GPS location of the camera, date and time of the day. However none of this information is necessary in the method presented in this dissertation.

Huerta et al. in [64] developed a solution and first they built a chromatic invariant colour cone model and an invariant gradient model for segmentation of potential shadows. Then, regions corresponding to potential shadows were grouped together by studying "a bluish effect" and an edge partitioning. Finally (i) temporal similarities between textures

and (ii) spatial similarities between chrominance angle and brightness are reviewed for all potential shadow regions in order to finally differentiate between shadows and non-shadow regions.

2.3.1 Semi-automatic and neural network based methods

Semi-supervised learning based methods have been developed to detect shadows. These methods require at least one shadow to be labelled by a user and then the remaining shadows will be identified. A very good example is the work by El-Zahhar [47]. They proposed to initially to extract colour, texture and gradient that are useful for differentiating between moving objects and their shadows. They then used a semi-supervised learning approach for shadow detection. In [39] authors designed a Hierarchical Mixture of MLP expert methods that uses a two-stage system for shadow detection and their proposed system works in both indoor and outdoor conditions.

Another approach is presented in [30] whereby chromaticity information is first used to create a mask of candidate shadow pixels. This is followed by employing gradient information to remove foreground pixels that were included incorrectly in the mask.

The neural network is another popular method to detect shadows which usually requires a significant amount of labelled training images [107]. Faghih and Moghaddam [49] used image statistics to model the accuracy

of Grey-Edge assumption so they could compensate the Grey-Edge algorithm error. They used Weibull distribution to describe image statistics based on image derivatives.

In [40] authors describe a multilayer neural network that can recover the illumination chromaticity given only one image of the scene. The network is previously trained with a set of images of scenes and the chromaticities of the corresponding scene illuminants. Experiments with other images demonstrate that the network performs better than many of the existing constancy methods. In particular, the performance of this solution is better for images with a relatively small number of colours. The *rg* chromaticity space is used as an input into neural networks. In this model, $r = R/(R + G + B)$ and $g = G/(R + G + B)$.

Agarwal et al. have presented a linear technique called RR (Ridge Regression) in[27] and compared their results with some of the existing Neural Networks and SVR (Support Vector Regression) systems. To model the chromaticity, *rg* model is used similar to [40].

Cavallaro, Salvador and Ebrahimi [41]have attempted to detect shadows from a sequence of images by processing three sources of information, namely colour, spatial, and temporal information of the scene. Foreground objects were first segmented from the background, and then shadows from the foreground objects were successfully detected. The performance of the above mentioned methods varies when dealing with changing lighting conditions. Usually these methods require additional manual

parameter adjustments so it can be adapted to different environments. To solve this issue mixture Gaussian is used in various studies to model shadows dynamically but good results are only achieved when the scene meets a series of assumptions[67].

A new online statistical learning method is presented in [63] to model the background appearance variations under cast shadows. This is based on the direct light sources and ambient dichromatic reflection model. The first stage that authors propose is using one Gaussian Mixture Model (GMM) to study the colour features. The second stage is building up one pixelbased GMM for each pixel to learn the local shadow features. This is going to be a slow process and requires numerous computer resources so to overcome the slow convergence rate in the conventional GMM learning, pixel-based GMMs are updated through confidence-rated learning. The proposed algorithm, can learn model parameters very quickly.

In [97] an algorithm is presented to detect and remove cast shadows in video sequences by analysing statistical dominance of the shadowed regions and non-shadowed regions. This method has some advantages: It does not require a colour space transformation and works in RGB space. It is data-driven and adapts very well to the changing shadow conditions and in recent studies [9],citeref108 and [1] this method has been improved by reducing the required training data.

Recently Khan et al [71] have developed a solution which can detect shadows from a single coloured image. Their solution automatically

learns the most relevant features in a supervised manner using multiple convolutional deep neural networks. The quality of their shadow-less image is very good but the limitation is their solution struggles to successfully distinguish shadows from dark objects. Double threshold neural network solution is presented in [68] to reduce the false detection of shadows in a single image where the hue and brightness of some non-shadow regions are similar or even lower than those of shadows. To solve the problem of detection shadows from dark objects, two different dynamic thresholds that iteratively updated are designed. The upper and lower limits of detecting shadows are determined respectively by a higher threshold that decreases gradually overtime and a lower one that increases. The detection result is obtained by a combination of two detection thresholds.

2.3.2 Texture and characteristic estimation methods

In recent years alternative methods have been developed to remove the effect of illumination from a single image. These methods are based primary on distinguishing between the texture of the objects and shadows. Bell and Freeman [38] took a learning base approach and generated a training set of images containing shading and reflectance variations. In [120] initially a part labelling which densely covers the object is defined and then Layout Consistent Random Field "LayoutCR" model is imposed asymmetric local spatial constraints on labels to ensure the consistent layout of parts.

Tappen, Freeman and Adelson [113] have used multiple cues to recover shading and reflectance intrinsic images from a single image. They used both colour and information and a trained classifier to recognise greyscale patterns in images and each image is then classified as being caused by shading or a change in the reflectance model.

A texture based method is presented in [106] where authors have decomposed an image into its shadow and intrinsic reflectance component by examining the texture information to gather constraints on reflectance between pixels that may not be close to another another pixel in the image. In this method texture can be represented by any model, such as filter responses or textons. However for simplicity and reducing the computing requirements hey represented the texture at a pixel as a vector of concatenated pixel values from its surrounding neighbour pixels.

Another texture based approach is presented in [69] where each pixel is modelled as a group of adaptive local binary pattern histograms that are calculated over a circular region around the pixel. This is achieved in two stages: background modelling and foreground detection. For the first step, each pixel of the background is modelled identically, which allows for a high-speed parallel implementation which can be very useful in real-time processing. In this method LBP was selected as the measure of texture because of its good properties. For foreground detection a very simple method is used. The histogram h is compared with the existing B background histograms using the same proximity measure as in the update algorithm. If the proximity is higher than the threshold T_P for at least

one background histogram, the pixel is classified as background. Otherwise, the pixel is selected as foreground.

Texture analysis is an important topic in image processing. Malik and Liuin [83] categorized different materials such as concrete, rug, marble, or leather on the basis of their textural appearance and they devid them into two surface attributes: (1) reflectance and (2) surface normal. In their work they provided a unified model to address both these aspects of natural texture. Get et al. [54] presented gloss as an appearance attribute that could reveal certain information about object properties and concluded that observers can apprehend the physical nature of the surface of real objects from features that are included in the BRDF and available in the gloss appearance. Obein [75] provided a measurement to calculate the strength if gloss which they explained it as attribute of visual appearance that originates from the geometrical distribution of the light reflected by the polished surface. they used the maximum likelihood difference scaling (MLDS) procedure by Maloney [85] to estimate gloss scales over an extended range.

Arora [32] presented a new approach to evaluate the illuminant chromaticity which does not need exact correspondences and has a better estimate of illuminant chromaticity. They used the evaluated chromaticity to project the input images on to a specular invariant colour space and showed that standard optical flow algorithms on this colour space significantly improves the results. Other works such as [98],[124] and [122] provide similar approach to remove the impact of glare from glossy texture.

Multi-view solutions have been extensively researched and they provide an alternative approach to the issue of irregular reflectance from glossy textures. Li and Liu [81] reconstructed 3D fine-scale surface models for non-Lambertian objects from multi-view multi-illumination image sets. Their solution provides good results but requires multiple images from different angel to produce the 3-D model. This is a very effective and popular approach and many other works such as [37] , [60] and [101] have used the same method. Although muti-view methods produce good results they require users to take pictures from different angels manually. Unlike these methods, our system detects the non-lambertian texture automatically and does not need any manual interaction.

Texture analysis is a very important topic in image and video processing and has many application such as road detection [29],[127],[91],[58] which is very popular in road traffic surveillance systems. Human detection is another popular use case for texture analysis in image processing. Kim and park [73] proposed an algorithm to recognize human presence with USB Web camera. Their method detects human movement using the circle detection and morphological methods. There are similar studies for human detection such as [70] and [105] where authors used neural network based training methods for human detection.

This thesis proposed a new solution to detect glossy and shiny textures in outdoor scenarios and this solution can be used to detect water, oil leakage metal and glass in natural environments.

2.3.3 Reflectance and glare analysis

An important topic discussed in this thesis is analysing the impact of glare and reflectance in images. The noise generated by excessive reflectance can cause major issues in tracking and object recognition system. The amount of light reflected by a surface, and how it is reflected, highly depends to the texture of the surface. The reflectance can be categorised into two different types:

- Specular reflection
- Deffuse reflectance

Specular reflectance occurs when the light reflects from glossy materials and creates a white patch. Deffuse reflectance happens on other scenarios where light reflects from matt and rough surfaces. Most materials have specular reflectance if they are polished but some materials such as water, mirror and glass can *only* have specular reflectance. In image processing applications, excessive specular reflectance can reduce the quality of the images and have negative impact on applications such as object recognition and tracking. The difference between how various materials reflect the light is illustrated in Figure 2.1.

Ortiz and Torres [94] proposed a method for the detection and elimination of specular reflectance in colour images of real scenes. They used a two-dimensional histogram which is used to relate the signals of intensity and saturation of colour images and also to identify the specularities in an area of the histogram. This is known as the Intensity-Saturation (MS) diagram, and it is constructed from the Intensity-Saturation-Hue (MSH) generalized colour space. The use of a new connected vectorial filter would

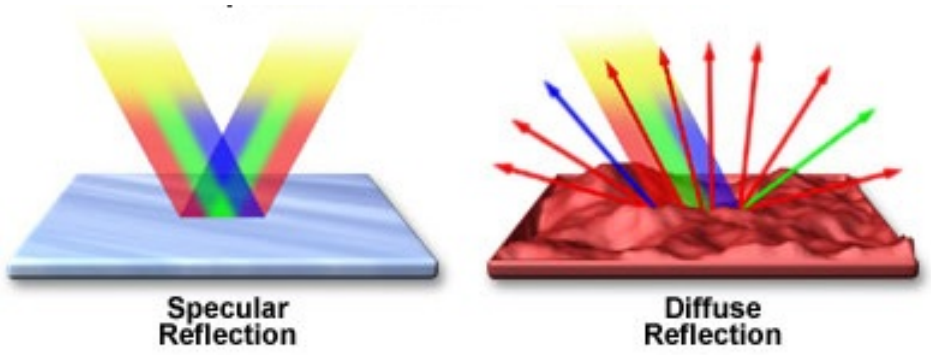


FIGURE 2.1: Defuse and specular reflectance

eliminate the specular reflectance previously detected. The chromatic reflection model proposed by Safer [104] is a common system that has been used in many methods to detect specularities. These models supposes that the interaction between the light and a dielectric material produces different spectral distributions in the object.

Based on this model, Lin et al [82] have designed a system for eliminating specularities in image sequences using stereo correspondence in a controlled environment. Bajcsy et al. [33] used a chromatic space based on polar coordinates that allows the detection of specular and diffuse reflections by means of the having some information about the captured scene. This is a simple and fast method but requires manual input from a user. Klinker and Tiala [74] used a pixel clustering algorithm which has been shown to work well in detecting brightness in images of plastic objects. Gershon [55] and Lee [80] use chromatic information for highlight identification. A similar solution is presented by Sato and Ikeuchi [100] who proposed temporal-colour space to extract the specular reflection and the body reflection.

Wangs [118] have designed an HMM based algorithm which is used to detect people in swimming pools. To achieve this a Bayesian approach using Hidden Markov model scheme is developed to enhance human detection for video surveillance system operating under aquatic environment. Pixels in current frame are marked as foreground or background by considering historical information of its surrounding frames and neighbouring pixels between the current processed frame and the reference image.

Another approach to reducing the noise caused by light specular reflection is putting a polarizing filter set in the front of the camera lens. This solution has been introduced by Fujikake et al in [35]. However, for this system to work polarisation angle of specular reflection should be calculated to be used as optical axis of the filter. This is not always an easy process and manual configuration is required. Nayar and Narasimhan [93] designed a novel method based on physical characteristics of light in order to remove weather effects from images. Their model is valid for various weather conditions including fog, haze and also provides the capability to recover from a single image the shapes and depths of sources in the scene. In addition, the weather condition and the visibility of the atmosphere can be estimated. These quantities can, in turn, be used to remove the glows of sources to obtain a clear picture of the scene. Although this solution tackles the problems of light scattering very well, it can't be used to address the issue of extensive reflectance in the image.

To improve image quality and reduce the impact of glare Talabala et al [48] used multiple captures with a high-frequency occlusion mask to

estimate and remove glare. Glare is reduced before image formation, providing high-quality reconstructions. The most significant weakness of their method is that it requires a significant volume of images to record a scene and it is only applicable to static scenes. A complete capture with their setup takes from 30 minutes to an hour. The other major limitation is that the mask must be nearly in focus to limit the mixture pixel region. This mandates either using a very small aperture, or placing the mask near the scene rather than on the camera lens like a filter. Veiling glare is a commonly-acknowledged issue in photography. The standard method for glare measurement involves photographing a central black target surrounded by a large uniform bright illuminant Matsuda [86] and Kuwabara [77] in 1952. ISO standard 9358 defines the ratio of luminance in the center of the target to the luminance of the illuminant as the veiling glare index (VGI). The standard also defines the glare spread function (GSF), which defines the amount of glare created by a small bright spot as a function of the distance from the center of the spot. Directly similar to [48], McCann and Rizzi [89] have measured glare in multiexposure HDR imaging.

Additionally there are many other research that can model background variation to reduce the image of foreground noise. For example Mikito [102] has analysed the property of neighbouring blocks to reduced the impact of moving trees and flags due to wind. Monnet et al [90] proposed a linear predictive model to model a dynamic sea shore environment. Any pixel that deviated significantly from its predicated value is declared as foreground. The main advantage of the predictive technique is that it reduces some uncertainties of a pixel's value by checking how it changes

over time.

Recently Tao et al [112] investigated the characteristics of pixel values from different viewpoints in colour space and they presented a new and practical approach that uses light-field data to estimate light colour accurately. Their algorithm will allow users to manually acquire depth maps using a consume camera which works very well on glossy material. The experiments illustrated in this research show high quality coloured images but the weakness is manual intervention is required. Without multiple viewpoints, most diffuse and specular separation methods assume the light source colour is known and consistent such as the work by Mallick [84] and Yoon [123] where they compute a data-dependent rotation of RGB colour space and showed that the specular reflection effects can be separated from the much simple effects for surfaces that can be modeled with dichromatic reflectance. Lee and Leonardis [34] and other studies such [110],[111] and [121] have also came up with with single image based methods where a modest impact of reflectance is removed from glossy materials. Kin et al [72] have also observed that for most natural images the dark channel can provide an approximate specular-free image. they propose a maximum a posteriori formulation which recovers the specular reflection and chromaticity despite of the hue-saturation ambiguity. The limitation of these methods is they are very sensitive to noise and they only work well in artificial and controlled environment.

Using advanced camera technologies for specular reflection analysis

The camera technology and its characteristics and capabilities have improved significantly in recent years. There are many methods which require specific type of camera for glare and specular reflectance analysis. Iwata has Ogata [65] have used high-speed cameras for this problem. system utilizing a high-speed camera and a strong flash for removing specular reflection. This method utilizes a principle of estimation using images with luminances varied by the flickering of a strobe. They used optical flow, a position of specular reflection compensated for the noise generated by inter-frame difference when the high–speed camera is moved. This study is based on reflectance model introduced by Tsuji [117] and [116]. In this model the strobe–off image is displayed as by I_n, and the strobe-on image is represented by I_e. The strobe-off image is taken specifically under natural light. This image contains a diffuse reflection component and specular reflection component. Expressing the strobe–off image by using Shafer[103] model enables the following equation to be derived:

$$I_n = I_{ns} + I_{nd} \tag{2.3}$$

where I_{nd} is the diffuse reflection component and I_{ns} is the specular reflection component.

In a similar study, Feris and Rashkar [50] used multiple image flashing to reduce the effect of specular reflectance in digital images. Their approach relies on a simple setup of a multi-flash camera take multiple pictures of the scene, each one with a differently positioned source of light.

They formulated the problem of specular highlights reduction as solving a Poisson equation on a gradient field from multiple input images. Their work belongs to an emerging class of computer graphics techniques that process multiple images taken with the same viewpoint but under different conditions, such as different illumination, exposure and focus such as [28],[43] and [99]. This solutions can remove specular reflectance as well as shadows in a controlled environment. But unlike these methods, our method which is presented in chapter 5, does not require specific type of camera and also it can remove reflectance and glare in normal real-life images.

2.4 Cloud Detection

Another important topic that has been discussed in this thesis is a new method to detect the presence of cloud automatically. Cloud detection is a very important issue for satellites images. Detecting cloud regions in remote sensing image (RSI) is a difficult challenge but it is very important meteorological forecasting and other RSI-related applications. Technically, this task is typically implemented as a pixel-level segmentation. However, traditional methods based on handcrafted or low-level cloud features often fail to achieve satisfactory performances from images with bright non-cloud and semitransparent cloud regions.

There are many studies in this area such as work by Yuan in [11]. Deep learning is another popular approach and [14] studies the potential of deep learning algorithms for cloud detection to improve state-of-the-art performance. A comparison between deep learning methods used with

classical handcrafted features and classical convolutional neural networks is performed for cloud detection is presented in this study. There many other studies based on deep learning methods such as [12], [26] and [2] which authors take similar approach.

Additional studies which investigate cloud detection is special cases such as [7] which focuses on coastal areas. Surya [20] proposes a new approach which uses wavelet transform, histogram analysis and clustering for cloud detection. In his work, remote sensing satellite images can be considered as a signal. Clouds are low frequency components in satellite image. Therefore wavelet analysis of satellite images is helpful to extract low frequency components and detection of clouds. The result analysis shows that the proposed method can detect clouds automatically from satellite images with high accuracy. Experimental analysis is performed on Landsat-7 Enhanced Thematic Mapper Plus (ETM+) sensor images and results are obtained. Tuplan [5] proposes a computer vision approach to cloud detection consisting of feature extraction and machine learning. Six image moments on local texture regions were extracted and fused within a classification algorithm for discrimination of cloud pixels. Three different popular classifiers were evaluated for efficacy. Tan et. al [10] proposed a 3 stage method for this problem. Firstly, the visual dictionary is learnt from the training features extracted using Maximum Response (MR) Filter. Second, Principle Component Analysis (PCA) is utilized to reduce the dimensions of the visual words for fast word search. Finally, the MR feature of an image patch is converted into the histogram of visual word. Other popular methods for cloud detections are: remote sensing image time series using Mean Shift algorithm [10], morphological

based approaches [18] and [6].

All of the above studies approach this problem from satellite images point of view, but on the other hand, the solution highlighted in this thesis detects clouds from ground level view point. Therefore satellite images are not needed and by analysing the reference shadows over a period time, we have successfully detected the presence of cloud and used that to generate high quality shadowless images. This has been explain fully in chapter 6.

2.5 Comparing the methods

We have compared our shadow detection method with the work of Weiss [119], Matsushita [88] and also other state-of-art methods on shadow removal from videos. The reason we selected Wiess and Matsushita's methods for comparison is both of these methods were designed for outdoor shadow removal from video and require multiple samples for model learning. Overall in outdoor scenarios, our method is very robust in many complicated cases such as noisy environment with moving objects as well as presence of moving clouds. But what makes our method stand out is its simplicity and the time it requires to detect a Reference Shadow successfully. The majority of other video based methods, require hours of training data or manual intervention (such as camera calibration). Nevertheless, we managed to detect shadows in as quick as 10 minutes in really complex noisy environment with numerous moving objects.

Table 2.1 highlights the advantages of our method over some of the existing shadow detection solutions.

TABLE 2.1: Outdoor Shadow Detection Methods

Method	Indoor/ Outdoor	Background/ Foreground	Initialisation Period	Requires GPS Location	Manual/ Automatic
Reference Shadow	Outdoor	Both	Minutes	No	Automatic
Weiss	Both	Both	35 frames in one day	No	Automatic
Matsushita	Outdoor	Both	Hours	No	Manual
Finlayson	Outdoor	Both	Requires camera calibration	No	Automatic
Cavallaro	Both	Foreground	Requires background subtraction	No	Manual
Huerta	Both	Foreground	Single Image	No	Manual
Antone	Outdoor	Both	Hours	Yes	Manual
El-Zahhar	Outdoor	Both	Semi-supervised	No	Automatic
Boroujeni	Both	Both	Hours	No	Manual
Hekkila	Outdoor	Both	Multiple	No	Manual
Russell	Outdoor	Cars	Semi-supervised	Yes	Automatic
Zui	Indoor	Both	Hours	No	Automatic
Carmen	Outdoor	People	Multiple	No	Automatic
Yanfeng	Outdoor	Both	Multiple	Satellite Image	Automatic
Samaras	Outdoor	Foreground	Hours	Yes	Manual

2.6 Conclusion

In this chapter the existing methods of shadow detection were presented and the strength and weakness of state of art was discussed. The findings from literature can be summarised as:

- The neural network based solutions can have good performance in shadow detection if the correct type of training data is provided to the model.
- HSV/HUV based systems are very simple and effective but they struggle to distinguish between shadows and dark objects.
- The methods which reply on camera characteristics have limited functionality and only work with a specific type of camera.
- The GPS location of the camera, time and date of the video is needed for some of the shadow detection systems.
- Most of the cloud detection solutions are based on pictures from satellites.
- Many of the existing system can only detect shadows from moving objects.
- Some of the existing solutions are sensitive to noise and will not perform very well where there are multiple shadows.

The shadow detection method which is introduced in this thesis can successfully detect shadows from moving objects as well as static shadows. It doesn't require the location or recoding time and no user input is

needed and performs very well in noisy environments with many moving foreground objects. Additionally a novel method to detect the presence of cloud is introduced and unlike the existing literature, it does not require satellite images.

Chapter 3

Shadow Detection

This thesis proposes a new method to detect day-time outdoor shadows. Due to the relative movement between the sun and an object the position and length of the shadow cast by the object changes during the day. In this chapter we investigate how this information can be used in real-life scenarios to assist us in detecting shadows.

3.1 Introduction

In previous section 2, existing methods and state of art was discussed. The key weakness of existing shadow detection systems are:

- Only foreground shadows are detected.
- Manual or semi-supervised user input is needed for shadow detection.
- Shadow detection systems are very sensitive to environment noise.

  To have a robust shadow detection system, we need to have a new solution that can address these limitations. The reference shadow solution which is presented in this thesis can detect both foreground

and background shadows. It is fully automatic and it doesn't need any user input and finally it performs very well in real-life noisy environments with many moving objects.

3.2 Definition of reference shadows

Shadow is created when an object blocks sunlight and, consequently some part of the image becomes darker. This is the shadow region. As our planet orbits the sun, the height of sun changes during the day and as a result the length of the shadows changes. In this chapter we investigate how this information could be used in real life scenarios for shadow detection in outdoor scenarios.

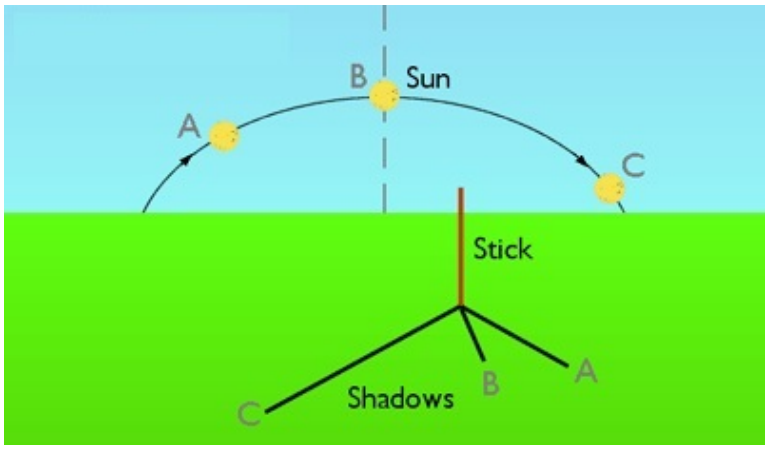


FIGURE 3.1: The shadow of a stick moves from A to B and then C during the day when the sun goes from A to B and then C.

