SHADOW REMOVAL UTILIZING MULTIPLICATIVE FUSION OF TEXTURE AND COLOUR FEATURES FOR SURVEILLANCE IMAGE

TEO KAH MING

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Master of Philosophy

School of Electrical Engineering
Faculty of Engineering
Universiti Teknologi Malaysia

SEPTEMBER 2018

ACKNOWLEDGEMENTS

First and foremost, I would like to use this great opportunity to express my sincere gratitude to my supervisor, Dr. Lim Cheng Siong for the continuous support of my study. His motivation, guidance, patience and valuable comments have helped me a lot on the long journey to finish my research and completing of this thesis. Without his continuous support, the objectives of this study would not be accomplished as expected.

Completing this work would have been more difficult were it not for the support and friendship provided by the Faculty of Electrical Engineering (FKE) staffs. I am indebted to them for their help. I would also like to extend my sincere thanks to Universiti Teknologi Malaysia (UTM) as well for granting the Research Student Grant (RSG) under the Research University Grant (GUP) to provide a financial support throughout my study.

Finally, I must express my very profound gratitude to my parents and to my brother for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

ABSTRACT

Automated surveillance systems often identify shadows as parts of a moving object which jeopardized subsequent image processing tasks such as object identification and tracking. In this thesis, an improved shadow elimination method for an indoor surveillance system is presented. This developed method is a fusion of several image processing methods. Firstly, the image is segmented using the Statistical Region Merging algorithm to obtain the segmented potential shadow regions. Next, multiple shadow identification features which include Normalized Cross-Correlation, Local Color Constancy and Hue-Saturation-Value shadow cues are applied on the images to generate feature maps. These feature maps are used for identifying and removing cast shadows according to the segmented regions. The video dataset used is the Autonomous Agents for On-Scene Networked Incident Management which covers both indoor and outdoor video scenes. The benchmarking result indicates that the developed method is on-par with several normally used shadow detection methods. The developed method yields a mean score of 85.17% for the video sequence in which the strongest shadow is present and a mean score of 89.93% for the video having the most complex textured background. This research contributes to the development and improvement of a functioning shadow eliminator method that is able to cope with image noise and various illumination changes.

ABSTRAK

Sistem pengawasan automatik bayangan sering mengesan bayang-bayang sebagai sebahagian daripada objek bergerak dan ini akan seterusnya menjejaskan pelbagai tugas pemprosesan imej seperti pengenalpastian dan penjejakan objek. Tesis ini membentangkan kaedah yang lebih berkesan seperti penyingkiran bayang-bayang bagi sistem pengawasan bangunan. Kaedah yang digunakan ini merupakan gabungan beberapa kaedah pemprosesan imej. Proses pertama menggunakan teknik penggabungan kawasan imej secara statistik untuk membahagikan imej dan seterusnya mendapatkan kawasan bayang-bayang yang berpotensi. Proses seterusnya merangkumi gabungan ciri-ciri pengesanan bayang-bayang termasuk Korelasi Silang secara normal, Ketetapan Warna Setempat dan pengesanan bayang Nilai Ketepuan Hue untuk menghasilkan potensi imej pemetaan bayang-bayang. Pemetaan ini seterusnya digunakan untuk mengenalpasti dan menyingkirkan bayang-bayang mengikuti kawasan bayangan berpotensi yang telah dibahagi. Rakaman video yang digunakan berasal dari Ejen Autonomi untuk Pengurusan Insiden Rangkaian Terhad yang meliputi keadaan luar dan dalam bangunan. Keputusan ujikaji menunjukkan bahawa kaedah yang digunakan dapat memperolehi pengesanan yang setara dengan beberapa kaedah pengesanan bayang yang lain. Kaedah yang digunakan menghasilkan skor min 85.17% untuk video yang mempunyai bayang-bayang paling ketara dan skor min 89.93% untuk video yang mempunyai latar belakang bertekstur yang paling kompleks. Penyelidikan ini menyumbang kepada pembangunan dan kemajuan sistem penghapusan bayang-bayang untuk mengatasi masalah gangguan imej dan pelbagai perubahan pancaran cahaya.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

$\lvert \boldsymbol{A} \rvert$	-	Matrix determinant of $\boldsymbol{A}$
$\boldsymbol{A}$	-	(Uppercase bold) Matrices
$\boldsymbol{A} \cdot\!* \boldsymbol{B}$	-	Element-by-element multiplication of matrices **A**, **B**
$\boldsymbol{A}^{Tr}$	-	Matrix transpose of $\boldsymbol{A}$
$\boldsymbol{a}$	-	(Lowercase bold) Vectors
$\boldsymbol{a} \cdot \boldsymbol{b}$	-	Scalar product of two vectors $\boldsymbol{a}$, $\boldsymbol{b}$
$\arg\max_{x}(\, f(x)\,)$	-	The value of x that causes function $f(x)$ to be maximum
$\arg\min_{x}(\, f(x)\,)$	-	The value of x that causes function $f(x)$ to be minimum
B	-	(Superscript, uppercase) Refers to the Blue channel in RGB image
B'	-	Normalized blue channel value, where $B' = {}^{B}/_{255}$
B^{GMM}	-	Number distributions that are chosen as background
$b^2(\cdot)$	-	A substitutional function
$\mathcal{C}_a$	-	Intensity strength of the light source
$\mathcal{C}_p$	-	Intensity strength of the reflectance light
C_{MAX}	-	A substitutional variable used in describing the RGB to HSV color conversion

C_{MIN}	-	A substitutional variable used in describing the RGB to HSV color conversion
C_j^{otsu}	-	Refers to class separation used in the conventional Otsu method, where j represents the divided classes
d	-	Color space dimension of the intensity at pixel $\boldsymbol{X}$
$\mathcal{d}$	-	Distance between the object and the illumination source
E^k	-	The illumination component of a color channel k described in the luminance model
$\exp(\,f(x)\,)$	-	Exponential function on function $f(x)$
F_i^{marked}	-	The available marked frame at frame index i in the video sequence.
F_t	-	Refers of a frame in a video sequence at time t
FN	-	False negative. Used to classify pixels which are incorrectly rejected
FP	-	False positive. Used to classify pixels which are incorrectly identified
$f^{SRM}(\mathcal{R}_{\boldsymbol{I}}^k, \mathcal{R}'^k_{\boldsymbol{I}})$	-	Gradient function to compute the numerical relationship value of two connected paired region $\mathcal{R}_{\boldsymbol{I}}^k$ and $\mathcal{R}'^k_{\boldsymbol{I}}$
$f_{\boldsymbol{X}}$	-	Number of occurrences for the specific intensity value $\boldsymbol{X}$ to appear in the entire image
G	-	(Superscript, uppercase) Refers to the Green channel in RGB image
G'	-	Normalized green channel value, where $G' = {}^{G}/_{255}$
H	-	(Superscript, uppercase) Refers to the Hue channel in HSV image
$\boldsymbol{I}$	-	(Uppercase bold) Refers to a color image