As illustrated in Figure 3.1, the length of a shadow L can be computed as:

$$L = h/tan(\alpha) \tag{3.1}$$

Where h is the height of the object, which casts the shadow and α is the angle between the sun and the horizon. The value of L for α between 0 to 180 degrees is displayed in Figure 3.2. This reveals that when the sun rises and shifts from east to west, the length of the shadow moves in a pattern that is very close to a line. The exception occurs in early morning and just after sunrise and also before the sunset when there is a non-linear pattern.

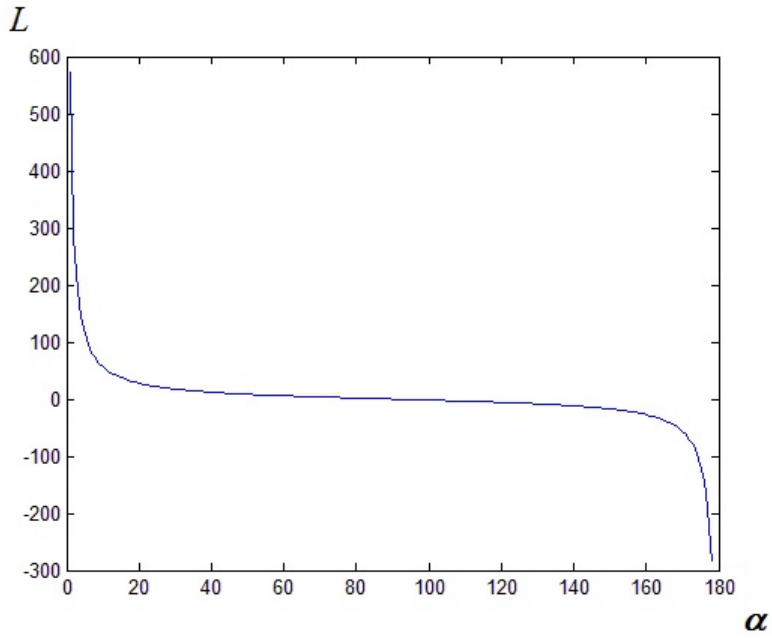


FIGURE 3.2: Length of reference Shadow from sunrise until sunset. The pattern is linear during most of the day; apart from sunrise and sunset.

We analyse and process the position of the centre of the shadow instead of its length. In a short period, the changes to the position of the centre of the shadow will have a similar pattern to the changes to the length. Our experiments confirmed monitoring the centre of the shadow will be sufficient to distinguish reference shadows from other dark objects

in the scene.

Additionally, the impact of other factors such as the shape of the object and camera view will be minimum in short sampling periods and will not change the trajectory of the centre (or corner or edge) of the shadow. Based on our experiments, using the corner of shadows will improve the performance of shadow detection system when the shadows are located at image boundaries. .

The reference shadow is the key element in detecting the remaining shadows in the image and it is determined as:

A dark part of the image which can be categorised as shadow with the highest confidence.

Once we have identified at least one reference shadow, the remaining shadows (from stationary and moving objects) can be detected as described in the next section. This is a vital advantage of this algorithm compared with other shadow detection systems, such as HSV based methods [115], [95] and [59] where they fail to distinguish between shadows and objects with similar dark colour.

Identifying one shadow provides very important information about the characteristic of the image and the scene:

- The trajectory of the centroid of reference shadow over time can be used to estimate the pattern of movement of other shadows.
- The reference shadow edges provide valuable data about the illumination changes between shadow and non-shadowed regions.

In normal outdoor conditions where the sun light is prominent, all the objects that block sun will produce shadows. Additionally, the objects

with the same angle and similar shapes will produce shadows which will have similar movement overtime. This is due to the long distance between these object and the source of the light (i.e. the sun) which means the α is the same for all of these objects.

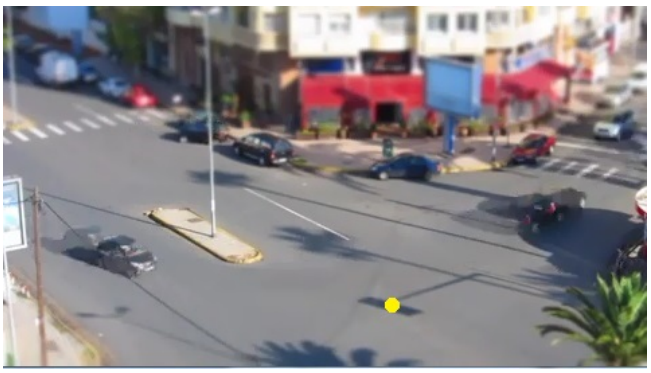
For example, all the stationary objects such as buildings, lamp-posts, traffic lights and people will have their shadows cast in the same direction. In normal video surveillance or traffic monitoring systems plenty of objects would create reference shadows. Based on our experience at least one reference shadow can be detected easily.

3.3 Detecting reference shadows

The height of sun varies during the day, therefore, the length and position of the shadows changes. This can be used as a very good cue to detect some of the shadows. To do this we need to monitor some aspects (such as centre or corner) of dark objects over a period. In most real-time scenarios, the result will be very noisy because other foreground objects move over the shadows and there might be some minor camera movement. Nevertheless, the patterns of shadows and dark regions will differ completely over time. To explain this in more detail, we will use the example in Figure 3.3, which is based on 3000 frames from a 20 minutes video from a surveillance camera which is pointing towards a busy junction in Morocco.

As highlighted in yellow, the centre of the shadow of a vertical object (possibly a road sign) has slightly moved to the left. This is because of the relative movement of the sun which starts from sunrise to sunset and

moves from east to west. Using this information to detect shadows would have many advantages. For example, there is no limit in the location, time of day or position of the camera. If the camera is stable and there is a single stationary object that blocks the sun, we can detect a reference shadow. In other sections we demonstrate that this solution can detect shadows automatically in minutes unlike many other methods such as **[11]**, **[14]** and **[20]** which requires manual intervention by users or hours of training.



(a) Frame 1

(b) Frame 3000

FIGURE 3.3: Trajectory of the centre of the shadow has changed in 3000 frames which is marked in yellow.

3.3.1 Pre-processing

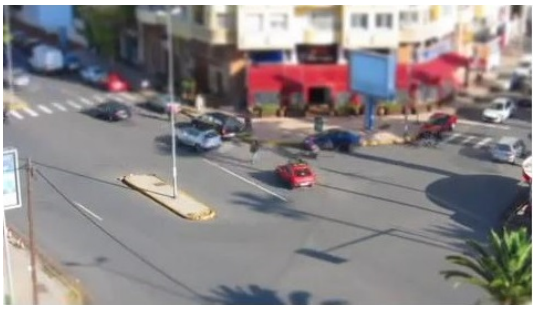
In this example, the following steps have been taken to detect a reference shadow:

1. The RGB image is converted to a binary image.
2. An averaging filter filters the binary image.
3. The position of the centre of the black regions is recorded across all frames after going through a Registration Process.

To successfully identify the reference shadows, we have made a couple of basic assumptions: 1) camera is stationary 2) shadows are darker than the background. These assumptions are true and applicable in all of the video surveillance application where the scene is viewed from CCTV cameras. The filtering process is illustrated in Figure 3.4. In this example, which is taken from a busy real-life scenario, many back regions are produced at the end of pre-processing sections as evidenced in 3.4(c) These black pixels could be moving objects, shadows or dark stationary objects. We now demonstrate how we can categorise these regions and successfully identify a reference shadow.

3.3.2 Registration process

Now that Figure 3.4.(c) is produced, we introduced a registration process to track and analyse the position of the centre of dark regions across all available frames and ensure the categorisation of dark regions has been completed with minimum error. To do this, we register only dark regions in each frame if both of the following two conditions are true:



(a) The original RGB image

(b) RGB image is converted to binary

(c) The binary image is then filtered by an averaging filter.

FIGURE 3.4: Pre-processing of RGB image.

1. The position of the centre of the dark region in the new frame should be very close to the position of the dark region in the previous registered frame.

2. The size (number of pixels) of the dark region in the new frame should be very close to the size of the dark region in the previous registered frame.

The first condition is in place because we know the centroid of shadow will move very slowly over a short period. This is because the value of α changes very slowly and hence the length of the shadow will change slowly as per L = h / tan (α). Since the length of the shadow is increasing or decreasing slowly, the centroid of the shadow will also move across x an y axis very slowly. The second condition is designed to filter the frames where a moving object goes over the cast shadow of the stationary object that has produced the reference shadow. These conditions make the shadow detection very robust and can successfully identify shadows in very busy environments. This has been demonstrated in the experiment section.

If any of the above conditions are false, we discard the dark region and do not register the coordination of the centre and select 0 for the value of x and y of the dark region in that frame. Moreover, we have also used a counter to determine how confident we are in our readings. When there is a new reading, the counter will be increased by one, and hence at each point in time, we know the confidence in each dark region by comparing the value of the counter of each dark region. In the example provided in Figure 4, more than 2900 different dark regions were detected across the 3000 frames but we only need to analyse the regions which have a high number of counters. In this example and by selecting regions with counter higher than 200, the number of dark region candidates reduced to around 50 from the initial 2900. This is illustrated in Figure 3.5. The process of registration is displayed in Figure 3.6. The registered dark region is highlighted in green and the discarded dark region is shown in red. In

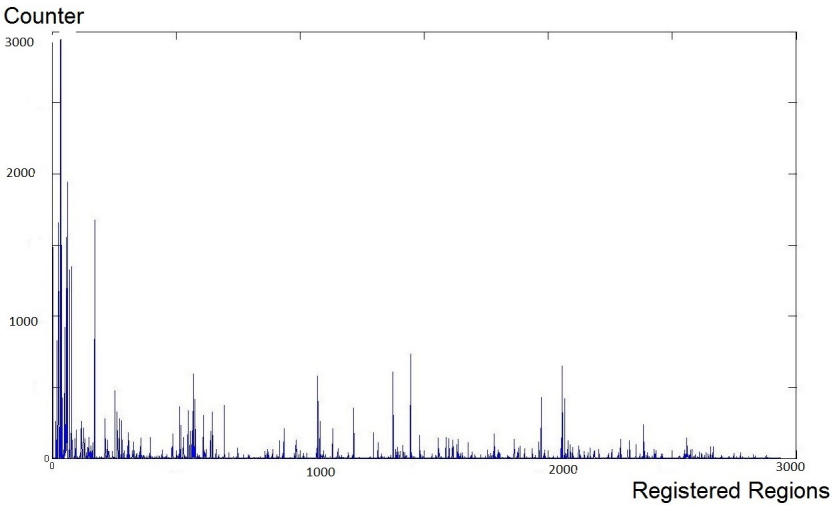


FIGURE 3.5: The value of Counter for each registered dark region. Around 2900 different dark regions have identified during the 20 minutes of video footage.

this example, the second condition of the registration process (dark region size) has been violated because the size of dark region has altered significantly in one of the frames and therefore the coordination of the centroid is discarded for that frame. In most cases, this is because a moving foreground object - possibly a car, has moved over the shadow. To calculate the size of dark region, the number of the pixels were counted using Matlab functions.

3.3.3 Selection process

When all the frames have gone through the pre-processing and registration process, we recorded the positions of the centre of all of the registered dark regions and stored this data in a matrix. The results are presented in Figure 3.7 In Figures 3.7-(b) and 3.7-(c), the x values of the centre of two dark regions are displayed. Region A is one shadow (or part of the shadow) of a stationary object; possibly a road sign and region B is the

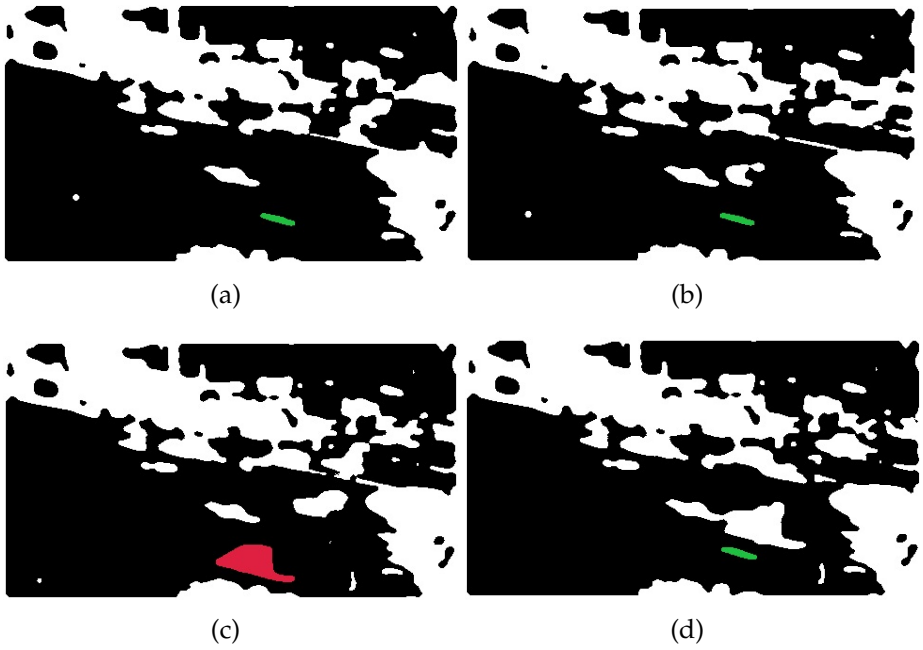


FIGURE 3.6: Four consequence frames. All of the dark regions (dark objects and shadows) are highlighted in white. The registered dark regions are displayed in green, and the unregistered dark region is displayed in red. The reason the dark region in the frame (c) is not registered is because another object (possibly a car) has gone over the dark region and hence the size of this dark region has changed significantly from previously registered frame.

dark car which is parked and not moving during the 20 minutes recording. These two regions are marked in Figure 7-(a). The gaps in the readings are due to the frames discarded during registrations process which was explained earlier.

As evidenced in Figure 7, the centre of the shadow region changes gradually over the period of time while the centre of the dark object stays the same and does not change.

Therefore the key question to answer when detecting shadows is ***how we can distinguish between dark objects and shadows*** and the answer is given in Figure 3.7(b) and 3.7(c). If we analyse the dark regions for a while and then check the trajectory of the centre (or corners) of dark objects, the

shadow and non-shadow regions will show different patterns.

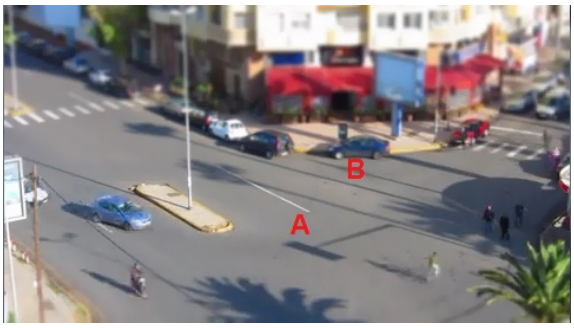
This is identical to our observation in the experiment from the surveillance video in Morocco. As highlighted in Figure 3.7.(a) the centre of the shadow region changes linearly during the 20 minutes observation at midday; thereby matching the pattern illustrated in Figure 3.2.

To implement this, we fit a line using the Hough transform model [62] from registered centroids and then check the angle of the line. If the angle of the line is horizontal, i.e. the centre has not changed over the time and the dark region is not a shadow. If the angle of the line is not 0, we have successfully identified a reference shadow.

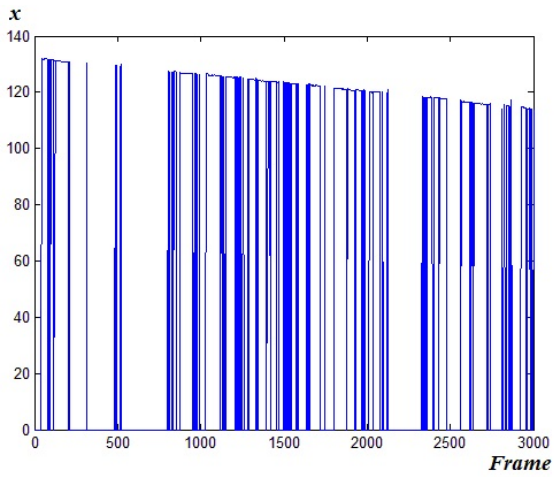
It should be noted that that in most real life cases, the measurements will be very noisy. For example, in Figure 3.7-(b) and after frame 500, there is a relatively large gap until the next reading. The reason for this is between frame 500 and 820, there is a traffic jam for few minutes and a car has stopped over the shadow A. Hence the coordinates are not registered and the 0 is used as the value as explained previously.

Another question to answer is how many frames should be processed to identify a reference shadow and the answer depends on the time of the day, time of the year and the location in which the video is taken. In our experiments, analysing 5-15 minutes of video footage was sufficient to detect a reliable reference shadow in noisy environments.

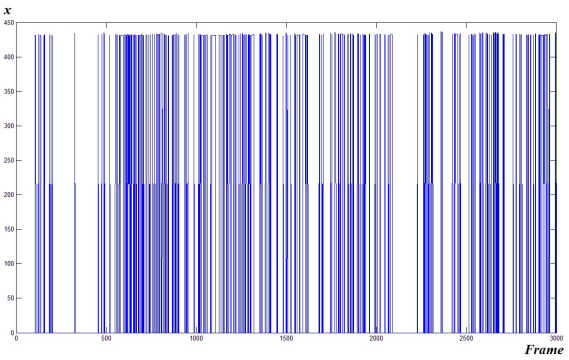
Additionally it is possible to identify more than one reference shadow and if this is the case, we can always select the one with the highest counter. For example in our video sample from Morocco, we have found two other reference shadows with the counter value of 1944 and 1677.



(a)



(b)



(c)

FIGURE 3.7: (a) Dark region A is a shadow; Dark region B is not a shadow and it is dark car. (b) x value of the centre of object A over the 3000 frames and (c) x value of the centre of object B over the 3000 frames.

It is important to point out that we do not need to identify all shadows and if only one reference shadow (which could be in fact a part of

a shadow) is identified, the rest of the shadows can be detected as explained in the next chapter. In real life scenarios where we would like to track objects using outdoor CCTV cameras and with very high confidence we can assume that at least one reference shadow exists. For example if the camera is pointed at a busy junction or a roundabout, there is a very good chance to have either a tree or traffic light or building in the image. These objects will obstruct light and will create reference shadows.

As explained above, this method initially identifies many candidates for reference shadow; for example in the video footage from Morocco, more than 2900 different dark objects are detected during the 20 minutes. This is displayed in Figure 3.5 Therefore the probability of identifying at least one reference shadow among several thousand candidates is very high.

To qualify as reference shadow, the object which creates the shadow should be stationary. For example on a windy day, leaves of tree will not yield reliable reference shadow. However in the same scenario, the tree trunk can successfully generate reference shadow.

Identifying one shadow is the most important step in shadow detection and removal. Once this stage is completed the remaining shadows can be detected easily by using various methods as explained in the following chapter. Reference shadow is a very simple but effective way for shadow detection in outdoor scenarios and its most important advantage to other systems is it works very well in complex lighting conditions as well as busy environments. In the next, we illustrate that this method can also be used to detect shadows at times close to sunrise and sunset to detect shadows. The sole difference between this scenario and other time

of the day is the centroid of the dark region shows a non-linear pattern. However this pattern is still different to other dark regions and it can be used to detect shadows and dark objects successfully.

3.3.4 Implementation

In this section, the methodology to implement automatic shadow detection in Matlab is explained. The key component of our algorithm is producing a matrix called *master*. This matrix is a database which contains the following information about all of the dark regions in all of the available frames:

Fixed Master matrix columns:

1. Object number
2. The first frame that the object was first observed.
3. The number of the time the object was observed (counter).
4. The frame number that the object was last seen.
5. Is the object still in the scene (1=in,0=out);
6. The number of the pixel in the object.
7. The object initial centroid X.
8. The object initial centroid Y.
9. The object new centroid X.
10. The object new centroid Y.

11. The object new pixels.

The size of Master matrix is *n x m* where *n* is total number of all the separate dark regions seen in all of frames. These dark regions could shadows or moving objects or noise etc. In a typical 10 minutes video, the size of the master database could be tens of thousands of rows but there is a *purge* function which is developed to remove the unwanted entries in the database. This function is developed to improve the matching capability so the system can easily deal with environments with lots of fast moving objects. *m* is 11 fixed columns plus 2 columns (Centroid X and Y) for any frame that the object is seen. An example from *master* database is presented in Figure 3.8

Object Number	First seen Frame	Counter	Last seen Fram	In or Out	Total Number of Pixels	Initial X	Initial Y	New X	New Y	New X	New Y	New X	New Y	New X	New Y	New X	New Y	New X	New Y
1	1	204	264	0	25138	174.685	70.63676	173.3393	70.74636	174.6925	70.63623	172.8043	66.08381	173.156	64.57092	173.5803	64.32259	173.7798	64.28987
2	1	2976	3000	1	6	1.666667	269.3333	1.666667	269.3333	1.666667	269.3333	1.666667	269.3333	1.666667	269.3333	1.666667	269.3333	1.666667	269.3333
3	1	2	2	0	791	59.74842	121.9709	59.74842	121.9709	59.74842	121.9709	0	0	0	0	0	0	0	0
4	1	2	2	0	105	107.1143	94.52381	107.1143	94.52381	107.1143	94.52381	0	0	0	0	0	0	0	0
5	1	2	2	0	303	130.9901	129.0099	130.9901	129.0099	130.9901	129.0099	0	0	0	0	0	0	0	0
6	1	3	31	0	891	172.4513	257.5947	174.8496	257.8855	172.4513	257.5947	174.8496	257.8855	0	0	0	0	0	0
7	1	2	2	0	69	189.2464	247.6377	189.2464	247.6377	189.2464	247.6377	0	0	0	0	0	0	0	0
8	1	2	2	0	220	215.7091	254.6636	215.7091	254.6636	215.7091	254.6636	0	0	0	0	0	0	0	0
9	1	2	2	0	543	254.2836	159.1087	254.2836	159.1087	254.2836	159.1087	0	0	0	0	0	0	0	0
10	1	2	2	0	490	254.3633	264.649	254.3633	264.649	254.3633	264.649	0	0	0	0	0	0	0	0
11	1	2	2	0	594	255.0505	221.1448	255.0505	221.1448	255.0505	221.1448	0	0	0	0	0	0	0	0
12	1	2	2	0	23	261.6957	180.8696	261.6957	180.8696	261.6957	180.8696	0	0	0	0	0	0	0	0
13	1	2	2	0	41	266.8537	148	266.8537	148	266.8537	148	0	0	0	0	0	0	0	0
14	1	2	2	0	10	277	187.5	277	187.5	277	187.5	0	0	0	0	0	0	0	0
15	1	2	2	0	14	279.1429	144.0714	279.1429	144.0714	279.1429	144.0714	0	0	0	0	0	0	0	0
16	1	260	557	0	109	301.9367	268.8987	297.0183	268.0275	301.9367	268.8987	302	268.8095	301.8171	268.8415	301.8171	268.8415	301.8171	268.8415
17	1	2	2	0	24	291.3333	214.8333	291.3333	214.8333	291.3333	214.8333	0	0	0	0	0	0	0	0
18	1	2	2	0	84	309.7262	230.0952	309.7262	230.0952	309.7262	230.0952	0	0	0	0	0	0	0	0
19	1	2	2	0	2	306	216.5	306	216.5	306	216.5	0	0	0	0	0	0	0	0
20	1	209	256	0	20	314	45.86364	314.1	45.95	314	45.86364	313.913	45.34783	313.913	45.34783	313.913	45.34783	313.913	45.34783
21	1	2	2	0	19	346.1053	42.57895	346.1053	42.57895	346.1053	42.57895	0	0	0	0	0	0	0	0
22	1	69	956	0	32	347.0938	109.9375	349.3125	108.0938	347.0938	109.9375	347.0938	109.9375	348	109.2333	347.7188	110	348.2414	109.3793
23	1	2	2	0	9	356.1111	142.4444	356.1111	142.4444	356.1111	142.4444	0	0	0	0	0	0	0	0
24	1	39	65	0	2465	397.7193	94.02662	397.3181	92.76876	397.7193	94.02662	397.848	94.92282	397.3905	94.88319	394.6406	93.28813	394.6513	93.57811
25	1	395	2197	0	11165	441.2337	202.3474	433.1088	193.0497	441.2337	202.3474	441.4724	200.6051	441.9329	197.1996	441.9389	196.2627	442.5754	195.8777
26	1	1655	1736	0	242	410.7323	50.81818	411.1612	50.0124	410.7323	50.81818	410.7163	50.70192	410.8439	50.69268	410.8439	50.69268	410.8439	50.69268
27	1	702	2516	0	123	411.4242	243.1288	413.5285	243.5366	411.4242	243.1288	411.3969	243.1527	411.4427	243.1069	411.1241	243.0365	411.336	242.704
28	1	2	2	0	1	417	118	417	118	417	118	0	0	0	0	0	0	0	0
29	1	346	2988	1	10	432	270	433.5	270	432	270	432	270	432	270	432	270	432	270
30	1	3	3	0	16	435.5	31	435.5	31	435.5	31	435.5	31	0	0	0	0	0	0

FIGURE 3.8: The first 30 rows of Master database is displayed. Dark objects with high counter value will be considered to be reference shadow. In this example the centroid of object number 1 is slightly changing in each frame so it could be a candidate. But object 2 is completely stationary so it can't be a shadow and it is probably a stationary dark object.

3.4 Conclusion

In this chapter, a novel method to detect shadows automatically is provided. As explained, the reference shadows can be detected by analysing how the length of dark regions changes over a period of time. For thousands of years, the length of shadows were used to identify the time using sundail and here the same concept is used in video processing to automatically distinguish shadows from dark regions. This is a very simple but effective approach which can be used in all outdoor environments.

Chapter 4

Shadow Removal

A new and robust method to identify at one shadow was presented in the previous chapter. In this section, we provide a novel solution for shadow removal and explain the advantages of this system. The objective of this solution is to provide shadowless images which can then be used by other applications such as tracking systems, shape recognition and so on.
In chapter 1, an introduction to this thesis was provided. Chapter 2 was literature review where the existing shadow detection methods were analysed and their strength and weaknesses were discussed and reviewed. In the previous chapter 3 and we presented a novel method which can be used to automatically detect shadows in outdoor environments which is called reference shadow. The advantage of reference shadow method is:

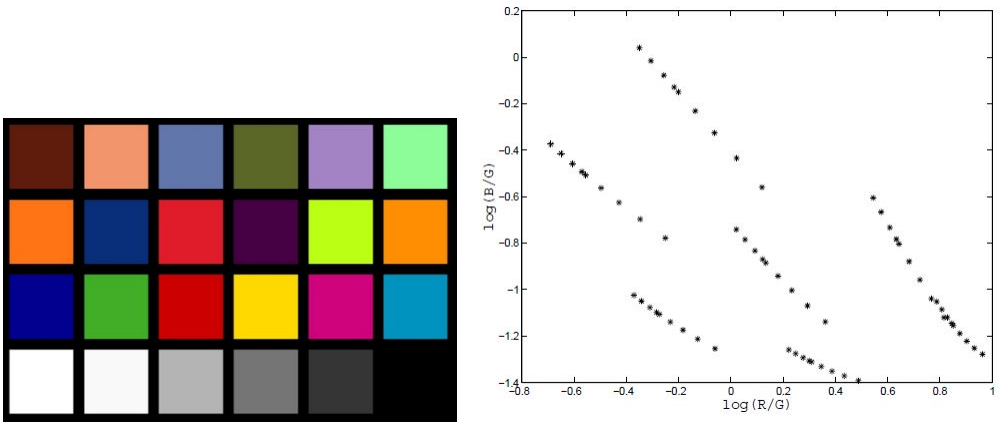
1. It detect shadows automatically and doesn't require user input.

2. Shadows from both foreground and background are detected.

3. It works very well in noisy environment with lots of moving objects.

4.1 Definition of illumination invariant image

If we imagine a set of coloured surfaces under Planckian light, in a controlled light box, say. If surfaces are Lambertian and not shiny, then for each colour the log of the chromaticity r,g will be presented as a single dot in a 2-d plot As displayed in Figure 4.1



(a) (b)

FIGURE 4.1: (a) Macbeth colour chart (b) Plot of each colour on Log(B/G) and Log(R/G)

Now that we have identified one reference shadow with high confidence, it is possible to detect the remaining shadows in the image[51]. The RGB colour ρ for each of the three channels K where K = (R, G, and B) is given by:

$$\rho_k = Ic_1\lambda_k^{-5}e^{-(c_2/T\lambda_k)}S(\lambda_k)q_k \tag{4.1}$$

Assuming that lighting can be approximated by Plank's law, constants c_1 and c_2 are equal to 3.74183 $\times$ 10^{-16} Wm^2 and 1.4388 $\times$ 10^{-2} mK. The

variable I controls the intensity and $S(\lambda_k)$ is surface spectral reflectance function. Now to form band-ratio chromaticity from colour values ρ,

$$r_k = \frac{\rho_k}{\rho_G} \tag{4.2}$$

We divide this by green and calculate $\frac{R}{G}$ and $\frac{B}{G}$ to remove the intensity information and by doing this, the intensity information I will be removed. Moreover, to removing the temperature term, the *log* of equation 4.2 will be taken to form:

$$r'_k = log(r_k) = log(s_k/s_N) + (e_k - e_N)/T \tag{4.3}$$

In this equation, $s_k = c_1\lambda_k^-5S(\lambda_k)q_k$ and $e_k = -c_2/\lambda_k$ and when temperature changes, the two vectors r' and k will form a straight line. In [51] camera calibration has been used to determine the vector direction $(e_k - e_N)$ in the space of logs of ratios. The invariant image is then formed by projecting 2-d colours into the direction orthogonal to the vector $(e_k - e_N)$. The result of this projection is a single scalar which is coded as a greyscale value. Therefore, the greyscale invariant is defined as:

$$gs = c_1R'(R) - c_2R'(B) \tag{4.4}$$

Where in RGB space, $R'(R) = log(R/G)$ and $R'(B) = Log(B/G)$. c_1 and c_2 are constants such that the vector [c_1 c_2] is in the direction is orthogonal to the lighting direction.

Experiments have confirmed images with a different level of illumination will map to the same greyscale invariant image. The most important benefit of this image is shadowless (which occur when there is a change in luminance) will disappear. Now if we divide equation (4.4) by c_1, we get the following:

$$gs' = R'(R) - CR'(B) \tag{4.5}$$

where $C = c_2/c_1$

4.2 Producing illumination invariant image using reference shadows

Others such as Finlayson [51] have used manual camera calibration to identify c_1 and c_2 but we use the reference shadow to automatically create greyscale invariant image by estimating the value of C. To create the invariant image without camera calibration, we generate a large number of gs' with different values of C in [0,1].

At some point, $[c_1, c_2]$ becomes orthogonal to lighting direction and

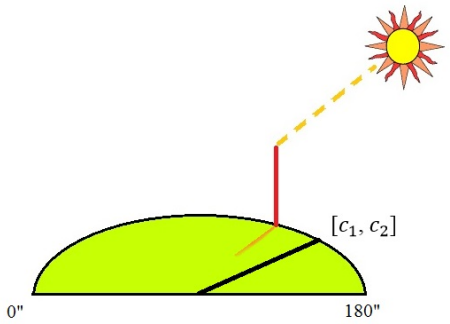


FIGURE 4.2: When $[c_1, c_2]$ changes with the lighting angle from 0 to 180, at some point it will become orthogonal with the lighting direction.

shadow edges will disappear. To automatically detect the shadow edges, we take the following steps listed in Algorithm 1.

Algorithm 1 Generating greyscale invariant image

1: Select $C \in [0\ ,\ 1]$ to generate gs' samples.
2: **For all** gs' **do**
3: Create Edge Map. (we used Sobel method).
4: Take multiple block samples from inside and outside of reference shadows boundary.
5: Find the gs' sample with the least difference between inside and outside the edge.
6: **End For**
7: The gs' sample identified in step 5 is gray-scale invariant image and the corresponding parameter C is set.

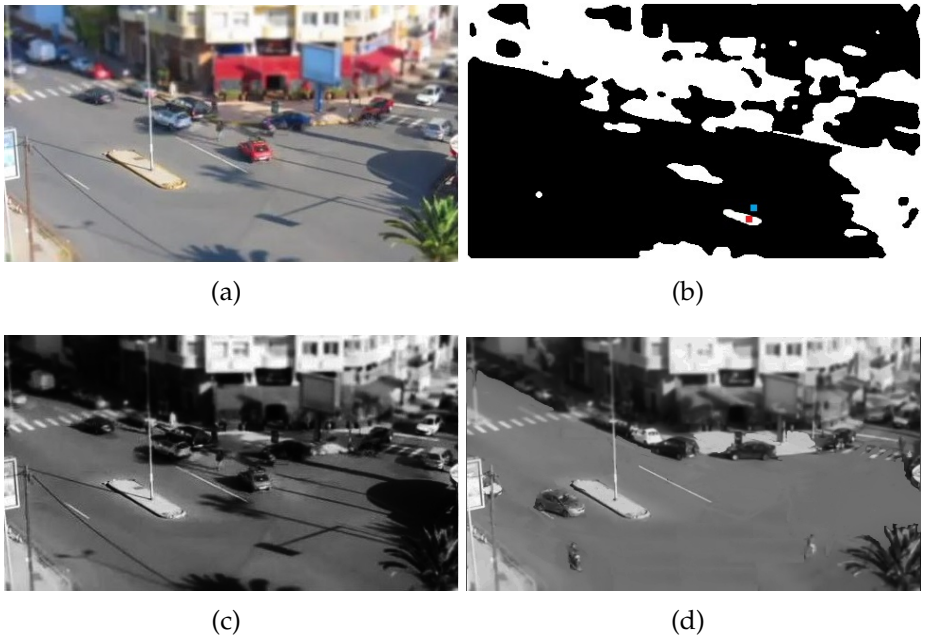
The following example is from a video taken from a stationary camera. We have used one frame per 2.5 seconds and in total 3000 frames were used to detect the reference shadow. In this case, we correctly detected the shadow located at the centre of the image as reference shadow and used that to generate the grayscale invariant image. In Figure 4.3, the process to generate the invariant image is displayed.

As displayed in Figure 4.3, we create the gs samples for uniformly distributed $C \in [0, 1]$, and for each sample, we compare the average of few blocks of pixels inside the reference shadow and compared them with multiple samples outside of it.

At some point the clear edge between the reference shadow and non-shadow areas starts to disappear and the average pixel value of the shadow region becomes very close to outside the shadow. This is an important threshold and at this point, the correct value of C is identified and it can be used for all other frames of the video to create the shadowless invariant image.

The reason we took multiple samples from inside and outside shadow area is that some of the samples may come from different object surfaces. Hence by having multiple samples, the impact of material change at the shadow edge will be removed.

Another example is shown in Figure 4.4 where the invariant illumination image is successfully produced but monitoring the edge of reference shadow. In this case, the reference shadow is at the bottom left corner of the screen.

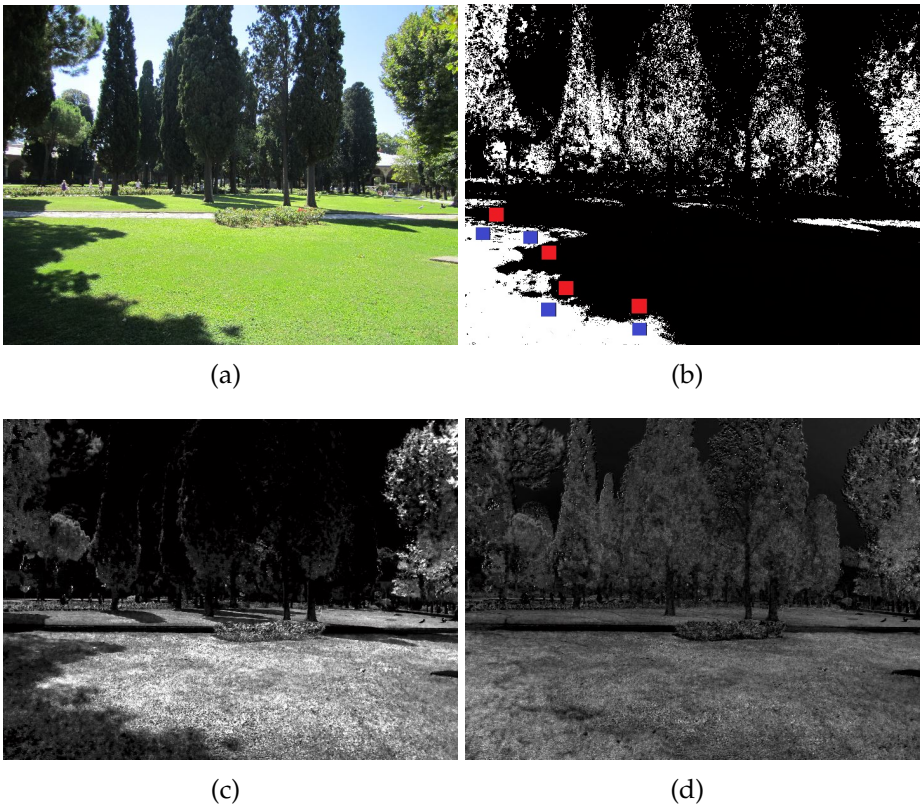


(a) (b)

(c) (d)

FIGURE 4.3: After processing 3000 frames, the shadow at the centre of the image is identified as a reference shadow. (a) RGB image (b) Comparing the values of a pixel block across the edge of reference shadow.(c) One of the gs samples where the reference shadow is not removed. In this image C=0.71 (d) One of the gs samples with the least difference between inside and outside reference shadow. In this sample C=0.42

Figure 4.3(d) and 4.4(d) show the outcome of the illumination invariance model. Although the shadows are completely removed, other illumination features in the scene are also removed and the final image seems to



(a) (b)

(c) (d)

FIGURE 4.4: After processing 1200 frames, the dark region at the corner of the image is detected as reference shadow. (a) RGB image (b) Comparing the values of a pixel block across the edge of reference shadow.(c) One of the gs samples where the reference shadow is not removed. In this image C=0.33 (d) One of the gs samples with the least difference between inside and outside reference shadow. In this sample C=0.61

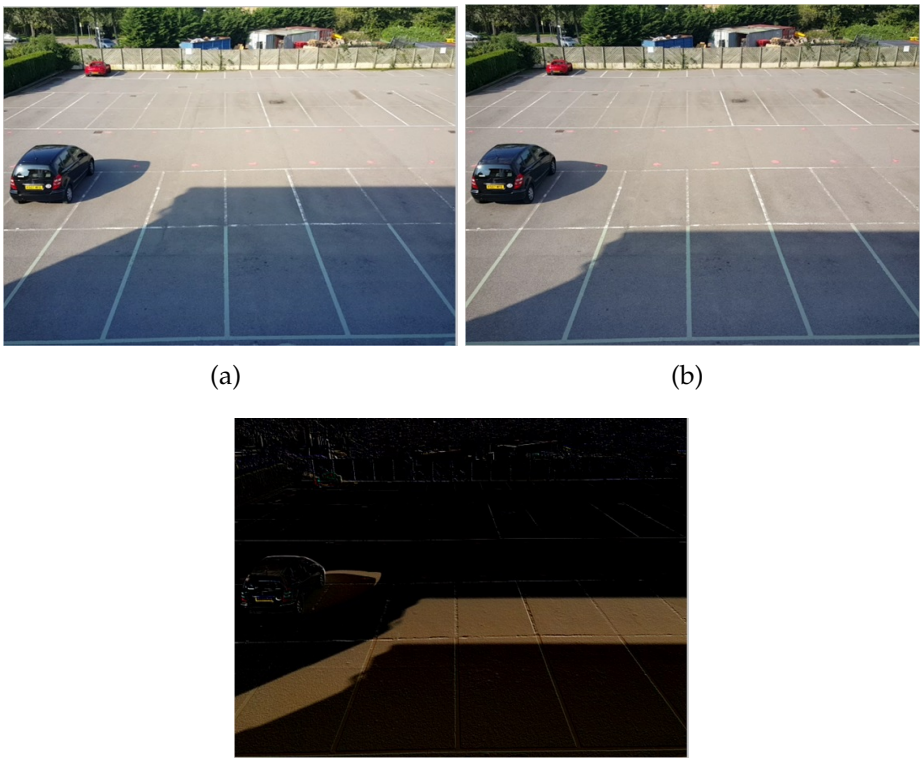
be unnatural grayscale compare to the original RGB image. To ensure this issue is resolved, the illumination invariance image is subtracted from the original RGB input in [51] to produce the edge of the shadow and then the RGB image is reconstructed. Although the final results look promising, this method will struggle to perform in busy environment with multiple shadows from moving and stationary objects and all of the examples and experiments presented in studies such as [51] are based on removing one single prominent shadow from a stationary object.

So to summarise, the applications where the grayscale image is sufficient, this method has great performance, and shadowless results can be produced very efficiently. However, for the systems which require all of the pixel data including colours, this system may not be suitable. In particular in a busy environment with lots of moving objects, producing a coloured shadowless image from the invariant image will be very difficult.

4.2.1 Analysing histogram changes during shadow removal process

Now that we have one shadow detected, we can analyse the histogram changes at the reference shadow edge. In previous sections, it was illustrated how important the shadow centroid is and here we utilise shadow edge to analyse the impact of illumination variance. Edges in an image usually have very important role in various image and video processing applications such as tracking and shape recognition because it is expected to have material changes at the edge. However, in the case of shadows, the edges exists because of illumination change and not material change. Hence these edges are not very useful for object and shape recognition systems. Here we explain how these edges can be used to provide important information about the lighting condition and chromaticity of the scene.

The following example is based on 45 minutes recording from a car park. The first, last and temporal difference of these two frames are shown in Figure 4.5



(a) (b)

(c)

FIGURE 4.5: (a) First frame of the video (b) last frame of the video (c) The temporal difference between these two frames.

Frame 4.5(c) contains very important information. The non-black areas show how the shadows have moved in 45 minutes, and these regions have been in shadow in the first frame, but they were outside shadow in the last frame. This means we have chromaticity and colour information of the same area when they were in shadow as well as outside shadow. Figure 4.6 displays the histogram of this region when it was in shadow and outside shadow.

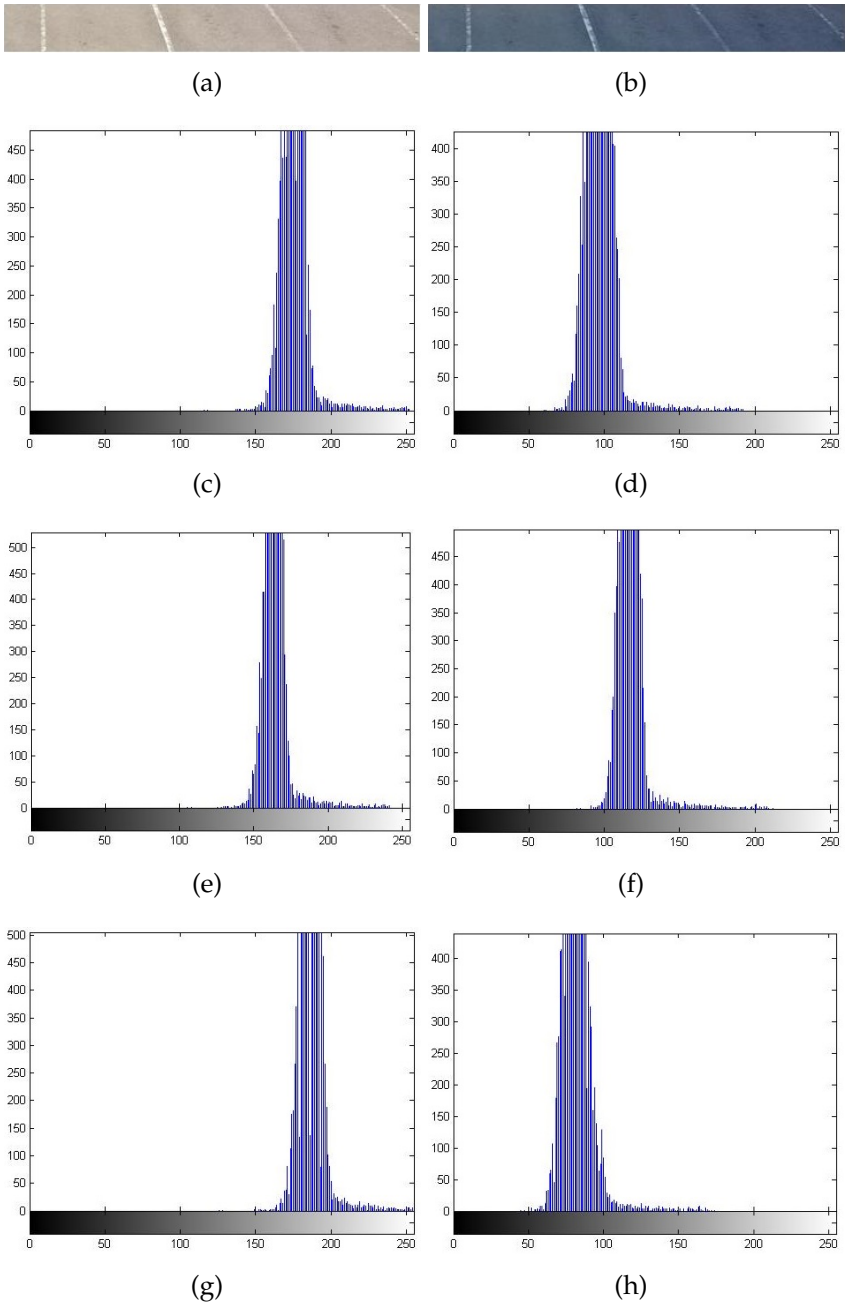


FIGURE 4.6: (a) and (b) show the same region when it inside and outside shadow. The histogram of the R, G and B components are displayed. When shadow covers the region, the histogram shifts to the left whilst the overall distribution of the histogram does not change significantly. Note the shift in blue channel is more than red and green channels.

Experiments show that shadows do not impact the distribution of the the histograms, they only shift the N-bins to the left. If we can find out

how the histogram bins have been shifted and calculate the ratio, we can apply the ratio to other shadows and find out how their RGB histogram would look like if shadow did not cover them. In summary, first we analyse the edge of reference shadow over time and calculate how the histogram of a block pixels have changed. We then apply the same ratio to other shadow regions. This assumption is correct for the scenarios where there is a single source of illumination ; for example normal outdoor condition where the sun light is dominant.

To have a robust comparison model, we can treat each histogram as a set of observations and then calculate the distance of each set using well-known methods such as Euclidean, Minkowski or Correlation distance. Given an *mx*-by-*n* data matrix X, which is treated as mx(1-by-n) row vectors x_1,x_2,...,x_{mx} and *my*-by-*n* data matrix Y, which is treated as textitmy (1-by-n) row vectors y_1,y_2,...,x_{my} the distance between the vector x_s and y_t are defined as follows:

Minkowski metric:

$$d_{st} = \sqrt[p]{\sum_{j=1}^{n} |x_{sj} - y_{tj}|^p} \tag{4.6}$$

for the special case of p = 2, the Minkowski metric gives the Euclidean distance, and for the special case of p=∞, the Minkowski metric gives the Chebychev distance. To compare the histogram set in shadow and non-shadow regions, we have used the Euclidean distance which is defined as:

$$d_{st}^2 = (x_s - y_t)(x_s - y_t)' \tag{4.7}$$

Calculating the Euclidean distance of R, G and B histograms of pixel block from reference shadow, provides a reference ratio ρ_r, ρ_g and ρ_b which can be used to shift and match the histogram of the reaming shadow of nearby pixels. Here we assume there is no material change at the shadow edge. This means we can build a database of pixel blocks and use that to detect and remove shadows of the moving foreground objects. In other words, let us assume we have two nearby pixel blocks A and B. Block A has been in reference shadow edge so we know how it looks like when it was inside and outside shadow; however, shadow has always covered block B during the video processing.