$\boldsymbol{I}^*$ - Theoretical image of $\boldsymbol{I}$

$\boldsymbol{I}^k$ - Refers to a specific color channel k in a color image

$\boldsymbol{I}_{back}$ - Background color image

$\boldsymbol{I}_{curr}$ - Current frame color image

i, j - Denotes a number that indicates the location or count of a variable in an indexed family or set

$[i, j]$ - A notation to indicate that all the pixel intensity values found in the image $\boldsymbol{I}$, are of the given specified range, such that $\{ \boldsymbol{X} \in \boldsymbol{I} \mid i \leq \boldsymbol{X} \leq j \}$

K - Number of Gaussian distribution used

K' - Least probable distribution in the K Gaussian distribution

k - Refers to a specified color channel available in image $\boldsymbol{I}$

$\vec{\boldsymbol{L}}$ - (Uppercase, bold with arrow) Direction of the illumination source

L - Maximum possible number of pixels (Gray level) in an image, $L = 256$ for an 8-bit image

l - Refers to the side of a square, used in describing the region of the neighboring pixel $\Omega_{\boldsymbol{u}}$

$lower$ - (Subscript) Specified lower bound threshold used to tune the value of the associated variable. α^V_{lower}, β^V_{lower}, τ^S_{lower}, τ^H_{lower}

$\boldsymbol{M}$ - (Uppercase bold) Refers to a binary image or mask

$\boldsymbol{M}_{fore}$ - Foreground binary mask

$\boldsymbol{M}_{gt}$ - Ground truth binary mask

M - Refers to the height or maximum number of rows in an image

M_{test} - Matching decision test

MAP	-	(Superscript) Denotes the resultant output after executing the modified linear combination technique
$Match_{i,t}$	-	Matching decision result on the i^{th} Gaussian at time t
$m_{k,t}$	-	Model representing a mixture of K Gaussian distribution at time t
$mean$	-	Mean value of both shadow detection and shadow discrimination rate
$\vec{\boldsymbol{N}}$	-	(Uppercase, bold with arrow) Direction of the object surface normal
N	-	Refers to the width or maximum number of columns in an image
$N[\,f(x)\,]$	-	The normalization operation on vector $f(x)$, so that all values of $f(x)$ lies within a specified range controlled by a upper and lower bound limit
$\mathcal{N}(\,A \mid \mu\,, \sigma^2)$	-	Standard Gaussian distribution of value A with mean μ and variance σ^2
$\mathbb{N}$	-	Notation for the set of natural number, where $\mathbb{N} = [0, +\infty)$
obj	-	(Subscript) Use along with a variable to indicate that the associated variable is classified as object, e.g. $\boldsymbol{M}_{obj}$, $\mathcal{R}_{obj}$, TP_{obj}, FN_{obj}, $\boldsymbol{u}_{obj}$, μ_{obj}
out	-	(Superscript) Denotes the resultant output generated after the spatial adjustment stage
$\Pr(\,A\,)$	-	Probability of event A
$\Pr(\,A \mid B\,)$	-	Posterior probability. Probability of event A given that event B had occurred
$p_{\boldsymbol{X}}$	-	Probability of occurrence for the specific intensity value $\boldsymbol{X}$ to appear in the entire image

$\mathcal{P}(\mathcal{R}_{\boldsymbol{I}}^{k}, \mathcal{R}'^{k}_{\boldsymbol{I}})$	-	Merging predicate function of two regions $\mathcal{R}_{\boldsymbol{I}}^{k}$ and $\mathcal{R}'^{k}_{\boldsymbol{I}}$
Q	-	Controls and quantify the statistical complexity of **I***
R	-	(Superscript, uppercase) Refers to the Red channel in RGB image
R'	-	Normalized red channel value, where $R' = {}^{R}/_{255}$
$\mathbb{R}$	-	Notation for the set of real number, where $\mathbb{R} = (-\infty, +\infty)$
$\mathcal{R}$	-	A region formed in the image
$\mathcal{R}_{\boldsymbol{I}}^{k}$	-	A bounded enclosed region of a specific color channel k, in image $\boldsymbol{I}$
$\mathcal{R}'^{k}_{\boldsymbol{I}}$	-	The adjacent bounded enclosed region of $\mathcal{R}_{\boldsymbol{I}}^{k}$ for a specific color channel k, in image $\boldsymbol{I}$, such that $\mathcal{R}_{\boldsymbol{I}}^{k}$ and $\mathcal{R}'^{k}_{\boldsymbol{I}}$ shares a common boundary
$\overline{\mathcal{R}_{\boldsymbol{I}}^{k}}$	-	The observed average for a specific color channel k, in region $\mathcal{R}_{\boldsymbol{I}}$
r_i	-	Criterion ratio of the i^{th} Gaussian
r^k	-	Reflectance component of a color channel k described in the luminance model
S	-	(Superscript, uppercase) Refers to the Saturation channel in HSV image
sha	-	(Subscript) Use along with a variable to indicate that the associated variable is classified as shadow, e.g. $\boldsymbol{M}_{sha}$, $\mathcal{R}_{sha}$, TP_{sha}, FN_{sha}, $\boldsymbol{u}_{sha}$, μ_{sha}p
$\mathcal{S}_{pairs}$	-	A list of connected pixels pair
T	-	Variable to describe the threshold value
T^{GMM}	-	Specified threshold value used for indicating the minimum portion of data that should be considered as background