Now if we calculate the Euclidean distance of the histogram of block A for each R,G,B channel when it was inside and outside of shadow and call these value ρ_r, ρ_g and ρ_b. Then we can shift the histogram of block B by distance ρ to identify the unknown histogram of block B if it was outside shadow.

The difference between this method and illumination invariant method that was previously mentioned is here we actually detect shadows based on their colour and texture and we must do this for each shadow. Conversely, the illumination invariant methods remove all of the illumination information of the image at the same time. In the histogram based method, we can automatically build a dataset which contains the colour information of the blocks when they are inside and outside shadow. This is illustrated in Figure 4.7

Another important observation is in sunny days, ρ_b is slightly bigger than ρ_r and ρ_g. This is because the main source of illumination in shadow region is blue sky. Hence, shadows are slighter bluer than non-shadow regions. This is Illustrated in Figure 4.6 (g) and 4.6 (h).

An important point to highlight here is both shadows, and foreground objects move. However, the shadows move significantly slower than usual foreground objects such as people or car. This can be a very useful clue to detect shadows of stationary and static objects. For example comparing two frames when they are seconds apart, would highlight foreground objects but shadow edges do not change significantly in the space of a few seconds. However, the temporal difference of two frames where the gap is around 5-15 minutes would be sufficient to detect the shadow edges. This obviously depends on the following points:

- Time of the day where the video is taken.
- The position of the camera and its distance from the shadows.

The time of the day is important because the length of shadows changes slower at midday compared with morning (after sunrise) or late in the afternoon before sunset. Also the position and location of the camera is a key factor when it comes to detection of shadow movement. If the camera is focused and zoomed directly at shadow edge, it can detect the changes in the shadow length in minutes. Based on our experiments and in normal CCTV camera configuration where the position of the camera is high, comparing two frames where they are 10 minutes apart would be enough

to see shadow edges move clearly.

4.3 Experiments

The following experiments are based on videos taken from a stationary camera. They show how the position of the centre of the reference shadow changes as the time passes. The goal of these experiments is to demonstrate the potential to detect at least one reference shadow in most real-life scenarios. Once the reference shadow is detected, the grayscale invariant image can be computed for the whole video.

It should be noted that changes to the length (and hence the centre) of the shadow will be less at midday compare to morning and afternoon. For example, we managed to identify the reference shadow by processing around 15 minutes of video. This time was reduced to 8 minutes when the video was taken in the same location at 5 pm. In all the following examples, at least one reference shadow was identified, and then it was used to generate the greyscale invariant image automatically. In all scenarios, the shadowless greyscale image was produced successfully and the final results were satisfactory. It should be noted that we identified the reference shadow by monitoring how the centre of the shadow is changing over time.

Another interesting finding is when the video is recorded after sunrise or before sunset, the trajectory of the centre of the shadow shows non-linear patterns. However, during the rest of the day as the length of the

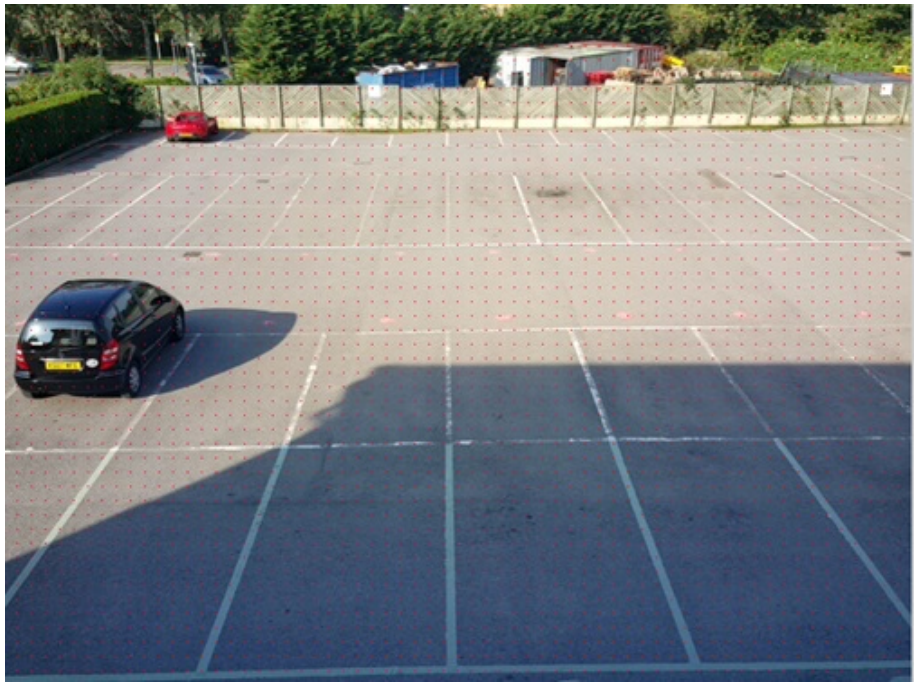
shadow changes linearly, the position of the centre shows linear pattern in short samples.

In both scenarios, the trajectory of shadow centre will be different to other dark areas of the image and hence the shadow can be distinguished from dark objects. Having said that, the quality of shadowless invariant image is of a lesser quality than the previous scenario. These experiments are presented in the following three subsections. The experiment conditions are outlined in Table 2. In the following examples, the reference shadow is marked by a white circle.

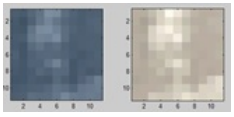
4.3.1 Experiment design

The way the experiments are structured is to test our shadow detection hypothesis and find out if reference shadows can be used to automatically detect shadows. Here we use reference shadows to detect shadows in different outdoor scenarios. First we start with simpler examples without any moving objects and then in other examples, it is illustrated that reference shadows can be used to detect shadows in very complex environments with significant number of moving objects.

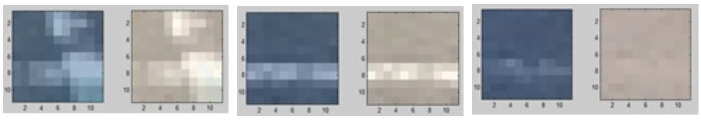
Additionally, it is illustrated that reference shadows can be used to successfully detect shadows close to sunrise and sunset. This is a significant improvement to the state of art and existing methods struggle to perform in these challenging and noisy circumstances.



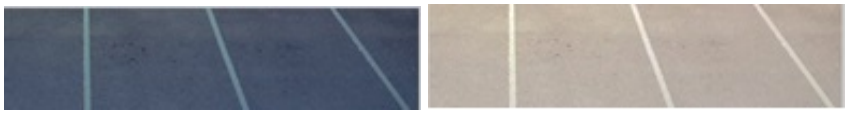
(a)



(b)



(c) (d) (e)



(f) (g)

FIGURE 4.7: (a) Image is divided into 10 x 10 pixel blocks (b) This block is at reference shadow edge and its distance ρ in histograms is calculated. (c),(d) and (e) are the result of shifting their histogram by ρ and the shadows are removed in each block. (f) Shadow region at the bottom corner of the image (g) shadow of the whole region is removed by shifting the R,G and B histogram of each block.

TABLE 4.1: Experiment Conditions

Figure	No. Frames	Video Length (mins)	Recording Time	Ref' Shadow Position	Trajectory of Ref' Shadow
(4.8)	300	45	Morning	Shadow of a chair	Linear
(4.9)	800	250	Morning-Afternoon	Shadow of a stick	Linear
(4.10)	300	60	Noon	Shadow at the corner of screen	Linear
(4.11)	250	30	Morning	Shadow of the cone	Linear
(4.12)	1700	20	Morning	Shadow of a lamppost	Linear
(4.13)	600	90	Morning	Shadow of a lamppost	Linear
(4.14)	70	100	Afternoon	Shadow of a building	Linear
(4.15)	3000	20	Afternoon	Shadow of a road sign	Linear
(4.16)	700	50	Early Morning	The dominant Shadow	Non-Linear
(4.17)	250	70	Late Afternoon	Shadow of a stick and clock	Non-Linear

4.3.2 Simple Videos without Moving Objects

Our method works great in the scenarios when there are no other objects in the video. In these scenarios, the reference shadow can quickly be identified, and the quality of the shadow-less image is also very good. The first two images show the first and the last frame of video. The third image shows the greyscale invariant image and the last image displays the trajectory of center of the reference shadow overtime. These videos are recorded during the day where shadow centre changes linearly. In this scenario, all the frames are registered for reference shadow identification as illustrated in Figures 4.8,4.9 and 4.10

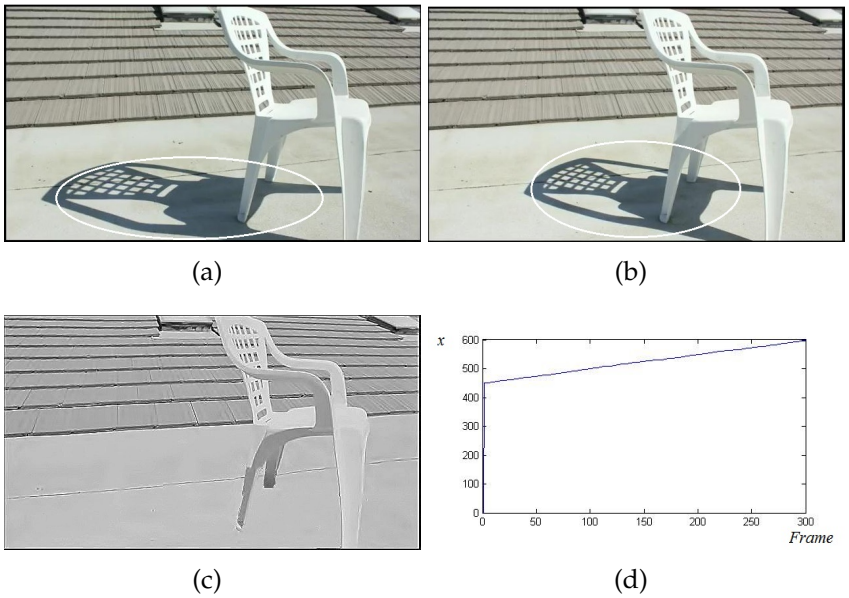


FIGURE 4.8: In this scenario there is a single object with clean background. All of the frames are registered and the centre of the shadow of the chair has a linear pattern in the 45 minutes of recording time. There is only one shadow in the video which is identified as reference shadow.

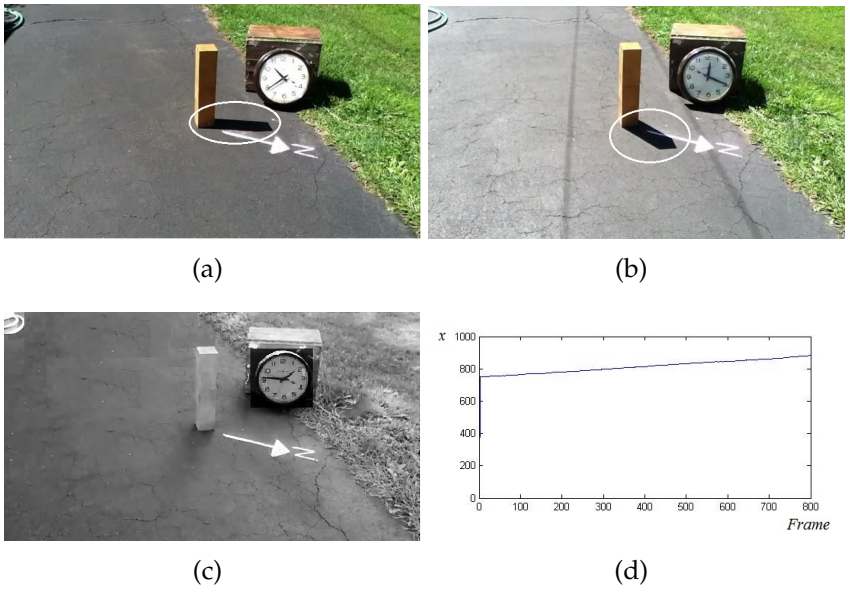
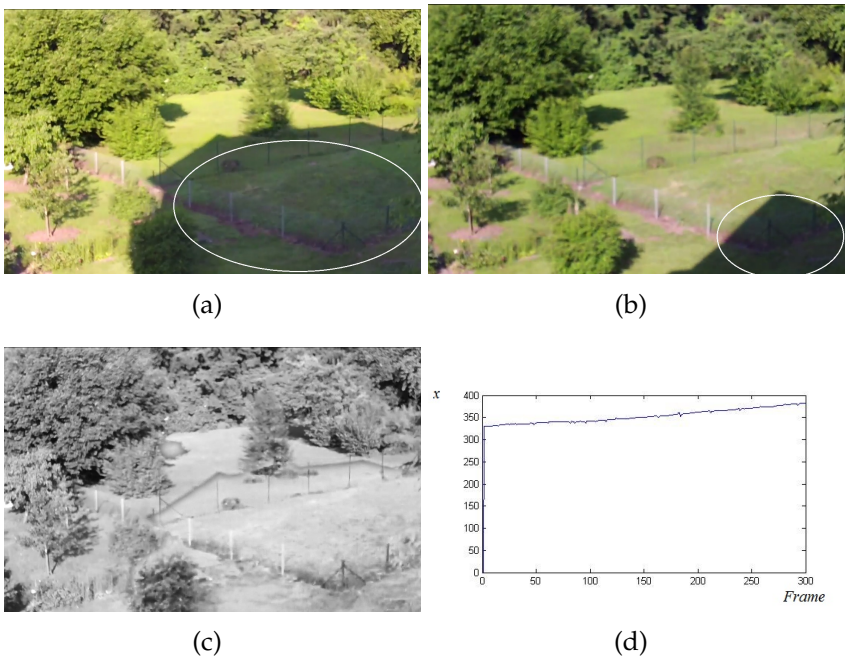


FIGURE 4.9: In this experiment there are two objects and the background is not homogenous. There are no moving objects and all of the frames are registered. In this example we used the shadow of the stick as the reference shadow.

4.3.3 Videos with moving objects

In this scenario, our method works exceptionally well and shadows from moving and stationary and static objects are detected. These cases are far more complex than examples given in other studies such as [119],[87] and [51]. The registration process is very robust and it successfully detects the reference shadow in a very noisy environment. In these real-life and noisy videos, the performance of our method is very impressive and any typical CCTV footage from the outdoor camera can be processed successfully and the shadows will be detected and then removed. In most of the following examples, multiple reference shadows were detected, but only was used to produce shadowless images.

In this scenario, when a moving object goes over the shadow, the coordination of the centre is not registered as displayed in Figures 13-17.



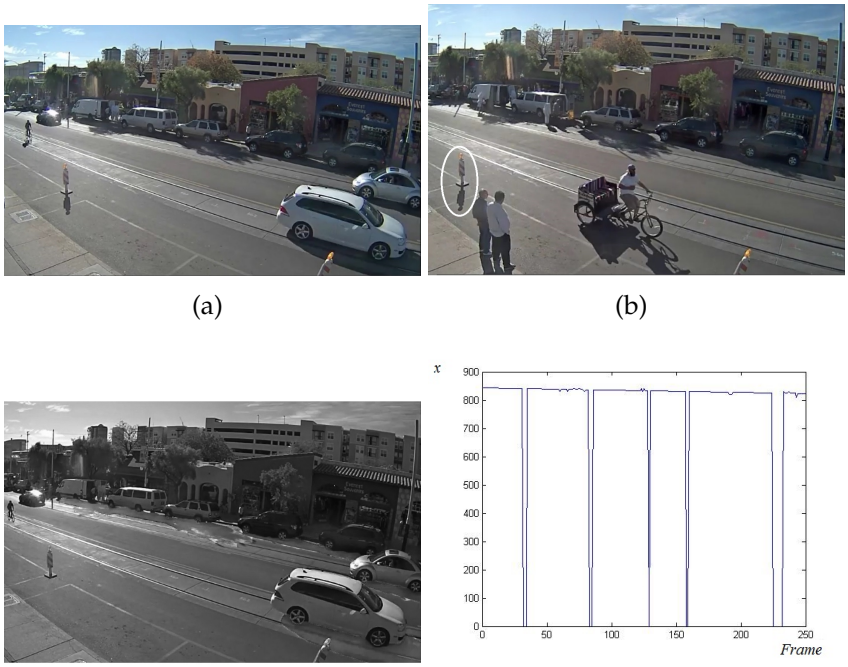
(a) (b)

(c) (d)

FIGURE 4.10: In this experiment the shadow is dominant and covers large part of scene. The background is not homogenous and the shape of the shadow gradually changes over the time. There are various objects in the scene such as trees. and the background is noisy.

4.3.4 Videos taken close to sunrise and sunset

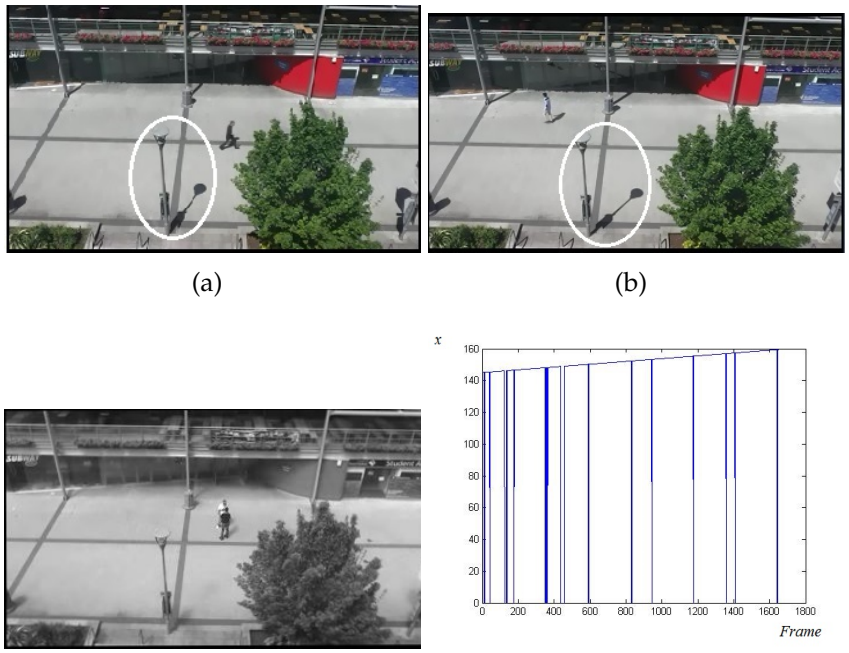
In this case the trajectory of the centre of the shadows does not have a linear pattern. This is the expected result as explained in Figure 2. In this scenario, the centre of shadow shows a non-linear pattern but still the reference shadow can be detected. Our experiments demonstrate the quality of the automatically detected invariant image is not as great as in the previous scenarios. However, to our knowledge, this is the sole method that can detect shadows even during poor lighting conditions when the illumination alters dramatically.



(a) (b)

(c) (d)

FIGURE 4.11: The reference shadow is the shadow of the traffic cone on the left side of the images. During the 30 minutes of recording time, the shadow of the cone has moved slightly to the left. In this example other moving objects go over the shadow few times hence some of the frames are not registered. In this scenario, there are various stationary and moving objects and the shadow can successfully be distinguished from other dark objects (such as dark cars).



(a) (b)

(c) (d)

FIGURE 4.12: The reference shadow is the shadow of the traffic cone on the left side of the images. During the 30 minutes of recording time, the shadow of the cone has moved slightly to the left. In this example other moving objects go over the shadow few times hence some of the frames are not registered. In this scenario, there are various stationary and moving objects and the shadow can successfully be distinguished from other dark objects (such as dark cars).

4.3.5 Comparison with other methods

To demonstrate how our system performs, we applied our method to the image set provided by Weiss [119] and Matsushita [88]. Regarding the quality of the shadowless images, there is no noticeable difference and both methods produce very similar results as displayed in Figure 21 and Figure 22. However, there is a significant difference between the two methods regarding the time required to produce the invariant image. For example, Weiss used 35 frames from sunrise to sunset but in our experience a video footage for around 10 minutes will be sufficient to detect shadows. The same results were observed when comparing our method

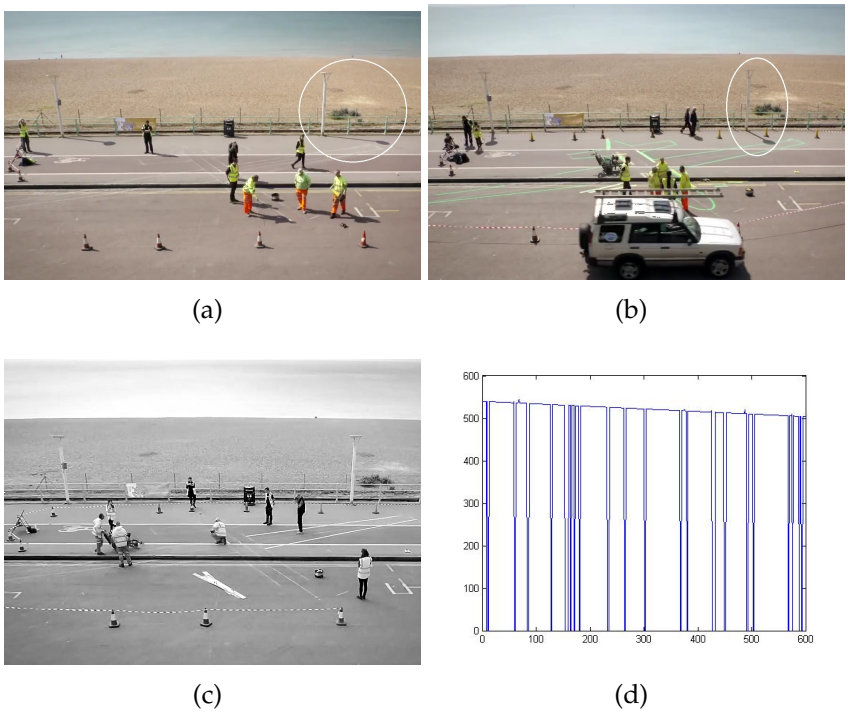


FIGURE 4.13: In this video the shadow of the cone was identified as the reference shadow with the highest confidence measure.

and Matsushita's [88]. Hence our method which is introduced in this research is a very strong solution for real-time application as shadows can be detected in few minutes.

4.3.6 Experiments summary

It was illustrated in this section that shadows can be detected automatically in different scenarios using reference shadows. Unlike most of exiting methods such as [3] ,[4] and [6], reference shadows perform exceptionally well in normal outdoor scenarios with lots of moving objects. Many of existing solutions can only detect shadows in perfect conditions and they are very sensitive to noise such [26] [57] and [12] but this limitation does not exist in reference shadow system. Additionally, reference

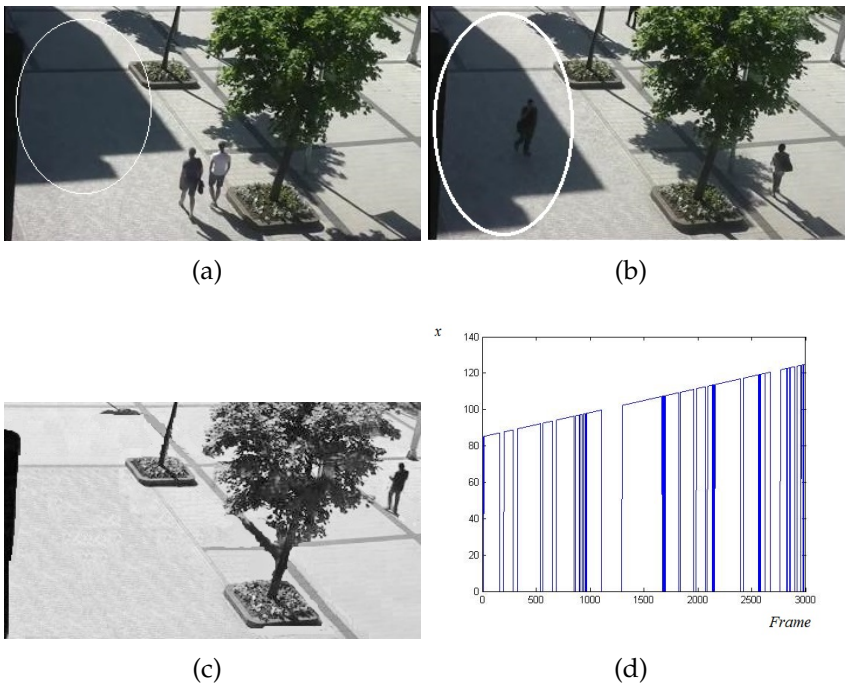


FIGURE 4.14: The reference shadow is dominant in this example and it is on the left side of the scene. When the pedestrians walk over the shadow, the size of the shadow changes and the frames are not registered. This is a typical outdoor CCTV footage and our method can successfully detect and remove shadow in such a noisy environment.

shadows have been used to successfully detect shadows near sunrise and sunset. Therefore to summarise and based on experiments illustrated in this this chapter, reference shadows can be used to detect shadows in the following scenarios:

1. Noisy environments with many moving objects.

2. Close to sunrise and sunset.

3. During the day when shadows disappear briefly because of the presence of cloud.

   In the next chapter 5, we explain how reference shadows can be used to detect some objects and events in the scene.

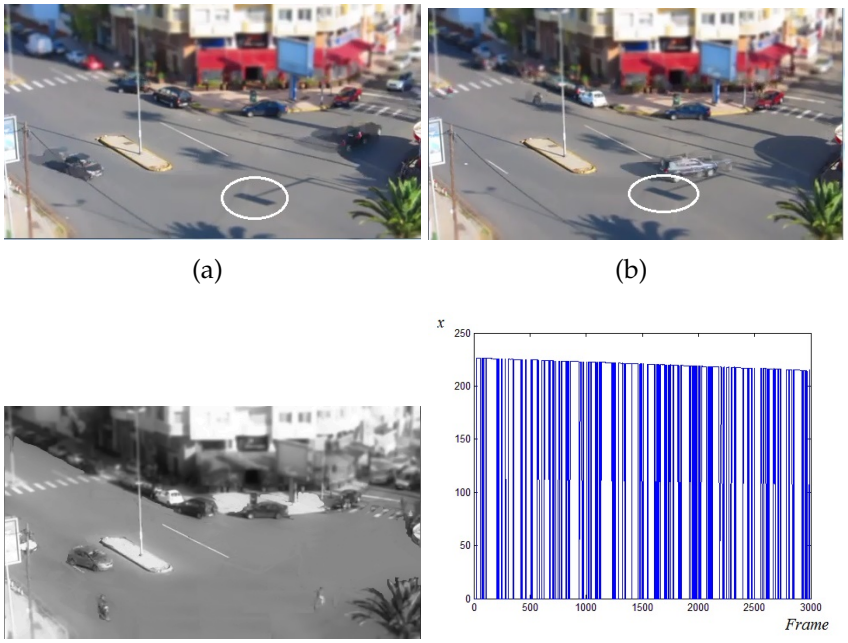
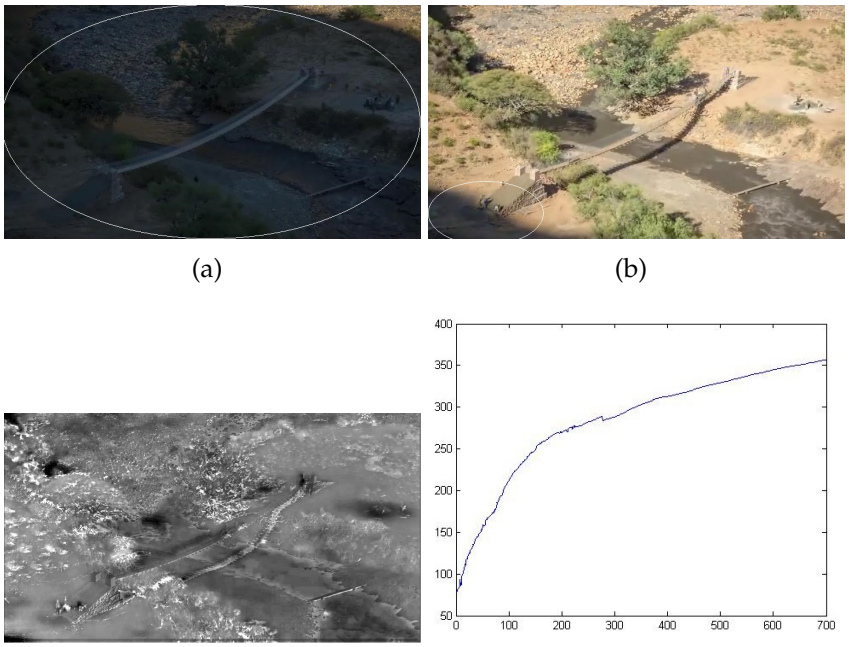


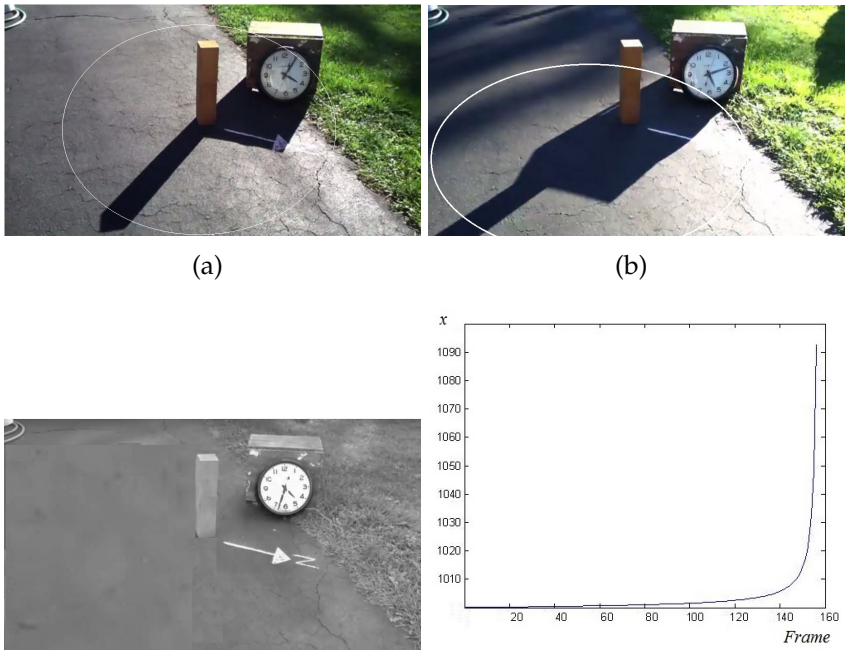
FIGURE 4.15: This is a CCTV footage from a very busy junction. There are many dark objects as well as numerous moving objects. In this video the reference shadow was identified as the shadow in the middle of the road (possibly from a road sign or a lamppost).



(a) (b)

(c) (d)

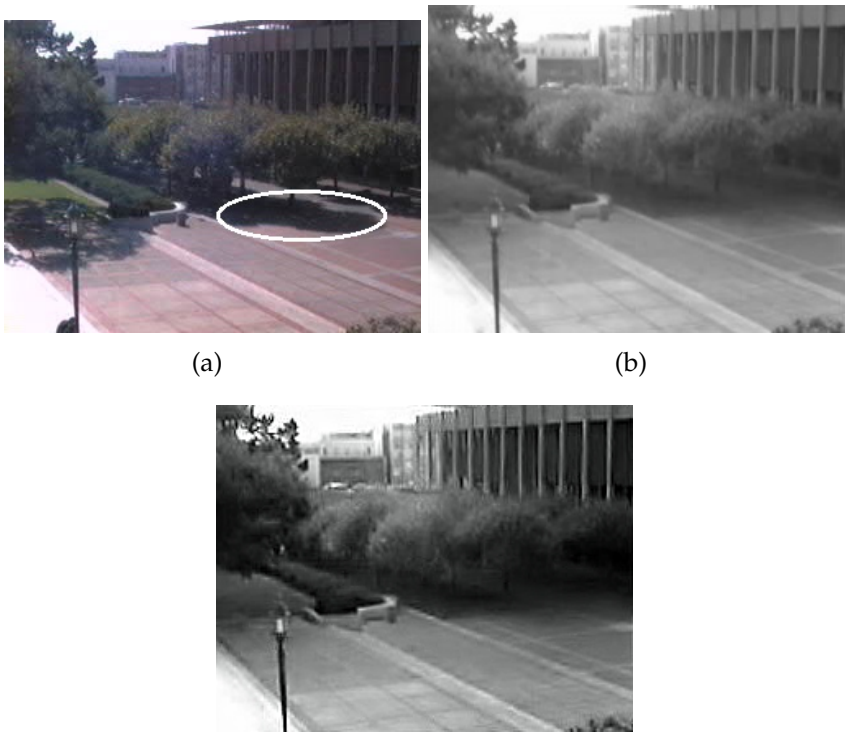
FIGURE 4.16: This video is recorded just after sunrise. The shadow on the valley changes rapidly over a short period of time so the centre of the shadow shows non-linear pattern. In this example the scene illumination is rapidly increased in less than one hour.



(a) (b)

(c) (d)

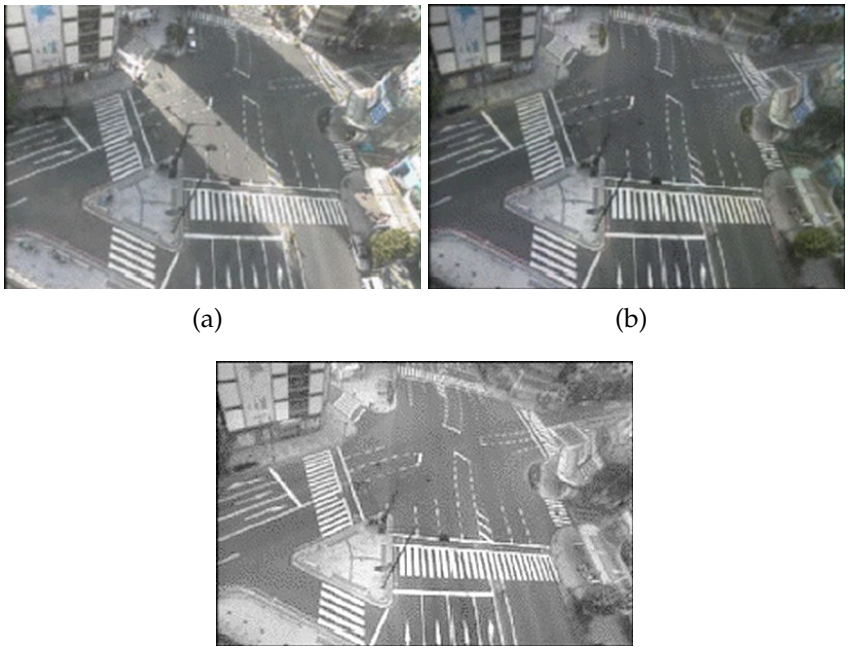
FIGURE 4.17: This example is just before sunset. In this scenario, the trajectory of the centre of the shadow has non-linear pattern over short period of time and illumination is rapidly reduced in one hour. In this experiment, two shadows are connected during the sampling process. Our method correctly detects the shadow in poor lighting conditions..



(a) (b)

(c)

FIGURE 4.18: (a) is an original imput image. (b) is shadowless image based on Weiss and (c) is based on reference shadow. The frame samples used in [119] are from morning until sunset. However video footage for around 15 minutes will be sufficient to detect reference shadow using our method. The final grayscale results are very similar but the advantage of our method is it will detect shadows a lot faster. The shadow highlighted in (a) has enough movement to be detected as per reference shadow method.



(a) (b)

(c)

FIGURE 4.19: (a) is an original image Matsushita [88]. (b) is shadow-less image based on Matsushita and (c) is based on reference shadow. They used mutiple in [88] are from different time of the day. However a short video footage from similar scenario was used to detect reference shadow using our method. To generate the invariant image, we used the shadow of the building in the middle of the image as reference shadow. Similar to previous example, the advantage of our method is it detect shadows a lot quicker and can be used when there are moving objects. Matsushita's method produces colour shadow-less image but assumes there are no moving objects in the scene.

Chapter 5

Texture and event detection based on reference shadows

In previous chapters, a novel method to automatically detect shadows was introduced and the advantages of reference shadows was illustrated in noisy environments with many moving objects. In this chapter was explain how reference shadows can be used to obtain more information about the scene.

By utilising reference shadows, it is possible to detect some of the features and objects in the video and discover some events. In this chapter we explain how it is possible to identify two very important features:

1. How to detect shiny materials the in the video automatically.

2. How to detect the presence of cloud in the video.

This will improve the awareness from the surrounding environment and can be used for texture detection.

5.1 Detecting non-Lambertian materials

By adopting a Lambertian model [61] of image formation so that if light source with a spectral power distribution (SPD) denoted *E(λ, x , y)* is incident on a surface which surface reflectance function is denoted *S(λ, x , y)* then the response of the camera sensors can be formulated as:

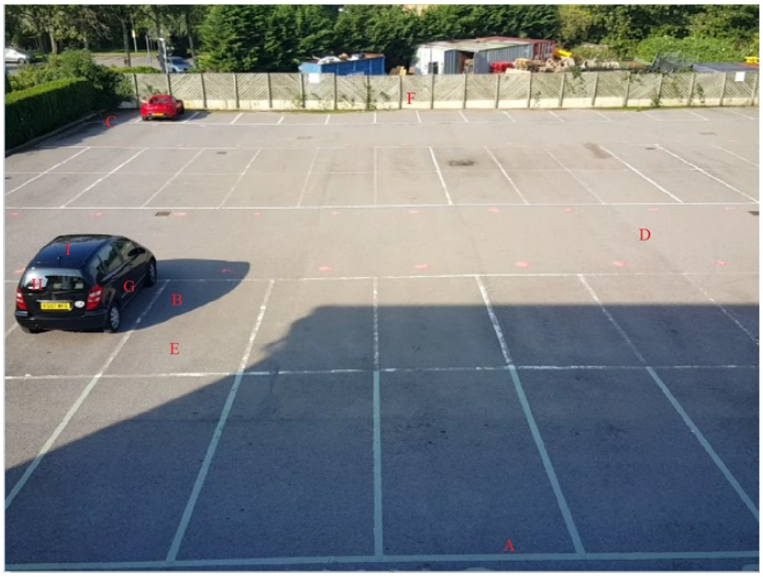
$$\rho_k(x,y) = \sigma(x,y) = \int E(\lambda,x,y)S(\lambda,x,y)Q_k(\lambda)d\lambda \tag{5.1}$$

where $Q_k(\lambda)$ is the spectral sensitivity of the k camera sensor, k = 1, 2, 3, and $\sigma(x,y)$ is a constant factor which represents the Lambertian shading term at a given pixel. In section 4.2, it was explained that assuming the surface has Lambertian characteristics, plotting Log (R/G) and Log (B/G) would form a line which will be orthogonal to lighting source. We previously showed how it is possible to generate grayscale illumination invariant model from reference shadows but here we present a novel idea to discover non-Lambertian materials such as glass, water, metal or any shiny material in the video under natural lighting condition. The approach taken here is producing a plot from Log (R/G) and Log (B/G) of different regions and various pixels in the video to see if they fit in the Lambertian model and then check if they produce a line. As explained in chapter 4, the result of this projection should be a single scalar which is coded as a greyscale value.

Here, we change the approach and claim that those objects that their chromaticity do not fit into the linear model of $gs = c_1R'(R) - c_2R'(B)$

should not be matt. This formula is generated from equation 5.1 Therefore projecting $R'(R) = log(R/G)$ and $R'(B) = Log(B/G)$ would be able to display if the surface chromaticity is fitting into Lambertian model or not. This is explained in the following figures:

(a) (b)



(c)

FIGURE 5.1: (a) First frame of the video (b) last frame of the video (c) 9 different pixels from three categories are selected for analysis. A, B, C are in shadow, D, E, F from none shadow, G, H , I are from the shiny parts of the car.

In this experiment nine pixels were analysed and the trajectory of their Log(R/G) and Log(B/G) were compared. The nine samples are from these three categories : 1) Shadow area 2) non-shadow 3) sample from

the black car. These samples are marked from A-I in Figure 5.1 (c). The duration of this video is one hour and the plot of Log (B/G) and Log (R/G) of these nine pixels over one hour, as depicted in Figure 5.2

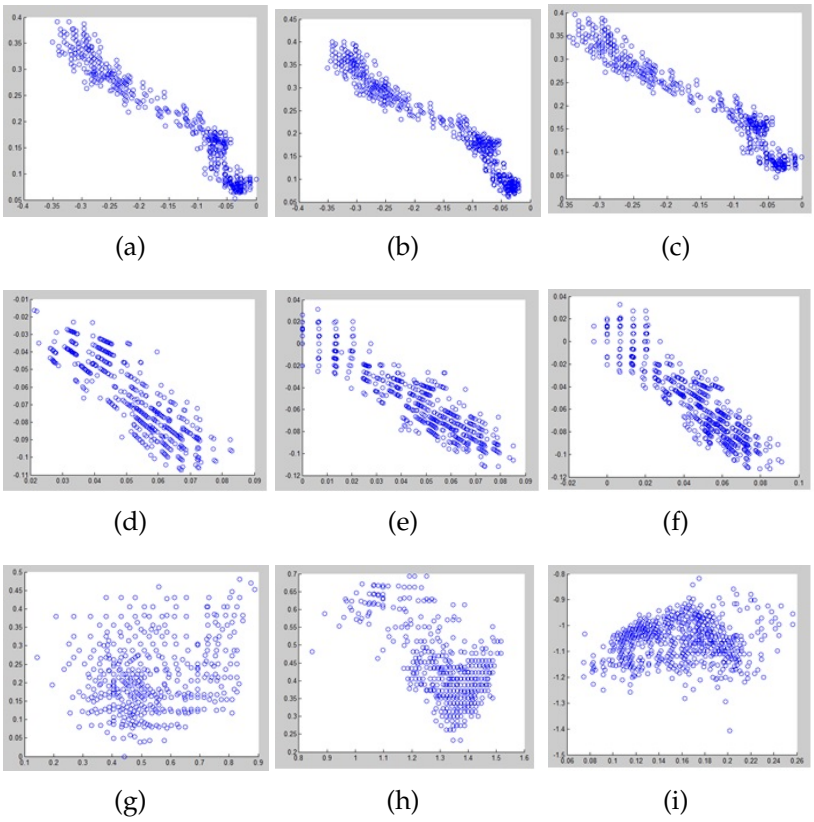
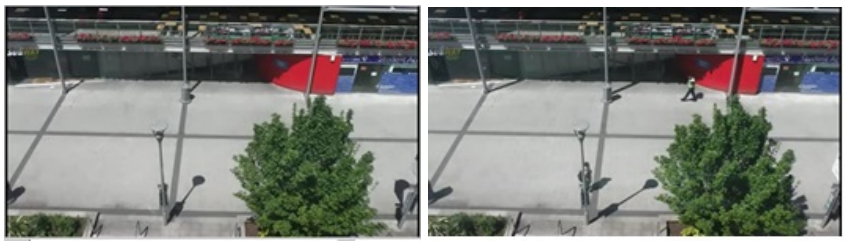


FIGURE 5.2: First row figures shows the trajectory of pixels (a) , (b) and (c) in Figure 5.1 over 1 hour of video which belongs to shadow region. The second row displays the pattern of pixels (d), (e) and (f) which are from the non-shadowed region of the image. 3rd row shows the trajectory of pixels (g), (h) and (I) which are of shiny metal and glass.

The important observation here is the pixels which belong to Lambertian surface model i.e. matt materials show linear trajectory over recording time. This assumption is true regardless of the sampling position whether the pixel belongs to shadow or non-shadow region. However,

the pixel which belongs to non-Lambertian surface model i.e. shiny materials such as water, glass or metal do not fit into this model and their chromaticity dose not reveal a linear pattern when the source of illumination changes.

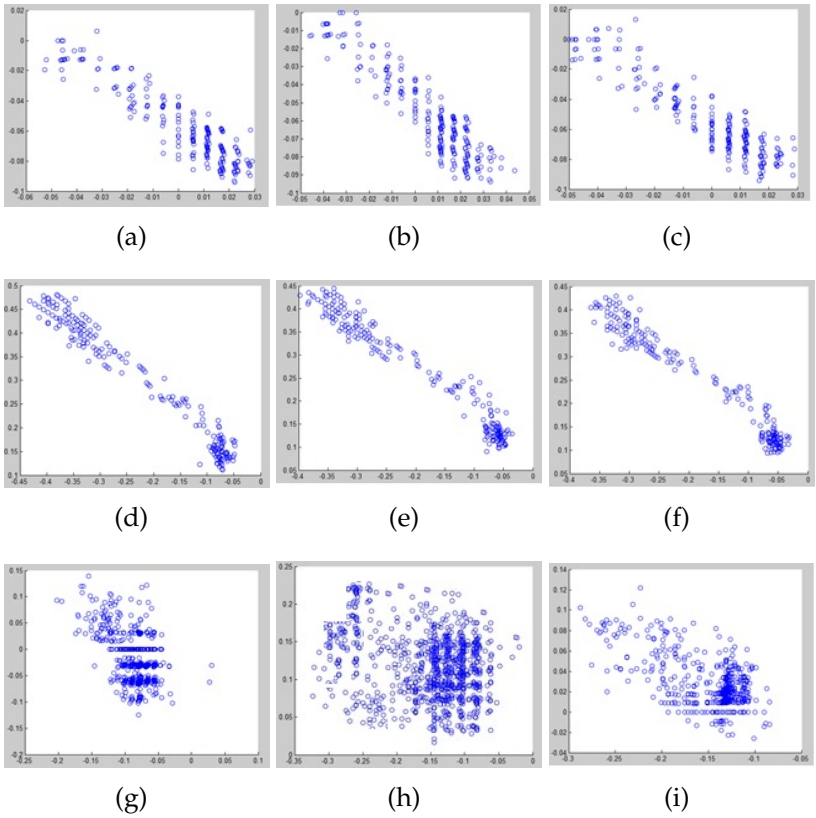
This behaviour is evident in Figure 5.2. The first two rows show the chromaticity of matt objects in both shadow and non-shadowed part of the image and the 3rd shows the same analysis for shiny objects. In this experiment, the matt regions show linear chromaticity over time while the shiny materials do not. This important characteristic can be used to discover and distinguish shiny and matt textures in the video.



(a) (b)

FIGURE 5.3: The 1st and last frame of the video taken outside Brunel university library.

Another example is shown in Figure 5.4 where pixels from matt surfaces show linear chromaticity over time (both shadow and non-shadowed regions). However, on the other hand, pixels from glass windows and doors do not fit into non-Lambertian model do not display any linear trajectory over time. Hence by plotting $R'(R) = log(R/G)$ and $R'(B) = Log(B/G)$ we can successfully distinguish matt and shiny textures over time because matt materials show linear pattern whilst shiny materials



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

FIGURE 5.4: The 1st row figures shows the trajectory of pixels taken in shadow region. 2nd row displays the pattern of pixels from non-shadowed region of the image. The last row images show the pattern of the trajectory of pixels from glass doors and windows.

don't fit into any linear model.

Another interesting observation is seeing the difference between Figure 4.2 and Figures 5.3 and 5.4. The trajectories illustrated in Figures 5.2 and 5.3 are from real-life conditions where the illumination changes are not as controlled and regulated as lab lighting conditions, as depicted in Figure 4.2 (b). In normal outdoor conditions, the illumination changes are consistence with the exception when there are clouds. For example, after sunrise, the scene becomes brighter until noon but if there are clouds, the image may become darker than previous frames. This why in normal

lighting condition, some patches can be seen in the trajectory of matt pixels over time, which are clear in rows 2 and 3 of Figure 5.2 and 5.3. In the next section, we explained how it possible to detect the presence of cloud in the image and what information can be obtained from this event.