T^{otsu}	-	A theoretical variable used in the conventional Otsu method
T_*^{otsu}	-	The output optimal threshold value determined by the conventional Otsu method
TN	-	True negative. For classifying pixels which are correctly rejected
TP	-	True positive. Used to classify pixels which are correctly identified
t	-	Refers to the time in the video sequence
$Total_x$	-	Total number of the pixel that falls under the category x.
$upper$	-	(Subscript) Specified upper bound threshold used to tune the value of the associated variable. α_{upper}^{V} , β_{upper}^{V} , τ_{upper}^{S} , τ_{upper}^{H}
$\boldsymbol{u}$	-	(Lowercase bold) Refers to a pixel position in spatial coordinate
V	-	(Superscript, uppercase) Refers to the Value channel in HSV image
$\boldsymbol{X}$	-	Pixel intensity value
x	-	Pixel spatial coordinate in the x-direction
$\mid x \mid$	-	Absolute value of a scalar x
y	-	Pixel spatial coordinate in the y-direction
α^{GMM}	-	Distribution learning rate 1
α^{V}	-	Specified lower bound threshold for the Value component
β^{V}	-	Specified upper bound threshold for the Value component
Δ_C	-	A substitutional variable used in describing the RGB to HSV color conversion

δ^{HSV}	-	Denotes a constant value used in HSV shadow cue algorithm, to avoid division by zero
δ^{LCC}	-	A numerically small scalar quantity value used by LCC algorithm to avoid division by zero
δ^{SRM}	-	Constant variable used to denote the maximum probability when $\mathcal{P}(\mathcal{R}_{\boldsymbol{I}}^{k}, \mathcal{R}'^{k}_{\boldsymbol{I}})$ is false
$\varepsilon(\boldsymbol{u})$	-	Weighting factor, represents the percentage of the receiving energy when the distant light source is partially occluded at pixel position $\boldsymbol{u}$
η	-	Shadow detection rate
λ^{GMM}	-	Matching threshold
$\mu_{\boldsymbol{I}}$	-	Mean intensity of the entire image $\boldsymbol{I}$
$\mu_{\mathcal{R}}$	-	Mean intensity of the selected region $\mathcal{R}$
μ_j^{otsu}	-	Refers to the mean used in the conventional Otsu method, where j represents the divided classes
$\mu_{i,t}^{GMM}$	-	Mean value of the i^{th} Gaussian at time t
φ_t	-	Frame attenuation value at time t
φ_T	-	Constant threshold value used to control the Frame attenuation value
ψ_t	-	Frame saturation value at time t
ψ_T	-	Constant threshold value used to control the Frame saturation value
$\varrho(\boldsymbol{u})$	-	Difference in intensity value of the pixel centered at position $\boldsymbol{u}$ with the mean value of the neighboring pixel intensity value at position $\boldsymbol{u}$
ρ^{GMM}	-	Distribution learning rate 2

$\boldsymbol{\Sigma}_{i,t}^{GMM}$	-	(Bold) Covariance matrix of the i^{th} Gaussian at time t
$\sum_{SUM}(\text{ A })$	-	Refers to the numerical count or number of pixels associated with the variable A. e.g. $\sum_{SUM}(\boldsymbol{I})$, $\sum_{SUM}(TP_{obj})$
$\sum_{SUM}(F_t)$	-	Refers to the cumulative frame count from 1 to t
σ_B^2	-	Refers to the inter-class variance used in the conventional Otsu method
$\sigma_{i,t}^{GMM}$	-	Standard deviation of the i^{th} Gaussian at time t
${\sigma^2}_{i,t}^{GMM}$	-	Variance of the i^{th} Gaussian at time t
${\sigma^2}_{init}^{GMM}$	-	Initial variance
τ^H	-	Specified threshold for the Hue component
τ^S	-	Specified threshold for the Saturation component
θ	-	Angle between the direction of the illumination source $\vec{\boldsymbol{L}}$ with surface normal $\vec{\boldsymbol{N}}$
ξ	-	Shadow discrimination rate
$\Omega_{\boldsymbol{u}}$	-	Refers to the neighboring pixels centered at pixel position $\boldsymbol{u}$
$\omega_{i,t}^{GMM}$	-	Normalized weight of the i^{th} Gaussian at time t
ω_{prior}^{GMM}	-	Prior weight
ω_j^{otsu}	-	Refers to the cumulative probability used in the conventional Otsu method, where j represents the divided classes
$\wedge$	-	Bitwise AND operation
$\oplus$	-	Bitwise XOR operation
$\forall x$	-	For all value of x

$\Leftrightarrow$	-	If and only if. e.g. given A $\Leftrightarrow$ B means A is true if B is true and A is false if B is false
$[\![\cdot]\!]$	-	Cardinal number of the image $\boldsymbol{I}$ or a region $\mathcal{R}$, e.g. $[\![\{3,5,7,9\}]\!] = 4$
$+1$	-	Increment operators to increase the counter $\#(\cdot)$ by 1
0	-	Increment operators to increase the counter $\#(\cdot)$ by 0

LIST OF ABBREVIATIONS

2D	-	Two dimensional
AND	-	Bitwise and operation
ATON	-	Autonomous Agents for On-Scene Networked Incident Management
BGS	-	Background subtraction
BGSLibrary	-	Background subtraction library
CAVIAR	-	Context Aware Vision using Image-based Active Recognition
CCM	-	Combined Color Model
CDNET	-	ChangeDetection.net
Chr	-	Chromaticity
GB	-	Gigabyte
DPGrimsonGMMBGS	-	DP Grimson GMMBGS
DPZivkovicAGMMBGS	-	DP Zivkovic AGMBGS
FFT	-	Fast Fourier Transform
Geo	-	Geometry
GMM	-	Gaussian Mixture Model
GNU GPL	-	GNU General Public License

HSV	-	Hue, Saturation, Value
i.e.	-	That is
e.g.	-	For example
i.i.d.	-	Independent and Identically Distributed
ITS	-	Intelligent Transportation System
LBMixtureOfGaussians	-	LB Mixture of Gaussians
LCC	-	Local Color Constancy
LrTex	-	Large-Region Texture
MATLAB	-	MATrix LABoratory
MFF	-	Multiple Feature Fusion
MixtureOfGaussianV1BGS	-	Mixture of Gaussian V1BGS
MixtureOfGaussianV2BGS	-	Mixture of Gaussian V2BGS
MOG	-	Mixture of Gaussian (Similar with GMM)
MRI	-	Magnetic Resonance Imaging
No.	-	Number
NCC	-	Normalized Cross-Correlation
OpenCV	-	Open source Computer Vision
p.d.f.	-	Probability Density Function
Phy	-	Physical
PM	-	Proposed method
r.v.	-	Random Variable
RAM	-	Random Access Memory
RGB	-	Red, Green, Blue
ROI	-	Region of Interest

SE	-	Structure Element used in morphological operation.
SRM	-	Segmentation Region Merging
SPD	-	Spectral Power Distribution
SrTex	-	Small-Region Texture
SZTAKI	-	SZTAKI Surveillance Benchmark
XOR	-	Bitwise exclusive or operation

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Surveillance is the monitoring of behavior, activities, or other changing information. It comes from French phrase, "*sur*" means "from above" and "*veiller*" means "to watch" (Lyon, 2001). It is used for managing, preventing and protecting the general public from danger, crisis and injury. Besides that, it has become crucial for detecting and identifying of anomalous or unusual events, especially after the rise in terrorist attacks such as the 2013 Boston Marathon bombing and 2015 Paris attacks. These attacks become the catalyst and main reason for the increasing attention in the recent years for the needs and importance of surveillance for public safety (Tian *et al.*, 2011).