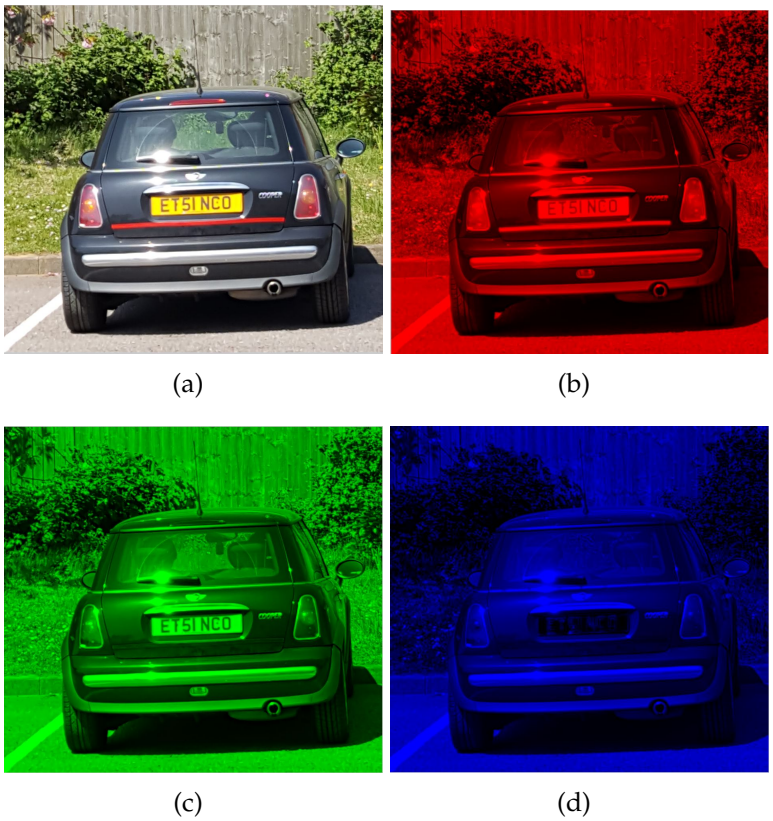
5.1.1 Distinguishing reflectance from white material

The difficulty in processing image and video in RGB domain is both white materials and specular reflection show as saturated patches in the scene so the R, G and B value of both areas are close to 255. This is illustrated in the following Figure 5.5. What this shows is just simply searching for saturated white regions will not be a good method to identify the reflections because the other white parts of the image will be incorrectly identified as reflectance.

As it can be seen in Figure 5.5 both the reflection and the white line on the road have saturated the R, G and B channels. This means selecting pixels based on their colour will not be sufficient to identify the glare and reflectance. Here we propose a new method which can be used to distinguish white materiel from reflectance successfully and we also illustrated how the glare and reflectance from each frame can be removed.

5.1.2 Logarithm of Chromaticity

In the previous section we explained that plotting *Log (R/G)* and *Log (B/G)* of Lambertian material would form a line which will be orthogonal to



(a) (b)

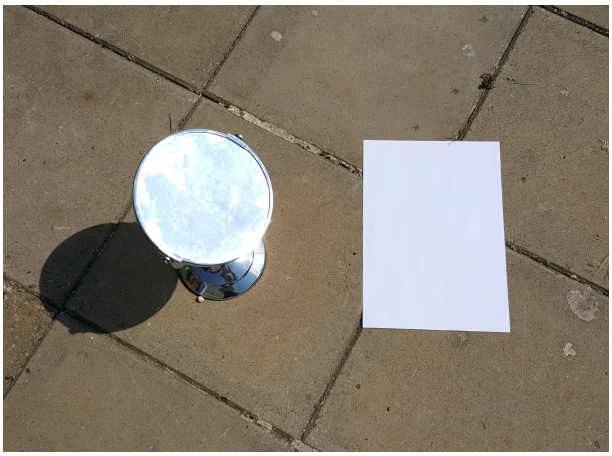
(c) (d)

FIGURE 5.5: Reflection in image saturates all three RGB channels so it can be detected incorrectly as a white material. The saturated pixels are shown in (e) which includes both white materials and reflection pixels.

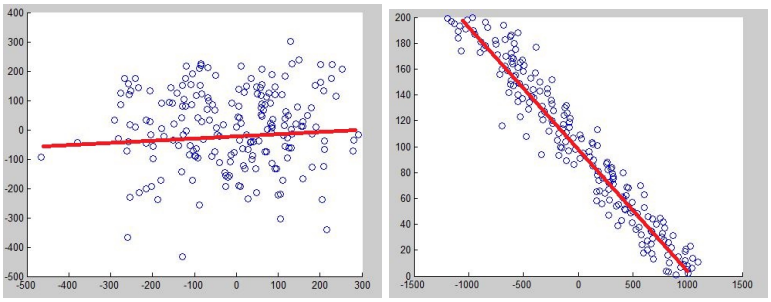
lighting source. We also explained that materials and textures that do not fit into the linear model of $gs = c_1R'(R) - c_2R'(B)$ are not matt. Therefore projecting $R'(R) = log(R/G)$ and $R'(B) = Log(B/G)$ would display whether or not the surface chromaticity is fitting into Lambertian model. Reflectance and glares are examples that do not fit within the following model:

$$\rho_k(x,y) = \sigma(x,y) = \int E(\lambda,x,y)S(\lambda,x,y)Q_k(\lambda)d\lambda \tag{5.2}$$

This methodology is illustrated in Figure 5.6. In this example, both the white paper and the mirror have saturated all three channels and the



(a)



(b) (c)

FIGURE 5.6: (a) The RGB value of both objects are close to 255 and both surfaces are white (b) logarithm of chromaticity of mirror over a period of time (C)logarithm of chromaticity of white paper which forms a line. In this example the σ of white paper is 89.15 but it is 221.89 for the mirror

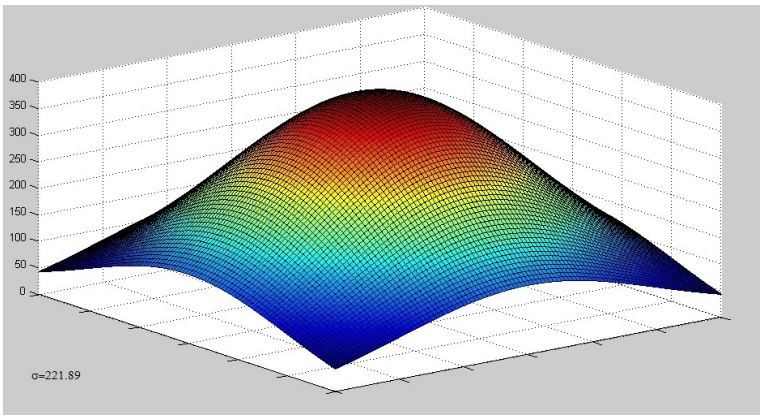
value of their RGB is close to 255. However during the 15 minutes recording, their logarithm of chromaticity shows a different trajectory and this can be used to distinguish the white paper (matt surface) from the mirror (shiny surface). To deploy this automatically, we attempt to form a line from $log(R/G)$ and $Log(B/G)$ values using a smoothing filter from the trajectory of both matt and shiny surfaces.

Therefore, to distinguish white matt objects from reflectance, we draw

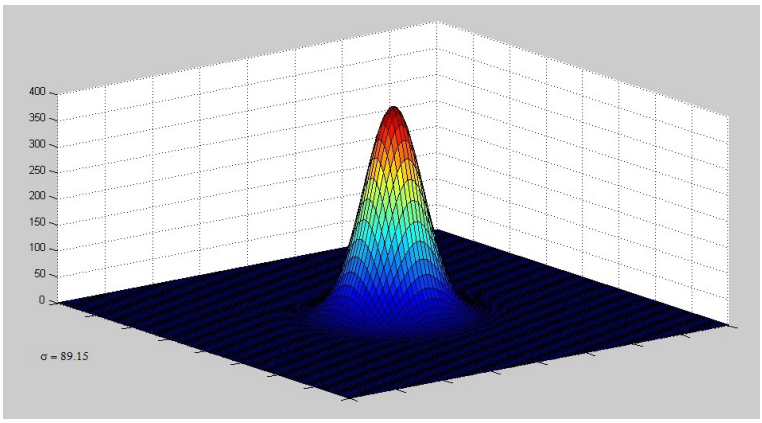
a line from scattered data and calculate the normal distribution of the logarithm of chromaticity of both surfaces. The σ of the shiny surface is significantly higher than matt material. There are many research materials and methods in the field of statistic about interpolating scattered data. For the simplicity, we calculated the average of the of the first few coordination to produce the first point of the line and used the same method to generate the last coordination of the line. This simple method works perfectly in the less noisy environment but in environments with moving objects and less reliable data, more complex methods will be needed to draw a line. This approach is explained in algorithm 2.

Algorithm 2 How to diffrentiate between white materials and reflectance

1: Plot the gradient of $R'(R) = log(R/G)$ and $R'(B) = Log(B/G)$ from saturated regions using the available frames.
2: Calculate the average of the first 10 percent available $R'(R) = log(R/G)$ and $R'(B) = Log(B/G)$ to generate the first coordination of the line.
3: Calculate the average of the last 10 percent of the available $R'(R) = log(R/G)$ and $R'(B) = Log(B/G)$ to generate the last coordination of the line.
4: Produce the normal distribution of all the logarithms of chromaticity.
5: The σ of the distribution of the shiny material will be higher than matt surface and this indication can be used to distinguish these two types of material.



(a)



(b)

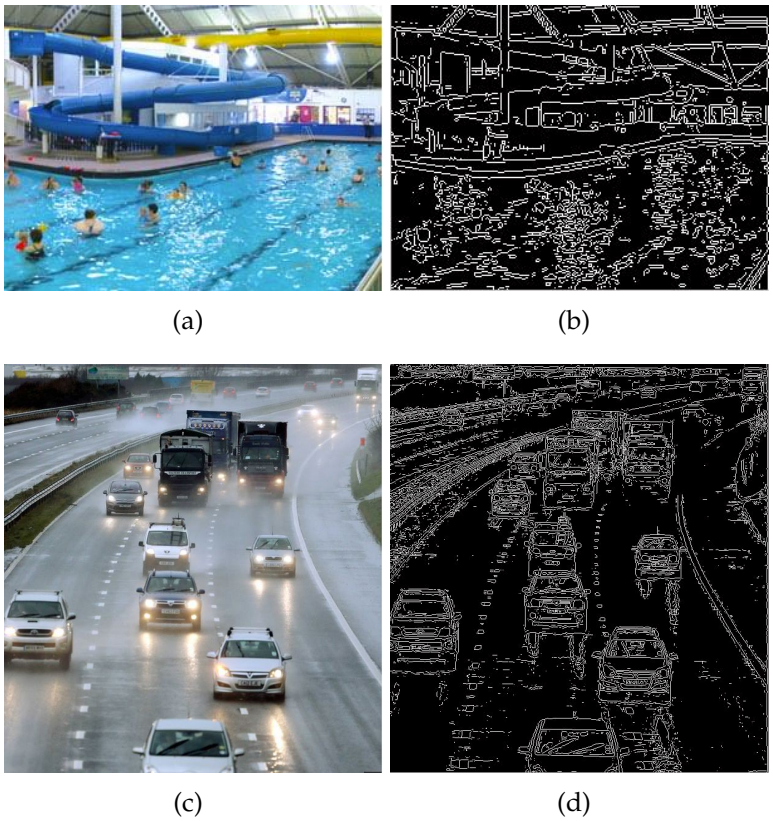
FIGURE 5.7: (a) The ratio of change of logarithm of chromaticity of a shiny surface (b) at The ratio of change of logarithm of chromaticity of a matt material.

5.2 Reflectance and glare analysis

The previous section presented a system to detect the materials with a shiny surface. This was used to identify water leakage, metal or glass automatically in the video. This section discusses and analyses the impact of expressive reflectance, which can have a negative impact on applications, such as tracking and video surveillance. In many real-time applications, particular outdoor environments such as roads and highways, highlights and glares caused by light specular reflections are challenges for object detection in some monitored areas as targets of interest are hidden and

covered by these lighting artefacts. Additionally, significant movements of the reflection region due to the dynamic background cause an increase in foreground error detections.

In Figure 5.8, some of the examples based on these scenarios are displayed. The reflection in these examples is generating additional unwanted edges which will have a negative impact on tracking applications and object detection systems.



(a) (b)

(c) (d)

FIGURE 5.8: (a) The reflection of light on the water is generating a lot of noise and unwanted edges. (b) The reflection of the light on the road surface is producing false moving component which would impact the performance of the car tracking system.

The problems of foreground noise reduction and hidden target detection have been researched extensively in the literature and to reduce the

noise caused by light reflection, polarising filter set in the front of the camera lens has been used. This system however, needed to detect incident polarisation angle of specular reflection as the optical axis of the filter [53]. One of the use cases of this application is video processing and tracking in an aquatic environment which is explained in [118]. In their method, they have used HMM models to identify people in swimming pool.

This is achieved by using a Bayesian approach using Hidden Markov model scheme is designed to enhance human detection and tracking performance for video surveillance system operating under noisy aquatic environment. Pixels in the current frame are classified as foreground or background by considering prior historical information of its surrounding frames and the likelihood probability between the current frame and the reference image.

5.2.1 Reducing glare and reflectance

Now that the regions and pixels which reflect the sun are successfully detected, we propose an algorithm which can be used to remove the reflectance from frames automatically without any user input. The use case of this system is for any typical video surveillance or tracking application where extensive reflectance and glare can negatively impact them. This is Illustrated in Figure 5.5, where the reflectance has produced additional edges. This will reduce the performance of object detection systems. Another application of this method is for image enhancement and image restoration systems where it is used to remove unwanted reflection from

video to produce better quality images.

To achieve this goal, the same methodology introduced in Chapter 4 will be implemented to differentiate between white objects and reflectance. Previously, we used reference shadows to produce illumination invariant image which resulted in having shadowless images but here we use the reflectance areas of the image as a reference to generate the same illumination invariance images. As discussed previously, the illumination invariance images lose all of the chromaticity information which includes both shadows and reflectance. This is summarised in the algorithm 3:

Algorithm 3 Generating greyscale Invariant Image for reflectance removal.

1: Identify the source of reflectance as described in Algorithm 2.
2: Select $C \in [0\ ,\ 1]$ to generate gs' samples.
3: **For all** gs' **do**
4: Analyse the chromaticity of reflectance region.
5: Find the gs' sample with the least chromaticity for the reflectance pixels.
6: **End For**
7: The gs' sample identified in step 5 is gray-scale invariant image and the corresponding parameter C is set.

The 2-D gray-scale invariant image which is produced by algorithm 3 loses of its illumination information. This includes shadows and reflectance. The key point to highlight here is we have proposed a system that can be used to differentiate between white matt materials and shiny surfaces.

This process is explained in following Figure 5.8. In this example the reflectance source is detected as described in algorithm 2 and then the value of chromaticity is compared with a white Lambertian surface. At

the point where Δ becomes maximum, the reflectance is reduced together with all the other illumination information in the image including the shadow of the car.

(a) (b)

FIGURE 5.9: (a) The white paint and the white reflectance on the car have similar RGB values.(b) at C=0.21 the delta between these two regions becomes maximum. At this point Illumination Invariant image is produced where shadows and reflectance are removed.

5.2.2 Experiments

In this section, some examples of how this system can be used to remove glare from images are presented. These are all from real-life video footage and the excessive reflectance is removed and illumination invariance image is produced which removes the illumination information from the image.

(a)

(b)

FIGURE 5.10: (a) The white paint and the white reflectance on the car have similar RGB values.(b) at C=0.21 the delta between these two regions becomes maximum. At this point illumination invariant image is produced where shadows and reflectance are removed.

5.3 Detecting the presence of clouds by analysing reference shadows

In the previous section it was explained how some of the features in the video could be detected and now we explain how some of the *events* can be discovered by analysing Reference shadows. Here we explain how it is possible to detect the presence of cloud by using the reference shadows and also the information that can be obtained from this event.

In the majority of image and video processing applications, such as tracking and object recognition functions, detecting the presence of cloud may not be a high priority in activity recognition; moreover,this event is not significant by itself. nonetheless, the presence of cloud has a major impact on the illumination condition of the camera viewpoint and more

(a) (b)

FIGURE 5.11: (a) Typical CCTV footage with modest reflectance on some of the cars (b) Producing 2-D illumination Invariance image would remove the shadows as well as the reflectance on the cars.

importantly would allow us to produce high quality shadowless images; especially when clouds completely block the sun and shadows disappear.

As discussed in chapter 2, most of the existing cloud detection methods are based from satellite images [5],[12],[26] and [2]. In this section, a new solution is provided which is based on standard CCTV video footage where the camera is positioned outside and it is stationary.

Furthermore, fast moving clouds alter the illumination conditions of the scene rapidly; thereby impacting tracking and feature recognition. To achieve this, first we introduce a ratio that can be used to identify shadows.

5.3.1 Blue channel proportion analysis

Pixels get darker when shadows cover them but in addition to this our experiments show that the shadow areas are bluer that non-shadowed regions. This is because the only source of illumination in shadow regions is sky, which is perceived as blue. When a pixel or a region is illuminated by a blue coloured light source, they tend to reflect blue light more and

(a) (b)

FIGURE 5.12: (a) An image with extensive glare and reflectance (b) The illumination invariance image is generated as explained by algorithm 3. The impact of illumination invariance is significantly reduced.

hence they look bluer. Conversely, the non-shadowed regions are illuminated by both sky and the sun where the illumination from the sun is far more dominant than the sky. In other words, the non-shadowed regions will not be as blue as shadows.

Consequently, the rate of change in R, G and B channels is not the same when shadows cover the the area. Although the intensity of the RGB channels reduces when a pixel is in the shadow region, the change is not proportional. So assuming P_{RGB} is the RGB values of a pixel in the non-shadow region and $P_{R'G'B'}$ is the similar value for the same pixel in shadow, we can claim that $P_{B'}/P_B > P_{R'}/P_R$ and $P_{B'}/P_B > P_{G'}/P_G$. The following examples show this ratio for shadow and non-shadowed regions are displayed in Table 5.1

TABLE 5.1: Analysing the blue ratio in an outdoor environment.

Figure	$P_{R'}/P_R$	$P_{G'}/P_G$	$P_{B'}/P_B$
5.9(a)	0.4844	0.5018	0.5876
5.9(b)	0.3178	0.4171	0.4491
5.9(c)	0.5217	0.5053	0.5829
5.9(d)	0.4019	0.3814	0.4682
5.9(e)	0.3596	0.3716	0.4151
5.9(f)	0.4163	0.4723	0.5876
5.9(g)	0.4844	0.5018	0.5876
5.9(h)	0.4087	0.3832	0.4329

As evidenced in these examples, the blue ratio is a good factor to narrow down the possible candidates of shadow regions and also can be used a reference to validate the presence of cloud which results in removal of the shadow. To use this, we check the blue channel of the shadow region when we suspect the clouds have appeared. What we should expect is the shadow region will become brighter but the $P_{B'}/P_B$ will not be increased as much as $P_{B'}/P_B$ and $P_{B'}/P_B$. This is because on the cloudy days, the sky will be no longer blue and the dominant source of illumination will be the grey sky.

Here we make an important assumption that when we refer to clouds, we assume the direct sun light is fully blocked and the shadows are completely removed. In other cases where there are thin clouds and the direct sun is partially blocked; this theory is not valid.

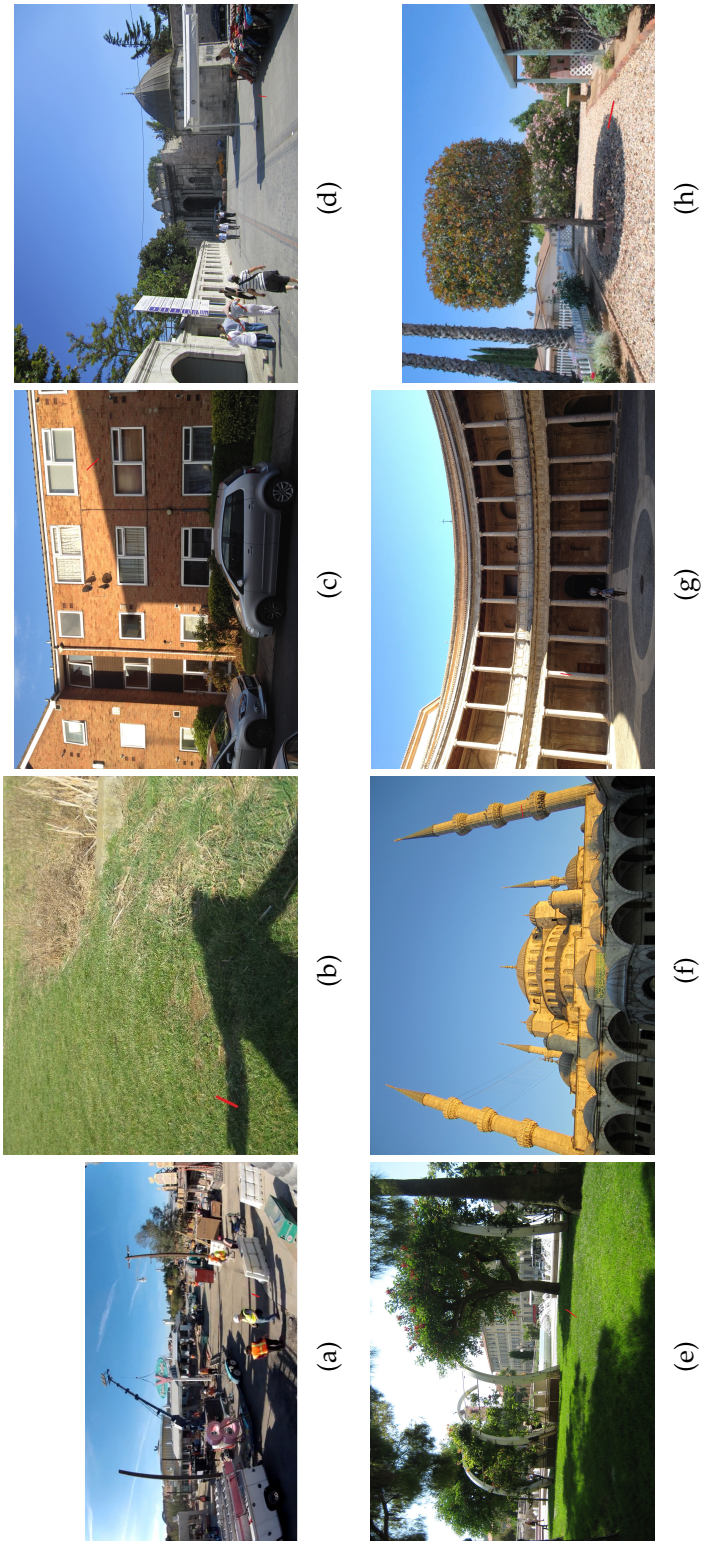


FIGURE 5.13: In this example 8 frames from real-life scenes are displayed. Pixel samples from shadow and non-shadowed regions are marked in red. The $P_{R'}/P_R, P_{G'}/P_G$ and $P_{B'}/P_B$ ratios are displayed in table 3.