A reliable automated surveillance system for security is needed. It is nearly impossible to monitor in real-time while analyzing massive information generated from various surveillance cameras installed at numerous locations. There are numerous surveillance systems have been developed such as attack detection (Raghavendra *et al.*, 2015), detecting abandoned and removed objects (Tian *et al.*, 2011), indoor surveillance (Liu *et al.*, 2015), anomaly detection (Sun *et al.*, 2017), human action recognition (Rahmani *et al.*, 2018), room occupancy detection (Candanedo and Véronique, 2016), and fall detection (Mubashir *et al.*, 2013).

A full fledge surveillance system generally comprises of three main steps: object detection, object recognition and lastly object tracking. These sequential steps rely heavily on the information of the previous step it undergoes and are interlink to one another. Hence the first step, which is the object detection phase, is crucial to have the highest possible accuracy. Detecting moving object in the scene, involves able to distinguish foreground objects from the background field regardless of the scene condition in a video sequence (Sobral and Vacavant, 2014). A false detection would have ended up in disastrous consequences for the subsequent image processing task such as object recognition. Therefore, a reliable detection of moving objects is critical for surveillance systems.

Arguably, the most common method to detect moving object in a video stream is the background subtraction (BGS) method (Sobral and Vacavant, 2014). This method does not require any prior knowledge of the moving objects in the video source (Sobral and Vacavant, 2014). The BGS process models the background image from the video source, then computes with the respective frame image to generate the foreground mask. The model background is then maintained and updated repetitively as the process continues until the video stream end. Among all BGS methods available, the Gaussian Mixture Model (GMM) by Stauffer and Grimson (1999) is the most widely used BGS method. This method utilizes multiple Gaussian as statistical means to compute the likelihood of a pixels to be classified as background or foreground pixel.

After the foreground mask is generated, it marks the beginning of the shadow detection method. For an effective shadow detection, several assumptions and shadow properties must be defined. The assumptions used are usually tailored and they are different from algorithms to algorithms. Both shadow properties and assumptions are the backbone of all the shadow detection methods.

The choice of shadow detection method highly depends on the prior assumptions and defined shadow properties defined earlier. However, the shadow detection method always has several similar traits and it can be easily noticed and identified. One of the most noticeable trait is that, the algorithm usually exploits the shadow's unique properties. For example, whenever a shadow is casted over a region, the texture property of the object remains unchanged. By knowing this property, shadows are easily identified and detected. Besides that, researchers normally take prior knowledge of the video source to their advantages, for a higher accuracy detection. Such case can be seen, for example, for a pedestrian surveillance system (Hsieh *et al.*, 2003). In these kinds of system, the orientation and position of the pedestrian are known beforehand. Hence it is much easier to detect shadow with known orientation. In some cases, the position of the illumination source is also exploited (Russell *et al.*, 2015). Knowing all the orientations beforehand is also the key in having a successful shadow detection.

1.2 Background of Study

Object detection is the most fundamental process used in image processing. It is used for detecting moving objects within an image, and the detection output is commonly used in the subsequent process such as object recognition, object tracking or object classification. Object detection has a wide range of applications from industrial automation quality control and military usage to human behavior analysis (Weinland *et al.*, 2011). Hence the detection process is important and require a high degree of accuracy for it to be useful for the subsequent process.

However, there are many challenging problems when dealing with object detection in the real-world, as commented by Sanin *et al.* (2012) and Russel *et al.* (2016). Firstly, these moving objects are usually having a wide range of texture properties and color range, hence blending the moving object into its surrounding background making it difficult to identify and model the object. Besides that, background environment conditions also influence the accuracy of the detection. Moving objects are normally found in complex and cluttered background. Furthermore, time varying changes such as changing from day to night and sudden lighting or illumination changes, drastically change the appearance of the object making the process even complex and harder for detection.

Yet, one of the common problems faced in object detection is incorrect identification and classification of shadows as part of the foreground moving object (Russel *et al.,* 2016). In a typical image, shadow and object share two important visual features. Firstly, shadow shares the same movement pattern as the moving object who casted it. The shadow is usually found adjacent to the moving object, sharing a common boundary. As for the second feature, shadow has a similar magnitude of intensity change as that of the foreground moving object (Nadimi and Bhanu, 2004). Therefore, without shadow identification process, the object detection method is likely to include the shadows as part of the object foreground. This falsely identified foreground, further reduce the accuracy of the subsequent image processing step, such as object recognition and object tracking.

1.3 Problem Statements

There are two main issues on how a falsely identified shadow as part of a moving object affects the subsequent image processing tasks as commented by Sanin *et al.* (2012). The first issue is a falsely included shadow as part of a moving object causes a change to the geometric properties of the moving object. Furthermore, the geometric properties are also affected by the duration of the day. A stronger surrounding luminance causes a shadow to be darker. The darker shadow making it as though as there is a new object presented in the video sequence. Adding the shadow along with the actual object distorts the actual object size and shape and subsequently affects the task of recognizing and classifying the moving object.

For the second issue, falsely identified shadow causes the object detection method to identify two or more moving objects as a single moving object entity. This occurs because shadow creates a false adjacency between the two moving objects which are close to one another. This false adjacency created by the shadow tricks the detection method to classify the moving objects as a single entity, instead of two different entities (Sanin *et al.,* 2012; Russel *et al.*, 2016).

Without any form of prevention on the aforementioned issues, the wrongly identified moving object further affects the subsequent image processing task such as object counting, object tracking, and object classification. Hence it is crucial for the object detection method to properly identify, segment and differentiate the moving object and its shadow (Prati *et al.*, 2003; Al-Najdawi *et al.*, 2012). Without identifying and removing shadow regions, the detected boundaries became unreliable and distorted, thus causing the surveillance system to lose track of the moving object.

Figure 1.1 illustrates what happens if the detected shadow regions are included as part of a moving object. Columns (*i*) and *(j)*, respectively illustrate tracking system detection result without and with a shadow detection method. Rows (*a*) shows the identified foreground masks. Lastly, Row (*b*) shows the tracking results with a unique ellipse boundary representing different object for both system without and with shadow detection method.

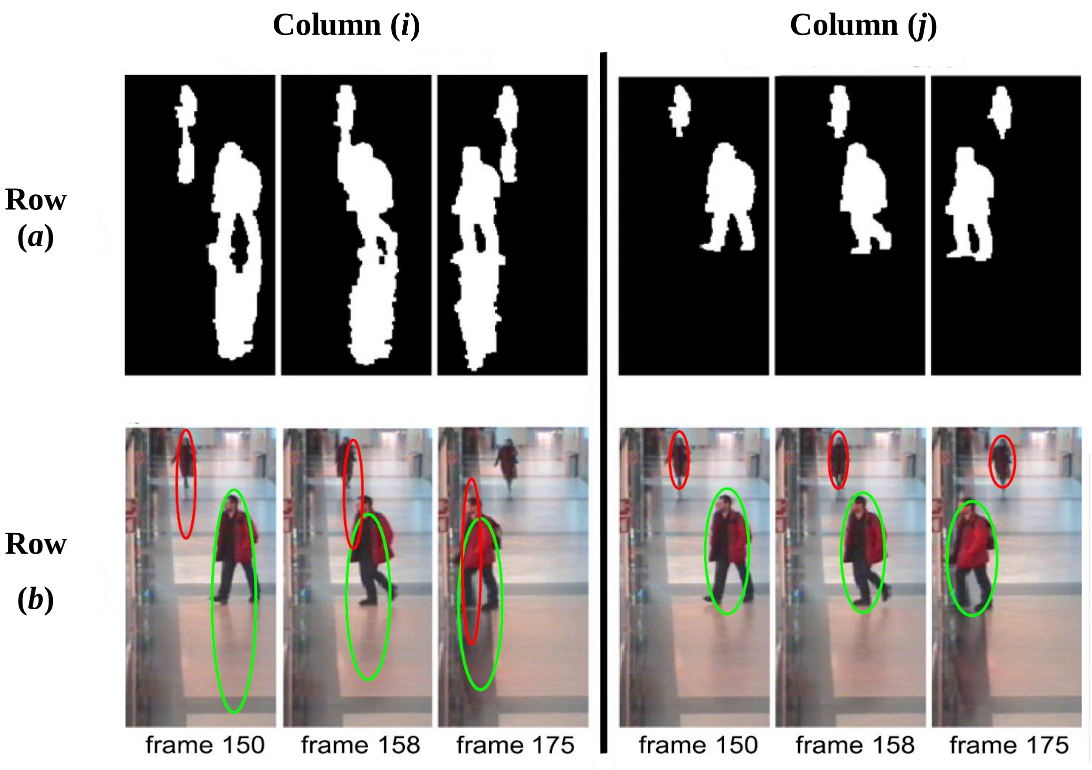

Figure 1.1: Illustrates a case where a tracking system tracks ongoing pedestrians, without and with shadow detection.

Apart from that, shadows can be formed under various illumination conditions such as under poor illumination lighting environment and even at night time. Therefore, a surveillance system has to be able to cater for various illumination changes. In the current shadow detection methods, even though the researchers have catered for various illumination conditions by evaluating their detection method in both indoor and outdoor scene, however their methods are still unable to accurately detect and eliminate shadow especially in the scenes with strong illuminations are present (Russel *et al.*, 2016). Besides that, most environment scenes have a single dominant illumination source. However, there are some environment scenes, such as in an indoor environment has multiple dominant illumination sources. In a typical indoor environment, the illumination sources are surrounded in a closed environment. Such situation leads to the creation of shadow by the reflected light off the wall or the floor. The reflected light from those surfaces are normally weak in comparison with the illumination source, and the amount of reflected light depends on the reflective index of the surface.

Lastly, image noise is also one of the issues that hinders the performance of the shadow detection. Image noise is random variation of brightness or color information in images, which can be found in most digital video recorder. Besides that, it causes the colors of any given scene do not remain for a long period of time due to the present of camera noise and illumination fluctuations (Horprasert *et al.*, 1999).

Therefore, incorrectly identify shadow as a part of a moving object resulting in producing incorrect detection. Hence there is a need for improving the accuracy of the shadow detection method especially for tackling environment scene that have image noise and illumination changes.

1.4 Objective of Research

The objectives of the research are defined as follows:

(i) To develop a multiplicative fusion shadow detection method that utilizes both texture and Chromaticity features together with image segmentation for surveillance images to tackle image noise and illumination changes.

(ii) To evaluate the accuracy of the developed method, by benchmarking it several commonly used shadow detection methods.

1.5 Scope of Research

Research scope is an agreement on the research work to be completed, also justifying the objectives of the research and setting deliverables and limitations of the research. The main focus of this research is to develop a highly accurate shadow eliminator method for a surveillance system, tackling unresolved shadow issues. The summary of the scope of research can be found in Table 1.1.

This developed method is limited to using Gaussian Mixture Model (GMM) background subtraction method by Stauffer and Grimson (1999). Details on the parameter setup are discussed in Chapter 3. Besides that, the list of commonly used shadow detection methods used for benchmarking are prepared by Sanin *et al.*, (2013). The list includes a wide range of shadow detection methods from Chromaticity based detection to texture based detections. In this research, the additional multiple feature fusion is added for a better benchmarking result.

For a fair evaluation, a standard video dataset is utilized. The video dataset used is the Agents for On-Scene Networked Incident Management (ATON), prepared and maintained by Martel-Brisson and Zaccarin, (2008) and Sanin *et al.*, (2012). Each video in the video dataset consist of several challenges in order to evaluate the robustness of a shadow detection method. The challenges tackled in this research is limited to the handling image noise, ability to cope with various illumination changes, flexibility of detection under various background scene and the ability to handle scene with multiple illumination sources. Details on the properties of the video sequence dataset and its associated challenges are presented comprehensively in Chapter 3.