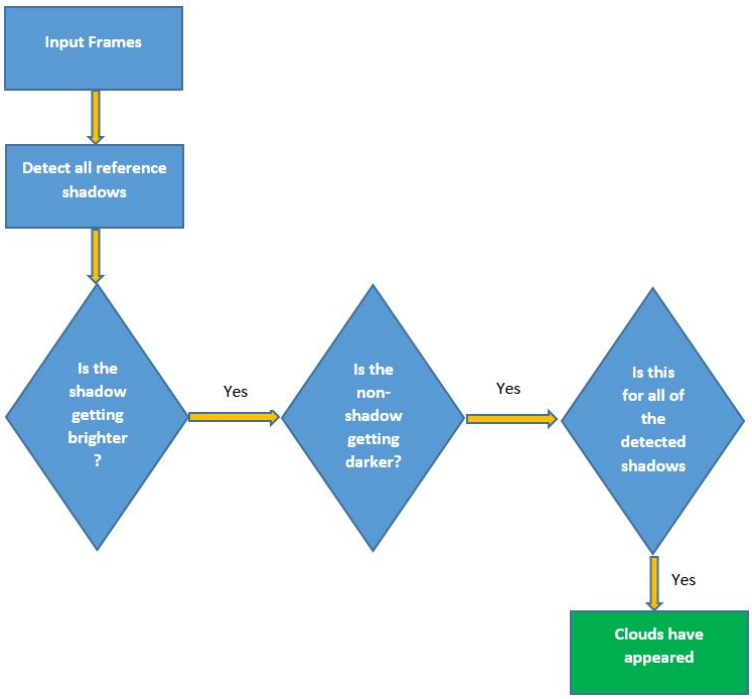
5.3.2 Cloud detection

Here we explain how we can detect the presence of cloud and also elaborate how this event can lead us to generate high quality shadowless images when there are no clouds in the scene. We focus on normal illumination condition where some part of the day there are no clouds and then the clouds gradually block the sun. This is implemented by meeting the following logical conditions to detect clouds:

1. When clouds appear, all the reference shadows start to disappear.
2. When clouds appear, the non-shadowed regions will become darker while the shadowed regions become brighter.
3. When clouds appear, the shadow regions will become less blue.

If the above three conditions are true, we have a very high chance that clouds are present. This is because in normal lighting conditional the sun illumination follows a specific pattern. From sunrise to noon, the illumination increases before decreasing gradually until sunset. Therefore, changes in this behaviour indicate an external factor has influenced the illumination source. In normal outdoor condition, the most important factor will be clouds. Here, we make some very important assumptions:

- The camera sensor is consistent during the day.
- Artificial light sources are negligible compared with illumination from the sun.

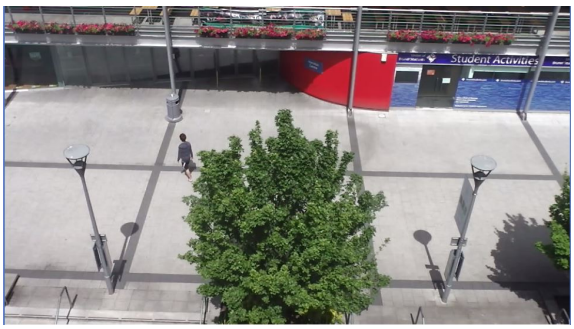


(a)

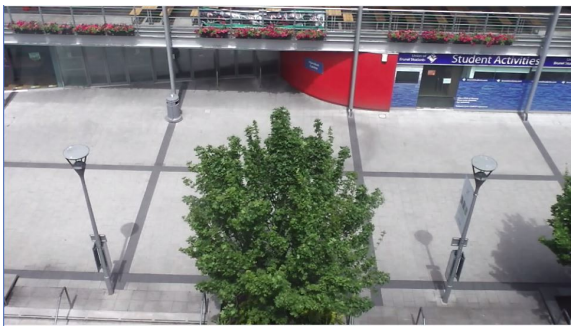
FIGURE 5.14: Detecting the appearance of the cloud. All of the detected shadows are removed simultaneously and the areas which were previously in shadow will become brighter. The pixels that used to be outside shadow will become slightly darker.

These two assumptions are very important and in normal outdoor CCTV conditions, these assumptions are correct. The following Figure 5.15 displays the impact that cloud can have on the illumination condition over a short period. This process is displayed in Figure 5.15. In this example, there are 25 frames per second and cloud appear in less than 30 seconds. The process of cloud detection is Illustrated in Figure 5.16. In this example, the shadow darkens gradually as the time passes. However after frame 3000, the process is reversed and the shadow region becomes righter. This is the first indication that clouds have started to appear which can be seen in Figure 5.15 (a). The second condition for cloud presence is the fact that the brightness difference between shadow and

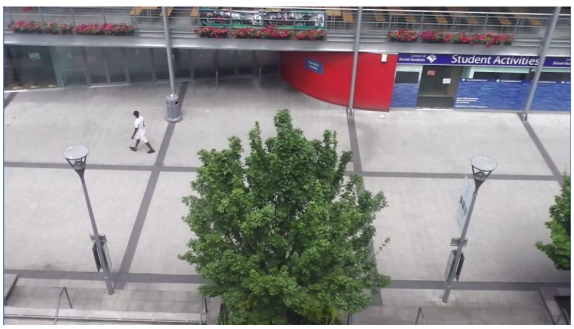
non-shadowed region suddenly start to reduce until it becomes zero. This is depicted in Figure 5.6 (b) where the edge of shadow and non-shadowed region has disappeared. In this experiment, the brightness difference of the lamppost shadow and the background is reduced to zero from around 120.



(a)

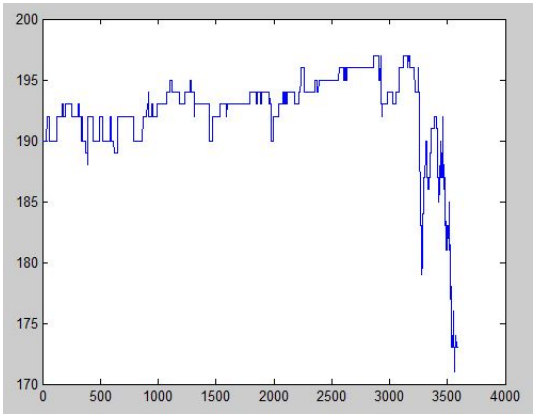


(b)

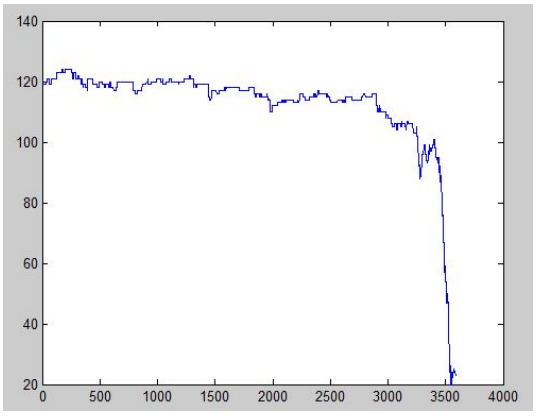


(c)

FIGURE 5.15: In this experiment, the clouds appear in less than 3 minutes. The impact is shown on reference shadows where they gradually disappear.



(a)



(b)

FIGURE 5.16: (a) shows the greyscale value of the reference shadow (in this example it is shadow of the lamp-post), the shadow is getting darker but suddenly it becomes brighter. (b) This figure shows the difference between the brightness value of reference shadow and the non-shadowed region at the edge of the reference shadow.

It should be highlighted that to have a robust system to detect cloud, we would need more than one reference shadow. This reduces the impact of foreground objects going over the shadows which may result in producing the wrong conclusion. In most of real-life outdoor scenarios, there will be multiple reference shadows because potentially many stationary objects can produce reference shadows. In these cases, cloud detection will be more accurate because the impact of the cloud will be the same on all the reference shadows. Hence if all the reference shadows start to fade

and disappear at the same time, we can conclude the presence of cloud with high confidence.
In the example that is displayed in Figure 5.15, there are two reference shadows from lamp-posts, three reference shadows from the bicycle racks at the button of the image. When the clouds appear, all of these shadows are removed simultaneously.

Now that we have successfully detected the presence of cloud, we can use this to discover and remove the shadows in other times where there are no clouds. This is explained in the next section.

5.3.3 Utilising cloud detection for shadow removal

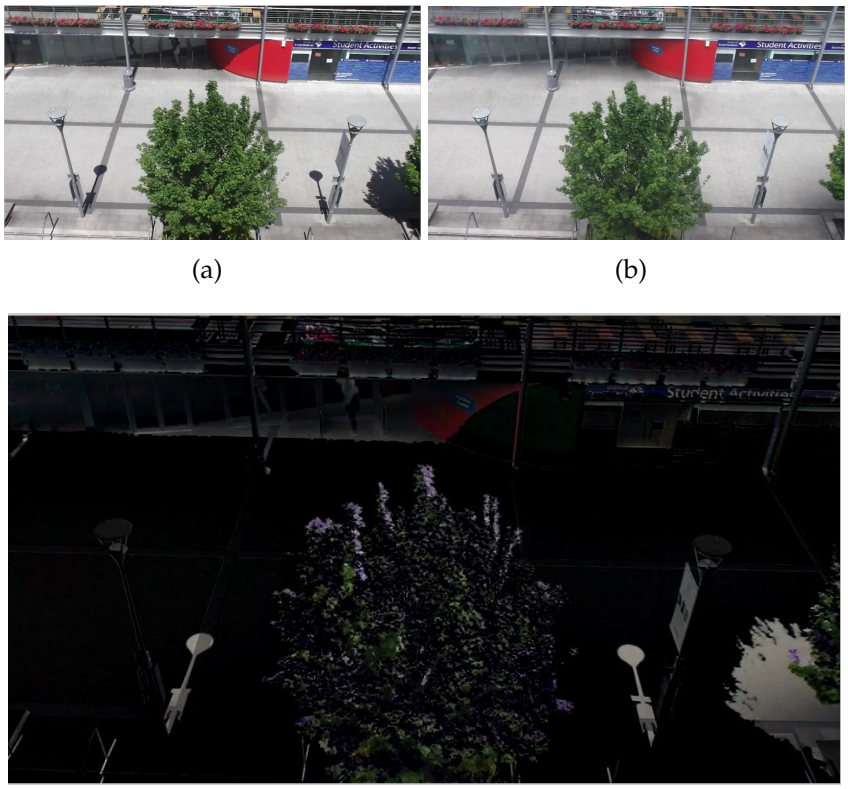
In previous section a novel method for cloud detection was presented. Here we explain how we can utilise this event to detect and remove all the shadows in the image and this would be the real benefit of cloud detection. As displayed in Figure 5.15 (c) when the clouds block the sun, all the shadows will disappear and the scene is illuminated by light from the sky instead of direct light from sun. This is a very important and useful attribute of clouds and since they block the direct light from the sun, the shadows are removed. So, when we know there are clouds, we can conclude there are no shadows; therefore, we can produce a shadowless reference image which can be used in other frames to detect shadows. To do this we have subtracted the cloud frame from non-cloud frames and the results are illustrated in Figure 5.17

In this example, two frames from a video sequence are displayed. One of the frames is from the time where there are clouds and the other frames shows the view when there are no clouds present. Figure 5.17 (c) displays the delta between these two frames and here two very important features are highlighted and everything else is filtered:

1. Moving objects.
2. Shadows.

Here, we have a procedure to segment shadows and moving objects from the rest of the image. Both two features are very important subjects

in many image and video processing applications such a shape and activity recognition and tracking. The importance of Figure 5.7(c) is the fact that shadows are shown and by using the shadowless reference image in Figure 5.17 (b), we can remove the shadows from image 5.17(a).



(a) (b)

(c)

FIGURE 5.17: (a) a frame illuminated with direct sun light (b) a frame where clouds block the direct sunlight and therefore shadows are removed. (c) Subtracting the two frames would highlight two important features: 1) moving objects 2) Shadows.

5.4 Experiments

In this section, we have presented the experiment results and showed how the source of reflection could be automatically detected by analysing various regions of the frame over a period. Additionally, we have provided experiment results that reveal how the clouds can be detected automatically by analysing the reference shadows.

The following figures display the experiment results of the cloud detection process explained in this chapter. The following experiments are from a video footages taken from stationary camera in the daytime. In each experiment, the location and position of the reference shadow are explained. Additionally, the brightness difference between regions inside and outside shadow is shown.

As it can be seen, when the could appear in the scene the region which was previously in shadow will become brighter and at the same time the area which used to be outside shadow will become darker. The illumination difference between shadow and non-shadow region reduces and these two regions will become almost identical when thick clouds completely block the sun.

5.4.1 Experiment design

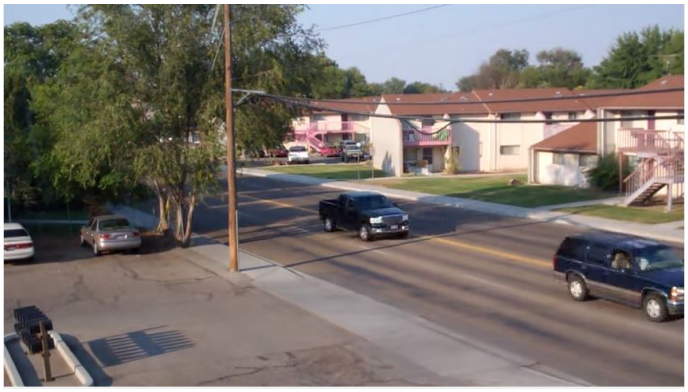
The experiments in this section are designed to illustrate that clouds can be detected by analysing the reference shadows. The hypothesis is clouds can be detected automatically using reference shadows and then high quality and natural shadowless images can be produced. The natural

shadowless image can be subtracted from other frames and then all other shadows in other frames of the video can be detected. This is shown in the following examples where all the frames are from standard CCTV footages. Unlike the existing literature such [11], [14] and [26] where clouds are detected from satellite images, here we propose a new novel solution which is based on cloud detection based on analysing the impact of clouds on shadow regions.

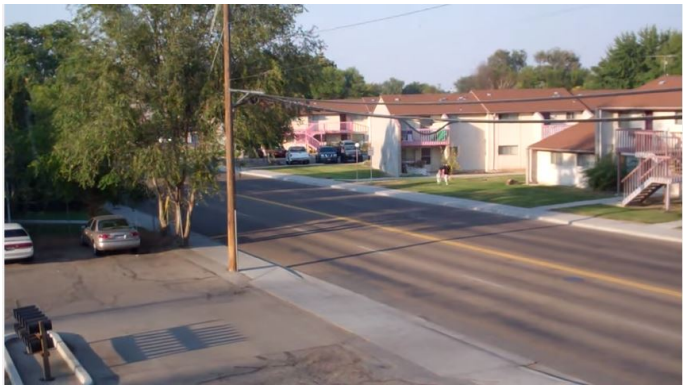
5.4.2 Conclusion

In this section a new method is presented to detect clouds. Also a novel solution is provided to identify none-Lambertian textures in the video automatically. Detecting shiny materials has many applications such as water and ice detection for road users and it can it can also be used to detect glass and metal in video.

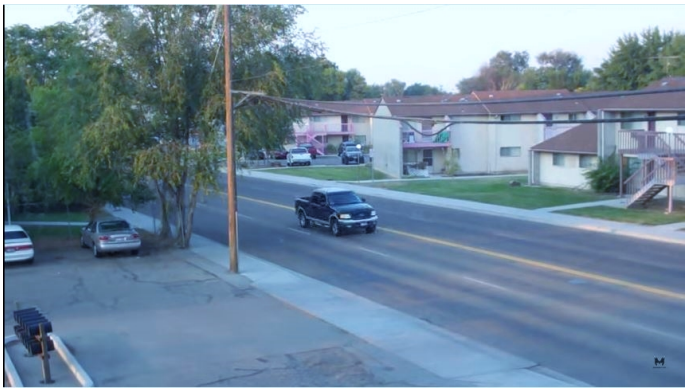
Cloud detection is very important in video processing due to the impact on illumination. When clouds appear, the sun will be blocked and therefore the shadows will be removed. This high quality shadow-less image can be used in other frames of the video to detect shadows and also moving objects.



(a)



(b)

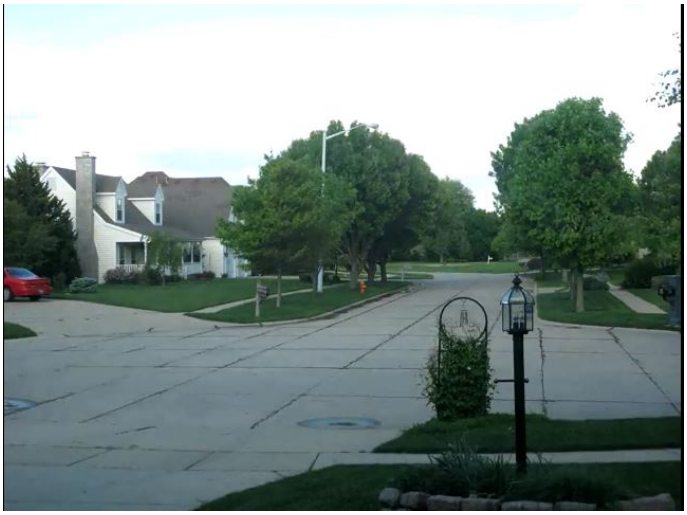


(c)

FIGURE 5.18: (a) a frame illuminated with direct sun light (b) a frame where clouds block the direct sunlight and therefore shadows are removed. (c) Subtracting the two frames would highlight two important features: 1) moving objects 2) Shadows.

(a)

(b)



(c)

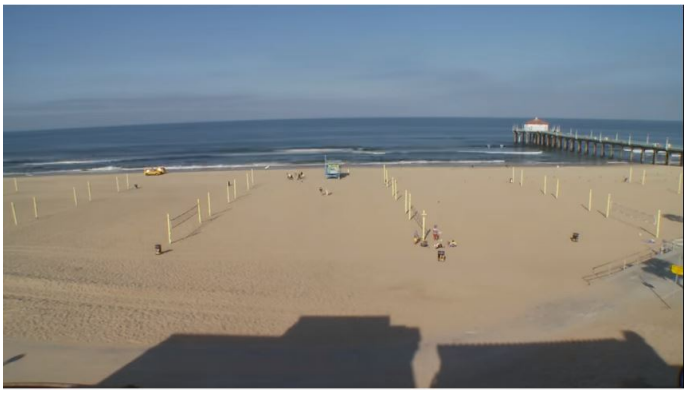
FIGURE 5.19: (a) a frame illuminated with direct sun light (b) a frame where clouds block the direct sunlight and therefore shadows are removed. (c) Subtracting the two frames would highlight two important features: 1) moving objects 2) Shadows.

(a)

(b)

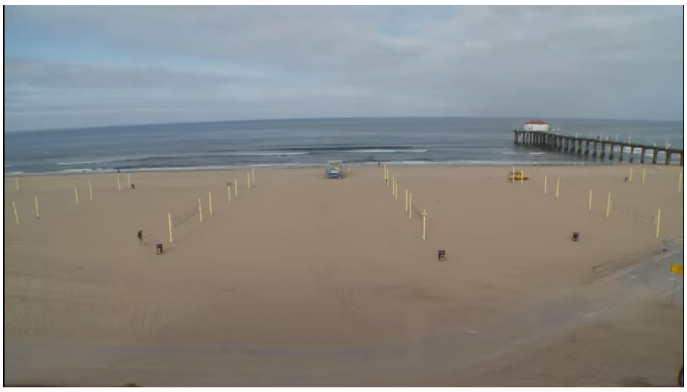
(c)

FIGURE 5.20: (a) a frame illuminated with direct sun light (b) a frame where clouds block the direct sunlight and therefore shadows are removed. (c) Subtracting the two frames would highlight two important features: 1) moving objects 2) Shadows.



(a)

(b)



(c)

FIGURE 5.21: (a) a frame illuminated with direct sun light (b) a frame where clouds block the direct sunlight and therefore shadows are removed. (c) Subtracting the two frames would highlight two important features: 1) moving objects 2) Shadows.

Chapter 6

Conclusion and further work

6.1 Summary

In this thesis, we have introduced a new method to automatically detect and remove shadows in videos taken from stationary cameras in an outdoor environment. Based on the observation that, owing to the relative movement of the sun, the length and position of a shadow change linearly over a relatively long period in an outdoor environment, we can conveniently distinguish a shadow from other dark regions in an input video. Then, we can identify the reference shadow as the one with the highest confidence of the above mentioned linear changes. This reference shadow is used to produce shadow-free invariant model[51], with which the shadow-free invariant images can be computed for all the frames in the input video.

6.2 Contributions

The main contributions of the work can be summarised as follows:

1. We have developed a novel method to identify the reference shadow. Without any prior knowledge, a shadow would appear the same as other dark regions in an image, e.g. an object with dark colours. However, if we observe the changing patterns of the position and length of a dark region, those for a shadow changes linearly while those for a dark object do not change. Thus the shadows can be distinguished from the dark objects. With simple confidence measured against this linearity and consistency of changes, we can choose one or a few reference shadows with the highest confidence measure. Experimental results have demonstrated this process of reference shadow identification is fairly accurate and reliable.

2. With the reference shadows, we can effectively produce shadowless invariant images. Previously this model fitting normally involves troublesome camera calibration, but with our new method, this process becomes straightforward and automatic.

3. Utilising the reference shadows, have detected the presence of cloud in the scenes and then used that to generate a reference image which is then used to produce coloured shadowless images for other frames.

4. the reference shadows, have detected the presence of cloud in the scenes and then used that to generate a reference image which is then used to produce coloured shadowless images for other frames.

The work-flow to generate shadowless images is illustrated in the following Figure.

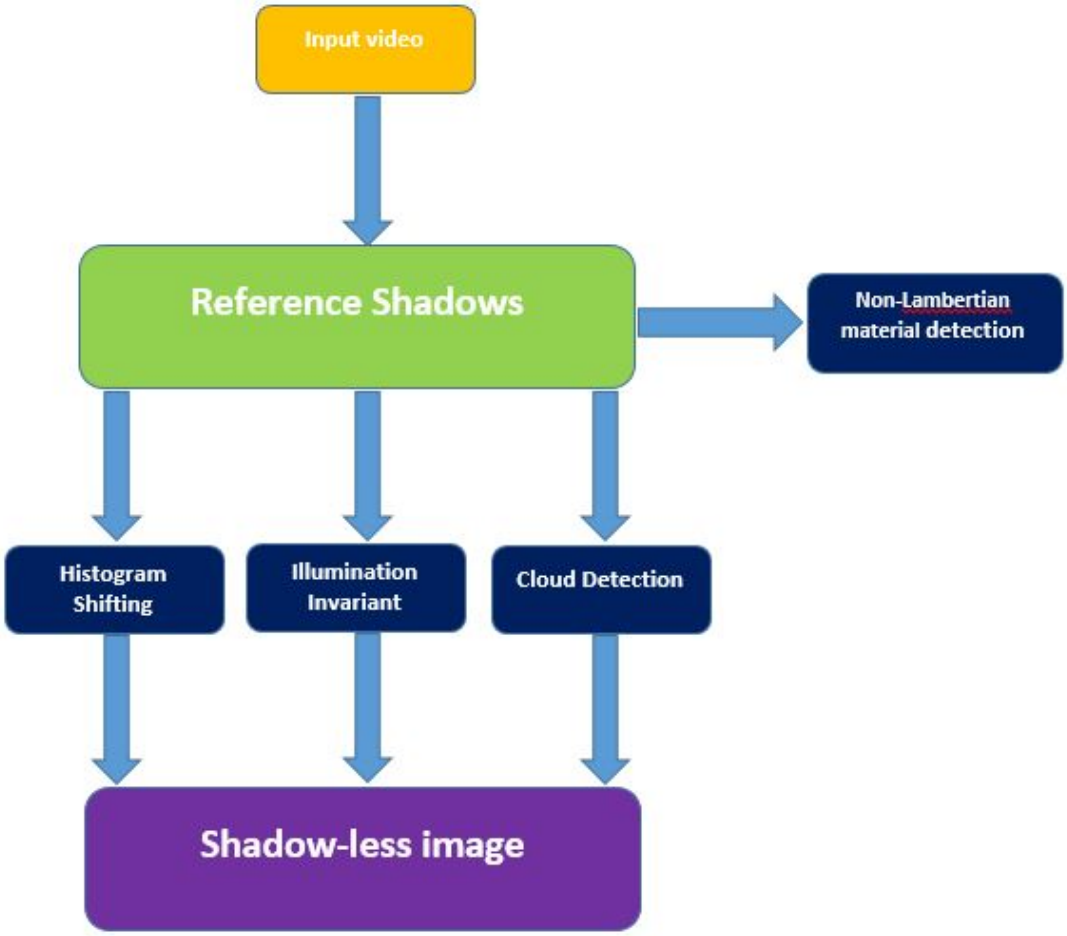


FIGURE 6.1: Video processing mechanism to produce shadowless image.

The results of our method have been compared with some of the state of the art systems in outdoor environments. So to summarise, the following subjects have been covered in this thesis:

6.2.1 Reference shadows

Reference shadows are the key component of the shadow detection algorithm mentioned in this document. They can be distinguished from other dark regions in the video by analysing the trajectory of their centroid or their over a period. The centroid of Reference shadows move linearly

over a short period of time and this is because of the change in position of the sun during the day.

To detect the reference shadow, we analysed and processed the position of the centre of the shadow instead of its length. In a short period, the changes to the position of the centre of the shadow will have similar pattern as the length. Our experiments confirmed monitoring the centre of the shadow will be sufficient to distinguish reference shadow from other dark objects.

Additionally, the impact of other factors such as the shape of the object and camera view will be minimum in short sampling periods and will not change the trajectory of the centre (or corner or edge) of the shadow. Based on our experiments, using the corner of shadows will improve the performance of shadow detection system when the shadows are located at image boundaries.

6.2.2 Illumination invariant images

Illumination invariant images do not have any illumination information such as shadows. We have used reference shadows to produce illumination invariant automatically unlike methods such as [51] where they used manual camera calibration. Illumination invariant images do not have any colour information and they are greyscale. To produce Illumination invariant images, To create the Invariant Image without camera calibration, we generate a large number of gs' with different values of C in [0,1]. At some point [c_1, c_2] becomes orthogonal to lighting direction and shadow edges will disappear. To automatically detect the shadow

edges, we monitored the edge of reference shadows and once the edge is disappeared, the Reference shadow is removed together with the rest of the shadows.

Furthermore, we analysed the edge of reference shadow over the time and processed the impact on the material background when it becomes engulfed in shadow. We formulated the impact of shadow on the histograms in R, G and B colour spaces and then applied the same model to the reset of the shadow region and successfully produced good quality coloured shadowless image. Here we assumed there is no material change at the shadow edge.

6.2.3 Detecting the presence of clouds

Detecting the presence of cloud may not have huge value in systems such as tracking or shape recognition however when there are clouds, shadows disappear. Hence knowing that this event has happened would allow us to produce high quality shadowless reference images which can be used to detect and remove shadows in other parts of the video where there are no clouds. In this thesis we have provided a robust method and the presence of clouds are successfully detected by analysing the Reference shadow characteristics.

To do this we used the reference shadows again over the time and analysed the illumination and chromaticity of Reference shadows when the clouds appear. It was observed that during such a scenario, the reference shadow region becomes brighter and the previous non shadow region. Additionally our experiments show that the shadow areas are bluer that

non-shadowed regions. This is because the only source of illumination in shadow regions is the sky which is perceived as blue. When a pixel or a region is illuminated by a blue coloured light source, they tend to reflect blue light more and hence they look bluer. Conversely, the non-shadowed regions are illuminated by both sky and the sun where the illumination from sun is far more dominant than sky. This means the non-shadowed regions will not be as blue as shadows.

6.2.4 Identifying shiny materials in video

Another important method which is presented in this thesis is a system which can automatically detect non-Lambertian surface materials. We have illustrated matt surfaces and shiny material display different behaviours under same illumination models. Lambertian surfaces have linear pattern but shiny materials such as water, glass and metal do not show the linear trajectory and we have used this characteristic to distinguish them from matt materials.

To implement this, we projected the logarithm of chromaticity ($R'(R) = log(R/G)$ and $R'(B) = Log(B/G)$) and calculated the gradient of the sample regions. It was demonstrated that ratio of the change in shiny material was significantly higher than matt surface and this test can be used to distinguish these different objects.

The same concept was also used to distinguish reflection glare from white matt objects and illumination invariant images where used to remove the impact of excessive reflectance from the video. This solution is used for applications such as water and leakage detection and so on.

6.3 Research limitations

The research in this thesis is focused on detection of shadows, clouds and shiny textures but there are some limitations and assumptions that should be considered:

1. This research is based on outdoor condition with natural light.
2. There should not be any secondary source of light.
3. The camera should be static.

The above assumption are correct for most outdoor standard CCTV deployments but the methods introduced in this thesis can not be used where there are multiple sources of artificial light. Additionally to detect reference shadows, the camera should be stationary therefore the proposed methods in this thesis can not be used with moving cameras such as autonomous cars.

6.4 Further work

In this thesis, a novel method to detect a single shadow (reference shadow) was presented. Additionally, the impact of illumination variance has been discussed and analysed. The difference between Lambertian and non-Lambertian surface (shiny and matt) has been illustrated and a novel method to distinguish was presented. The reference shadow was then used to produce Illumination Invariant gray scale image. Some of the further research topics that can be followed are listed below:

- Illumination Invariant image loses all the illumination information of the image including colour information. Additional procedures will be needed to produce coloured shadowless image.
- When a single shadow is detected, other systems such as Neural Networks can be used to train the learning model and detect and remove the reference shadows.
- How different of glossy textures can be detected.
- How can shadows be identified when the camera is not static such as autonomous cars.

The above topic features are only some aspects of the further research that can be conducted in this field. Moreover, the integration of our shadow detection system with tracking and object recognition solutions would be valuable for video processing applications.

Bibliography

[1] B. Min and H. Wei and W. Chen and L. Yao. " Shadow detection in color aerial sensing images based on adaptive pulse coupled neural network". In: (2014), pp. 521–525.

[2] B. Zhong and W. Chen and S. Wu and L. Hu and X. Luo and Q. Liu. "A Cloud Detection Method Based on Relationship Between Objects of Cloud and Cloud-Shadow for Chinese Moderate to High Resolution Satellite Imagery". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10.11 (2017), pp. 4898–4908.

[3] C. Wang and L. Deng and Z. Zhou and M. Yang and B. Wang. " Shadow detection and removal for illumination consistency on the road". In: (2017), pp. 198–203.

[4] D. J. R. Del Carmen and R. D. Cajote. " Moving shadow detection and removal for video-based traffic monitoring". In: (2016), pp. 1–7.

[5] D. Tulpan and C. Bouchard and K. Ellis and C. Minwalla. "Detection of clouds in sky/cloud and aerial images using moment based texture segmentation". In: (2017), pp. 1124–1133.

[6] G. C. Iannelli and P. Gamba and X. Li and H. Shen. “Improving urban extent extraction from VHR optical data by means of cloud detection and image reconstruction”. In: (2016), pp. 7307–7309.

[7] H. Lv and Y. Wang and Y. Yang. “Thin cloud detection using spectral similarity in coastal and blue bands of Landsat-8 data”. In: (2017), pp. 4677–4680.

[8] H. Xu and Z. Yang and M. Tian and Y. Sun and G. Liao. “ An Extended Moving Target Detection Approach for High-Resolution Multichannel SAR-GMTI Systems Based on Enhanced Shadow-Aided Decision, year=2018, volume=56, number=2, pages=715-729, month=Feb,” in: *IEEE Transactions on Geoscience and Remote Sensing* ().

[9] J. Ji and X. Jiang and W. Sun. “ Shadow detection using double-threshold pulse coupled neural networks”. In: (2016), pp. 1971–1975.

[10] J. Qian and Y. Luo and Y. Wang and D. Li. “Cloud Detection of optical remote sensing image time series using Mean Shift algorithm”. In: (2016), pp. 560–562.

[11] K. Yuan and G. Meng and D. Cheng and J. Bai and S. Xiang and C. Pan. “Efficient cloud detection in remote sensing images using edge-aware segmentation network and easy-to-hard training strategy”. In: (2017), pp. 61–65.

[12] L. Gómez-Chova and G. Mateo-García and J. Muñoz-Marí and G. Camps-Valls. “Cloud detection machine learning algorithms for PROBA-V”. In: (2017), pp. 2251–2254.

[13] Liao Jie. " Moving shadows detection algorithm and implementation based on dual background modeling". In: 01 (2015), pp. 1557–1561.

[14] M. Le Goff and J. Y. Tourneret and H. Wendt and M. Ortner and M. Spigai. "Deep learning for cloud detection". In: (2017).

[15] M. Russell and J. J. Zou and G. Fang. " A Novel Region-Based Method for Moving Shadow Detection". In: (2016), pp. 1–6.

[16] M. Russell and J. J. Zou and G. Fang. " Human shadow detection for real-time applications". In: (2016), pp. 1610–1613.

[17] N. Mo and R. Zhu and L. Yan and Z. Zhao. " Deshadowing of Urban Airborne Imagery Based on Object-Oriented Automatic Shadow Detection and Regional Matching Compensation". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11.2 (2018), pp. 585–605.

[18] S. Danda and A. Challa and B. S. Daya Sagar. "A morphology-based approach for cloud detection". In: (2016), pp. 80–83.

[19] S. Musleh and M. Sarfraz and L. Niepel. "A Comparative Study on Shadow Detection Methods Based on Features". In: (2018), pp. 1–6.

[20] S. R. Surya and M. A. Rahiman. "Cloud detection from satellite images based on Haar wavelet and clustering". In: (2017), pp. 163–167.

[21] T. F. Y. Vicente and M. Hoai and D. Samaras. " Leave-One-Out Kernel Optimization for Shadow Detection and Removal, year=2018,

volume=40, number=3, pages=682-695, month=March," in: *IEEE Transactions on Pattern Analysis and Machine Intelligence* ().

[22] W. Yanfeng and G. Ningsheng and G. Xilong. " Research on method for moving shadow detection". In: (2017), pp. 1–4.

[23] Wei Ji and Y. Zhao. " Moving cast shadow detection using joint color and texture features based on direction and distance". In: (2016), pp. 439–444.

[24] X. F. Yang and W. C. Siu. " Vehicle detection under tough conditions using prioritized feature extraction with shadow recognition". In: (2017), pp. 1–5.

[25] X. Kang and Y. Huang and S. Li and H. Lin and J. A. Benediktsson. " Extended Random Walker for Shadow Detection in Very High Resolution Remote Sensing Images". In: *IEEE Transactions on Geoscience and Remote Sensing* 56.2 (2018), pp. 867–876.

[26] Z. Li and H. Shen and H. Li and L. Zhang. "Automatic cloud and cloud shadow detection in GF-1 WFV imagery using multiple features". In: (2016), pp. 7612–7615.

[27] V. Agarwal, A. V. Gribok, and M. A. Abidi. "Machine learning approach to color constancy". In: *Neural Networks* 14.3 (2007), pp. 559–563.

[28] D. Akers et al. "Conveying Shape and Features with Image-Based Relighting." In: (2003), pp. 349–354.

[29] H. Amini and B. Karasfi. "New approach to road detection in challenging outdoor environment for autonomous vehicle". In: (2016), pp. 7–11.

[30] C. Andres and B. Lovell. "Improved Shadow Removal for Robust Person Tracking in Surveillance Scenarios," in: (2010), pp. 141–144.

[31] M. Antone and M. Bosse. "Calibration of outdoor cameras from cast shadows". In: (2004), pp. 3040–3045.

[32] C. Arora and M. Werman. "Optical flow for non Lambertian surfaces by cancelling illuminant chromaticity". In: (2014), pp. 1977–1981.

[33] R. Bajcsy, S. Lee, and A. Leonardis. "Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation". In: (1996), pp. 241–272.

[34] R. Bajscy, S. Lee, and A. Leonardis. "Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation". In: (1996), pp. 241–272.

[35] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. "Systems and Experiment: Performance of Optical Flow Techniques". In: (1994).

[36] H. Barrow and J. Tenenbaum. "Recovering intrinsic scene characteristics from images". In: *Computer Vision systems* (1978), pp. 3–26.

[37] R. Basri and D. W. Jacobs. "Lambertian reflectance and linear subspaces". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25.2 (2003), pp. 218–233.

[38] M. Bell and W. Freeman. “Learning local evidence for shading and reflection”. In: 1 (2001), pp. 670–675.

[39] H. Boroujeni, N. Charkari, and A. Jalilvand. “A robust moving shadow detection algorithm based on semi-supervised hierarchical mixture of MLP-experts”. In: (2011), pp. 141–146.

[40] V. Cardei, B. Funt, and K. Barnardi. “Estimating the Scene Illumination Chromaticity Using a Neural Network”. In: *Optical Society of America* 9.1 (2002), pp. 2374–2386.

[41] A. Cavallaro, E. Salvador, and T. Ebrahimi. “Detecting Shadows In Image Sequences”. In: (2004), pp. 15–16.

[42] A. Chakrabarti and K. Hirakawa. ““Computational color constancy with spatial correlation”. In: (2010).

[43] M. Cohen, A. Colburn, and S. Drucker. “Image stacks technical Report”. In: (2003), pp. 727–730.

[44] F. Cozman and E.Krotkov. “Robot Localization using computer vision”. In: (1997).

[45] K. Deb and A. Suny. “Shadow Detection and Removal Based on YCbCr Color Space”. In: (2014), pp. 23–33.

[46] M. Ebner. “‘Color constancy based on local space average color”. In: (2009), pp. 283–301.

[47] M. El-Zahhar, A. Karali, and M. ElHelw. “A semi supervised learning-based method for adaptive shadow detection”. In: (2012), pp. 384–353.

[48] E.Talvala et al. “Veiling Glare in High Dynamic Range Imaging”. In: (2013).

[49] M. Faghih and M. Moghaddam. “Neural Gray Edge: improving gray edge algorithm using neural networks”. In: (2011), pp. 1705–1708.

[50] R. Feris and R. Rashkar. “Specular reflection reduction with multi-flash imaging”. In: (2004), pp. 316–321.

[51] G. Finlayson, S. Hordley, and M. Drew. “On the removal of shadows from images”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28.1 (2006), pp. 59–68.

[52] Graham D. Finlayson, Mark S. Drew, and Cheng Lu. “Entropy Minimization for Shadow Removal”. In: (2009), pp. 35–57.

[53] H. Fujikake et al. “Video camera system using liquidcrystal polarizing filter to reduce reflected light,” in: (1999).

[54] G. Ged et al. “Recognizing real materials from their glossy appearance”. In: *Journal of Vision* 10.18 (2010), pp. 451–465.

[55] R. Gershon, A. D. Jepson, and J. K. Tsotsos. “Highlight identification using chromatic information”. In: (1987), pp. 161–170.

[56] A. Gijsenij, R. Lu, and T. Gevers. “‘Color constancy based on local space average color”. In: (2012), pp. 697–708.

[57] A Gijsenij and J. Weije. “‘Computational color constancy: survey and experiment”. In: (2011), pp. 2475–2489.

[58] J. Hariyono and K. H. Jo. “Detection of pedestrian crossing road using action classification model”. In: (2015), pp. 21–24.

[59] M. Heikkila and M. Pietikainen. "Adaptive shadow detection using global texture and sampling deduction". In: (2013), pp. 115–122.

[60] A. Hertzmann and S. M. Seitz. "Example-based photometric stereo: shape reconstruction with general, varying BRDFs". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27.8 (2005), pp. 1254–1264.

[61] B. K. Horn. "Robot Vision". In: (1986).

[62] P. Hough. "Method and means for recognizing complex patterns". In: (1960).

[63] J. Huang and C. Chen. "Moving cast shadow detection using physics-based features". In: (2009), pp. 2310–2317.

[64] I. Huerta, M. Holte, and T. Moeslund. "Detection and removal of chromatic moving shadows in surveillance scenarios". In: (2009), pp. 1499–1506.

[65] S. Iwata and K. Ogatai. "Specular reflection removal with high-speed camera for video imaging". In: (2015), pp. 1735–1740.

[66] N. Jacobs and N.Roman. "Consistent temporal variations in many outdoor scenes". In: (2007).

[67] A. Jalilvand and N. Charkari. "A novel coloured local invariant descriptor based on SURF". In: (2011), pp. 214–219.

[68] J. Ji, X. Jiang, and W.i Sun. "Shadow detection using double-threshold pulse coupled neural networks". In: (2016), pp. 1971–1975.

[69] K. Jiang et al. "A texture-based method for modeling the background and detecting moving objects". In: (2006), pp. 657–662.

[70] H. M. Kahily and A. P. Sudheer. "Real-time human detection and tracking from a mobile armed robot using RGB-D sensor". In: (2016), pp. 1–6. DOI: 10.1109/STARTUP.2016.7583953.

[71] S. Khan and M. Bennamouni. "Automatic Shadow Detection and Removal from a Single Image". In: (2015), pp. 431–446.

[72] H. Kim, S. Hadap H. Jin, and I. Kweoni. "Specular reflection separation using dark channel prior". In: (2013), pp. 1460–1467.

[73] Yong Sam Kim et al. "Real time detection of moving human based on digital image processing". In: (2007).

[74] G. Klinker, S. A. Shafer, and T. Kanade. "Image segmentation and reflection analysis through color". In: (1988), pp. 229–244.

[75] K. Knoblauch and F. Viéot. "Difference scaling of gloss: Nonlinearity, binocularity, and constancy". In: *Journal of Vision* 10 (2010), pp. 189–203.

[76] S. Koppal and S. Narasimhan. "Clustering appearance for scene analysis". In: (2006).

[77] G. Kuwabara. "On the flare of lenses". In: (1952), pp. 451–463.

[78] J. Lalonde, S. Narasimham, and A. Efros. "What do the Sun and the Sky Tell Us About the Camera". In: (2009).

[79] J. Lalonde, S. Narasimhan, and A. Efros. "What do the sun and the sky tell us about the camera?" In: (2008), pp. 354–367.

[80] S. Lee and R. Bajcsy. “Detection of specularity using color and multiple views”. In: (1992), pp. 643–653.

[81] G. Li, Y. Liu, and Q. Dai. “Multi-view photometric stereo of non-Lambertian surface under general illuminations”. In: (2011), pp. 1–6.

[82] S. Lin et al. “Diffuse-specular separation and depth recovery from image sequences”. In: (2002), pp. 2352–2364.

[83] J. Malik and T.Leung. “Representing and Recognizing the Visual Appearance of Materials using Three-dimensional Textons”. In: *International JOurnal of Computer Vision* (2001), pp. 29–43.

[84] S. Mallick et al. “Beyond lambert: Reconstructing specular surfaces using color”. In: (2005), pp. 619–626.

[85] L. Maloney and J. Yang. “Maximum likelihood difference scaling”. In: *Journal of Vision* 6 (2003), pp. 573–568.

[86] K. Matsuda and I. Nitoh. “Flare as applied to photographic lenses”. In: (1972), pp. 198–216.

[87] Y. Matsushita et al. “Estimating intrinsic images from image sequences with biased illumination”. In: (2012), pp. 274–276.

[88] Y. Matsushita et al. “Illumination Normalization with Time-Dependent Intrinsic Images for Video Surveillancel”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26.10 (2004), pp. 1336–1341.

[89] J. McCann and A.Rizzu. “Veiling glare: the dynamic range limit of HDR images.” In: (2007), pp. 1450–1466.

[90] A. Monnet et al. "Background modeling and subtraction of dynamic scenes," in: (2003), pp. 1305–1312.

[91] M. D. E. Munajat, D. H. Widyantoro, and R. Munir. "Road detection system based on RGB histogram filterization and boundary classifier". In: (2015), pp. 195–200.

[92] S. Narasimhan, G.Wang, and C. Nayar. "All images of an outdoor scene". In: (2002).

[93] S. N. Narasimhan and S. K. Nayar. "Shedding light on the weather". In: (2003), pp. 665–672.

[94] F. Ortiz and F. Torres. "Automatic Detection and Elimination of Specular Reflectance in Color Images by Means of MS Diagram and Vector Connected Filters". In: (2006), pp. 681–688.

[95] N. Otsu. "A threshold selection method from gray level histograms". In: (1979), pp. 62–66.

[96] R. Perez, R. Seal, and J. Michalskey. "All-weather model for sky luminance distribution—preliminary configuration and validation". In: (1993).

[97] F. Porikli and J. Thornton. "Shadow flow: a recursive method to learn moving cast shadows". In: (2005), pp. 891–898.

[98] V. Prinet, D. Lischinski, and M. Werman. "Illuminant chromaticity from image sequences". In: *The IEEE International Conference on Computer Vision (ICCV)* 3 (2014), pp. 311–319.

[99] R. Raskar et al. “A nonphotorealistic camera: depth edge detection and stylized rendering using multi-flash imaging”. In: (2004), pp. 679–688.

[100] Y. Sato and K. Ikeuchi. “Temporal-color space analysis of reflection”. In: (1993), pp. 570–576.

[101] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur. “Multiplexing for Optimal Lighting”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29.8 (2007), pp. 1339–1354.

[102] M. Seki et al. “Background subtraction based on cooccurrence of image variations”. In: (2003), pp. 54–74.

[103] A. Shafer. “Using color to separate reflection components”. In: (1985), pp. 210–218.

[104] S. A. Shafer. “Using color to separate reflection components”. In: (1985), pp. 210–218.

[105] A. Sharma, A. Singh, and R. Rohilla. “Color based human detection and tracking algorithm using a non-Gaussian adaptive Particle filter”. In: (2016), pp. 439–442.

[106] L. Shen, P. Tan, and S. Lin. “Intrinsic image decomposition with non-local texture cues”. In: (2008), pp. 1–7.

[107] R. Stanikunas, H. Vaitkevicius, and J. Kulikowski. “Almost complete colour constancy achieved with full-field adaptation”. In: *Neural Networks* 17.3 (2004), pp. 327–337.

[108] J. Stumpfel, A.Wenger, and A.Tchou. “Direct HDR capture of the sun and sky”. In: (2004).

[109] K. Sunkavalli and W. Matusik. "Factored time-lapse video". In: (2007).

[110] R. Tan and K. Ikeuchi. "Separating reflection components of textured surfaces using a single image". In: (2005), pp. 178–193.

[111] R. Tan, L. Quan, and S. Lini. "Separation of highlight reflections on textured surfaces". In: (2006), pp. 1855–1860.

[112] M. Tao, J. Su, and T.Wang. "Depth Estimation and Specular Removal for Glossy Surfaces Using Point and Line Consistency with Light-Field Cameras". In: (2016), pp. 1155–1164.

[113] M. Tappen, W. Freeman, and E. Adelson. "Recovering intrinsic images from a single image". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (2005), pp. 1459–1472.

[114] A. Trebi-Ollennu and T.Cheng. "Design and analysis of a sun sensor for planetary rover absolute heading detection." In: (2001).

[115] V. D. Tsai. "A Comparative Study on Shadow Compensation of Color Aerial Images in Invariant Color Modelss". In: (2006), pp. 1661–1671.

[116] T. Tsujii. "High-speed stroboscope for specular reflection removal of DC illumination". In: (2010), pp. 2704–2709.

[117] T. Tsujii. "Specular reflection removal on high-speed camera for robot vision". In: (2010), pp. 1542–1547.

[118] Junxian Wang et al. "Specular reflection removal for human detection under aquatic environment," in: (2004).

[119] Y. Weiss. "Deriving intrinsic images from image sequences". In: (2001), pp. 68–75.

[120] J. Winn and J. Shotton. "The Layout Consistent Random Field for Recognizing and Segmenting Partially Occluded Objects". In: (2006), pp. 37–44.

[121] Q. Yang, S. Wang, and N. Ahujai. "Real-time specular highlight removal using bilateral filtering". In: (2006), pp. 1460–1467.

[122] Q. Yang et al. "A Uniform Framework for Estimating Illumination Chromaticity, Correspondence, and Specular Reflection". In: *IEEE Transactions on Image Processing* 20.1 (2011), pp. 53–63.

[123] K. Yoon, Y. Choi, and I. Kweon. "Fast separation of reflection components using a specularity-invariant image representation". In: (2006), pp. 973–976.

[124] K. Yoon and I. Kweon. "Correspondence search in the presence of specular highlights using specular-free two-band images". In: *Computer Vision – ACCV* 7 (2006), pp. 761–770.

[125] Y. Yu and J.Malik. "Recovering photometric properties of architecturalscenes from photographs". In: (1998).

[126] X. Yuan, M. Ebner, and Z. Wang. "Single-image shadow detection and removal using local colour constancy computation". In: (2014), pp. 118–126.

[127] Y. Zhang et al. "Road and vehicle detection in highway scene for automotive FMCW antenna array radar". In: (2015), pp. 1–5.