Next, two evaluation tests are performed on the video dataset used to determine the accuracy the shadow detection. The tests include the shadow detection and discrimination test (Prati *et al.*, 2003) and the desaturation test (Sanin *et al.*, 2012). The shadow detection and discrimination test evaluate the percentage of accuracy for both object and shadow identified by the shadow detection method. Lastly, the desaturation test was carried out to test the performance capability of the algorithm when colour information is reduced.

Lastly, the developed shadow detection method is coded in C++ programming language, utilizing the open source computer vision (OpenCV) library version 2.4.10. It is developed using Microsoft Visual Studio 2013 as integrated development environment.

1.6 Significances of Research

The significances and findings of the research can be highlighted in several aspects. The most important aspect of this research is to further improve the accuracy of other image processing systems especially in terms of security and public safety. This research can be coupled with other image processing tasks such as human action recognition (Rahmani *et al.*, 2018), abandoned object detection (Tian *et al.*, 2011) and anomaly detection (Sun *et al.*, 2017) to be installed in places such as at banks and transit terminals. Thus, further improving the accuracy is useful for security purposes.

Table 1.1: Research scope for this research.

Description	**Scope**
Background subtraction technique	GMM background subtraction method by Stauffer and Grimson (1999)
Video dataset	ATON video dataset, prepared and updated by: • Martel-Brisson and Zaccarin (2008), and • Sanin *et al.*, (2012).
Video source challenges	• Handle image noise • Illumination changes • Multiple illumination source • Various background scene • Execution time
Benchmarked shadow detection method, prepared by Sanin *et al.*, (2012)	• Chromaticity, by Cucchiara *et al.* (2003) • Geometry, by Hsieh *et al.* (2003) • Physical, by Huang and Chen (2009) • Large-Region Texture, by Sanin *et al.* (2010) • Small-Region Texture, by Leone and Distante (2007) • Combined Color Model, by Sun and Li (2010) • Multiple Feature Fusion, by Dai *et al.* (2013)
Evaluating the performance of the shadow detection method	• Shadow detection and discrimination test, by Prati *et al.* (2003) • Desaturation test, by Sanin *et al.* (2012)

1.7 Thesis Outline

The remaining of this thesis is structured in the following manner. Literature reviews are presented in Chapter 2. Firstly, discussion on the shadow feature definition are presented then followed by reviews on background subtraction. After that, the four individual shadow detection technique classification, which are the Chromaticity, Geometry, Physical and Texture are analyzed in detail. Next, segmentation based shadow detection methods and combination of multiple techniques are analyzed and reviewed. Lastly, critical reviews on the identified shadow detection methods are analyzed and compared.

The first part in Chapter 3 describes the overview of the research methodology, then followed by introducing the video sequence dataset with detail explanations on the challenges provided by each video in the dataset. Next, the overview of the shadow eliminator process is discussed and each of the method used is explained in details. This includes background subtraction, image segmentation, individual shadow identifying methods and the developed combination technique. Lastly, evaluation techniques are introduced to quantify the shadow elimination method in terms of detection accuracy.

The results of the research are presented and discussed in Chapter 4. First the challenges for each of the video for benchmarking are neatly explained. Next the quantitative analysis for benchmarking are shown and explained. Finally, the results for this research are shown and compared with other commonly used shadow detection methods. A comprehensive discussion on this research in comparison with other research is provided.

Lastly, Chapter 5 summarizes the entire research. The research work is concluded, and several recommendations are emphasized for future work.

REFERENCES

Al-Najdawi, N., Bez, H.E., Singhai, J. and Edirisinghe, E.A. (2012). A Survey of Cast Shadow Detection method. *Pattern Recognition Letters,* 33(6), 752-764.

Amato, A., Mozerov, M.G., Bagdanov, A.D., and Gonzàlez, J. (2011). Accurate Moving Cast Shadow Suppression Based on Local Color Constancy Detection. *IEEE Transactions on Image Processing*, 20(10), 2954-2966.

Atev, S., Masoud, O., and Papanikolopoulos, N. (2004). Practical Mixtures of Gaussians with Brightness Monitoring. *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems*. 3-6 October, Washington, WA, USA.

Benedek, C. and Szirányi, T. (2007). Study on Color Space Selection for Detecting Cast Shadows in Video Surveillance. *International Journal of Imaging Systems and Technology*, 17(3), 190-201.

Benedek, C.S. and Szirányi, T. (2008). Bayesian Foreground and Shadow Detection in Uncertain Frame Rate Surveillance Videos. *IEEE Transactions on Image Processing*, 17(4), 608-621.

Bi, S., Liang, D., Shen, X. and Wang, Q. (2007). Human Cast Shadow Elimination Method Based on Orientation Information Measures. *Proceedings of IEEE International Conference on Automation and Logistics*, 18-21 August, Jinan, China, 1567–1571.

Bouwmans, T., El Baf, F. and Vachon, B. (2008). Background Modeling using Mixture of Gaussians for Foreground Detection - A Survey. *Recent Patents Computer Science*, 1(3), 219–237.

Bullkich, E., Ilan, I., Moshe, Y., Hel-Or, Y. and Hel-Or, H. (2012). Moving Shadow Detection by Nonlinear Tone-Mapping. *19th International Conference on Systems, Signals and Image Processing (IWSSIP)*, 11-13 April, Vienna, Austria, 146–149

Candanedo, L.M. and Feldheim, V. (2016). Accurate Occupancy Detection of an Office Room from Light, Temperature, Humidity and CO2 Measurements using Statistical Learning Models. *Energy and Buildings*, 112(1), 28-39.

Chen, C.C. and Aggarwal, J.K. (2010). Human Shadow Removal with Unknown Light Source. *Proceedings of the 20th International Conference on Pattern Recognition (ICPR)*, 23-26 August, Istanbul, Turkey, 2407–2410.

Chen, C.T., Su, C.Y. and Kao, W.C. (2010). An Enhanced Segmentation on Vision-based Shadow Removal for Vehicle Detection. *1st International Conference on Green Circuits and Systems (ICGCS 2010)*, 21-23 June, Shanghai, China, 679-682.

Cho, J.H., Kwon, T.G., Jang, D.G. and Hwang, C.S. (2005). Moving Cast Shadow Detection and Removal for Visual Traffic Surveillance. *18th Australian Joint Conference on Artificial Intelligence (AI 2005): Advance in Artificial Intelligence*, 5-9 December, Sydney, Australia, 746–755.

Choi, J.M., Chang, H.J., Yoo, Y.J., and Choi, J.Y. (2012). Robust Moving Object Detection Against Fast Illumination Change. *Computer Vision and Image Understanding*, 116 (2012) 179–193.

Cooley, J.W. and Tukey, J.W. (1965). An Algorithm for the Machine Calculation of Complex Fourier Series. *Mathematic of Computing*. 19(90), 297-301.

Cucchiara, R., Grana, C., Piccardi, M. and Prati, A. (2003). Detecting Objects, Shadows and Ghosts in Video Streams by Exploiting Color and Motion Information. *Proceedings of the 11th International Conference on Image Analysis and Processing*, 26-28 September, Palermo, Italy, 360-365.

Cucchiara, R., Grana, C., Piccardi, M. and Prati, A. (2003). Detecting Moving Objects, Ghosts, and Shadows in Video Streams. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10), 1337-1342.

Dai, J., Qi, M., Wang, J., Dai, J. and Kong, J. (2013). Robust and accurate Moving Shadow Detection Based on Multiple Feature Fusion. *Optics and Laser Technology*, 54(1), 232-241.

Duda, R.O. and Hart, P.E. (1973). Pattern Classification and Scene Analysis. *New York: Wiley*.

Elboher, E. and Werman, M. (2013). Asymmetric Correlation: A Noise Robust Similarity Measure for Template Matching. *IEEE Transactions on Image Processing*. 22(8), 3062-3072.

Elgammal, A., Duraiswami, R., Harwood, D. and Davis, L.S. (2002). Background and Foreground Modeling using Nonparametric Kernel Density Estimation for Visual Surveillance. *Proceedings of the IEEE*, 90(7), 1151–1163.

Emre Celebi, M., Kingravi, H.A., Iyatomi, H., Alp Aslandogan, Y., Stoecker, W.V., Moss, R.H., and Joseph M. (2008). Border Detection in Dermoscopy Images using Statistical Region Merging. *Skin Research and Technology*, 14(3), August, 347-353.

Fang, L.Z., Qiong, W.Y. and Sheng Y.Z. (2008). A Method to Segment Moving Vehicle Cast Shadow based on Wavelet Transform. *Pattern Recognition Letters*. 29(16), 2182-2188.

Finlayson, G.D., Drew, M.S. and Lu, C. (2009). Entropy Minimization for Shadow Removal. *International Journal of Computer Vision*, 85(1), 35–57.

Fung, G.S.K., Yung, N.H.C., Pang, G.K.H. and Lai, A.H.S. (2001). Effective Moving Cast Shadows Detection for Monocular Color Image Sequences. *Proceedings of the 11th International Conference on Image Analysis and Processing*, 26-28 September, Palermo, Italy, 404-409.

Gao, J., Tian, Y. and Liu, Y. (2009). Ellipsoidal Method for Shadow Detection Based on Normalized RGB Values. *International Conference on Networking and Digital Society* (*ICNDS '09)*, 30-31 May, Guiyang, Guizhou, 104-107.

Glaister, J., Alexander W. and David A.C. (2014). Segmentation of Skin Lesions from Digital Images using Joint Statistical Texture Distinctiveness. *IEEE Transactions on Biomedical Engineering*, 61(4) June, 1220-1230.

Gomes, V., Barcellos, P. and Scharcanski, J. (2016). Stochastic Shadow Detection using a Hypergraph Partitioning Approach. *Pattern Recognition*, September, 63, 30-44.

Goyette, N., Jodoin, P.M., Porikli, F., Konrad, J. and Ishwar, P. (2012). Changedetection.net: A New Change Detection Benchmark Dataset. *Proceedings of the IEEE Workshop on Change Detection (CDW-2012) CVPR-2012*, Providence, RI, 16-21 June.

Grest, D., Frahm, J.M. and Koch, R. (2003). A Color Similarity Measure for Robust Shadow Removal in Real Time. *Proceedings of the Vison, Modeling, and Visualization Conference (VMV 2003)*, 19-21 November, München, Germany, 253–260.

Guan, Y.P., (2010). Spatio-temporal Motion-Based Foreground Segmentation and Shadow Suppression. *IET Computer Vision*, 4(1), 50–60.

Guo, R., Dai, Q. and Hoiem, D. (2011). Single-Image Shadow Detection and Removal using Paired Regions. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2011)*, 20-25 June, Providence, RI, 2033-2040.

Haines, T.S. and Tao, X. (2014). Background Subtraction with Dirichletprocess Mixture Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(4), December, 670-683.

Hamad, A.M. and Tsumura, N. (2012). Background Updating and Shadow Detection based on Spatial, Colour, and Texture Information of Detected Objects. *Optical Review*, 19(3), 182–197.

Hel-Or, Y., Hel-Or, H. and David, E. (2011). Fast Template Matching in Non-Linear Tone-Mapped Images. *IEEE International Conference on Computer Vision (ICCV 2011)*, 6-13 November, Barcelona, Spain, 1355-1362.

Herodotou, N., Plataniotis, K.N. and Venetsanopoulos, A.N. (1998). A Color Segmentation Scheme for Object-Based Video Coding. *Proceedings of the IEEE Symposium in Advances in Digital Filtering and Signal Processing*, 5-6 June, Victoria, BC, Canada, 25-29.

Horprasert, T., Harwood, D. and Davis, L. (1999). A Statistical Approach for Real-Time Robust Background Subtraction and Shadow Detection. *IEEE ICCV'99 Frame-Rate Workshop*, 99(1999), 1-19.

Hsieh, J.W., Hu, W.F., Chang, C.J. and Chen, Y.S. (2003). Shadow Elimination for Effective Moving Object Detection by Gaussian Shadow Modeling. *Image and Vision Computing*. 21(6), 505-516.

Huang, D.Y. and Wang, C.H. (2009). Optimal Multi-Level Thresholding using a Two-Stage Otsu Optimization Approach. *Pattern Recognition Letters*, 30(3), 275–284.

Huang, J.B. and Chen, C.S. (2009). Moving Cast Shadow Detection using Physical based Features. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009)*, 20-25 June, Miami, FL, USA, 2310-2317.

Jiang, K., Li, A.H., Cui, Z.G., Wang, T. and Su, Y.Z. (2013). Adaptive Shadow Detection using Global Texture and Sampling Deduction. *IET Computer Vision*, 7(2), 115-122.

Joshi, A.J. and Papanikolopoulos, N.P. (2008). Learning to Detect Moving Shadows in Dynamic Environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(11), 2055–2063.

Kaewtrakulpong, P. and Bowden, R. (2001). An Improved Adaptive Background Mixture Model for Real-Time Tracking with Shadow Detection. *European Workshop on Advanced Video Based Surveillance Systems (AVSS)*, 1, 135-144.

Lalonde, J.F., Efros, A.A. and Narasimhan, S.G. (2010). Detecting Ground Shadows in Outdoor Consumer Photographs. *European Conference on Computer Vision*, ECCV 2010, 5-11 September, Heraklion, Crete, Greece, 322-335.

Lang, F., Yang, J., Zhao, L. and Li, D. (2012) Hierarchical Classification of Polarimetric SAR Image based on Statistical Region Merging. *Proceeding. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Melbourne, VIC, Australia, 25 August – 1 September, 147–152.

Lee, D.S. (2005). Effective Gaussian Mixture Learning for Video Background Subtraction. *IEEE Transection on Pattern Analysis and Machine Intelligent*, 27(5), May, 827–832.

Leone, A. and Distante, C. (2007). Shadow Detection for Moving Objects based on Texture Analysis. *Pattern Recognition*, 40(4), 1222-1233.

Lewis, J.P. (1995). Fast Template Matching. *Vision Interface 95 Canadian Image Processing and Pattern Recognition Society*. 15-19 May, Quebec City, Canada, 120-123.

Linda G.S. and George C.S. (2001). Computer Vision. *Prentice-Hall*: New Jersey. 279-325.

Ling, Z., Lu, X., Wang, Y. and He, X. (2013). Adaptive Moving Cast Shadow Detection by Integrating Multiple Cues. *Chinese Journal of Electronics*, 22(4), October, 757-762.

Liu, L., Zhang, W., Deng, C., Yin, S. and Wei, S. (2015). BriGuard: A Lightweight Indoor Intrusion Detection System based on Infrared Light Spot Displacement. *IET Science, Measurement & Technology*. 9(3), 306-314.

Liu, Z., Huang, K. and Tan, T. (2012). Cast Shadow Removal in a Hierarchical Manner using MRF. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(1), 56-66.

Liu, Z., Huang, K., Tan, T. and Wang, L. (2007). Cast Shadow Removal Combining Local and Global Features. *IEEE Conference on Computer Vision and Pattern Recognition*, *(CVPR '07)*, 17-22 June, Minneapolis, MN, USA, 1-8.

Lyon, D. (2001). Surveillance Society. Buckingham: *Open University Press*.

Mahmood, A. and Sohaib, K. (2010). Exploiting Transitivity of Correlation for Fast Template Matching. *IEEE Transactions on Image Processing*, 19(8), 2190-2200.

Martel-Brisson, N. and Zaccarin, A. (2008). Kernel-based Learning of Cast Shadows from a Physical Model of Light Sources and Surfaces for Low-Level Segmentation. *26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008)*, 23-28 Jun., Anchorage, AK, 1-8.

Martins, I., Carvalho, P., Corte-Real, L., & Alba-Castro, J.L. (2018). BMOG: Boosted Gaussian Mixture Model with Controlled Complexity for Background Subtraction. *Pattern Analysis and Applications*, 1, April, 1-14.

McDiarmid, C. (1998). Concentration. In *Probabilistic Methods for Algorithmic Discrete Mathematic*s. Springer Berlin Heidelberg, 195-248.

McFeely, R., Glavin, M. and Jones, E. (2012). Shadow Identification for Digital Imagery using Colour and Texture Cues. *IET Image Processing*, 6(2), 148–159.

Meher, S.K. and Murty, M.N. (2013). Efficient Method of Moving Shadow Detection and Vehicle Classification. *AEU - International Journal of Electronics and Communications*, 67(8), 665–670.

Mubashir, M., Shao, L., and Seed, L. (2013). A Survey on Fall Detection: Principles and Approaches. *Neurocomputing*, 100(1), 144-152.

Nadimi, S. and Bhanu, B. (2004). Physical Models for Moving Shadow and Object Detection in Video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(8), 1079-1087.

Nicolas, H. and Pinel, J.M. (2006). Joint Moving Cast Shadows Segmentation and Light Source Detection in Video Sequences. *Signal Processing: Image Communication*, 21(1), 22–43.

Nielsen, F. and Nock, R. (2003). On Region Merging: The Statistical Soundness of Fast Sorting, with Applications. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.* 18-20 June, Madison, WI, USA, II-19-26.

Nock, R. (2001). Fast and Reliable Color Region Merging Inspired by Decision Tree Pruning. *Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR, 2001)*, Kauai, HI, USA, 271-276.

Nock, R. and Nielsen, F. (2004). Statistical Region Merging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(11), 1452-1458.

Nock, R. and Nielsen, F. (2005). Semi-Supervised Statistical Region Refinement for Color Image Segmentation. *Pattern Recognition*, 38(1), 835-846.

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62-66.

Peng, B, Lei Z. and David Z. (2011). Automatic Image Segmentation by Dynamic Region Merging. *IEEE Transactions on Image Processing*, 20(12), May, 3592-3605.

Phong, B. (1975). Illumination for Computer Generated Pictures. *Communications of the ACM*. 18(6), June, 311-317.

Porikli, F. and Thornton, J. (2005). Shadow Flow: A Recursive Method to Learn Moving Cast Shadows. *10th IEEE International Conference on Computer Vision (ICCV'05),* 1(1), 17-21 October, Beijing China, 891-898.

Power, P.W. and Schoonees, J.A. (2002). Understanding Background Mixture Models for Foreground Segmentation. *Proceedings Image and Vision Computing New Zealand 2002*, 26–28 November, University of Auckland, Auckland, New Zealand, 10-11.

Prati, A., Mikic, I., Trivedi, M.M. and Cucchiara, R. (2003). Detecting Moving Shadows: Algorithms and Evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 25(7), 918-923.

Qin R., Liao S., Lei Z., and Li S. (2010). Moving Cast Shadow Removal Based on Local Descriptors. 2*0th International Conference on Pattern Recognition (ICPR 2010)*, 23-26 August, Istanbul, Turkey, 1377–1380.

Rafael C. Gonzalez, R.C. and Woods, R.E. (2002). Digital Image Processing (2nd Edition). *Prentice Hall*: New Jersey.

Raghavendra, R., Raja, K.B. and Busch, C. (2015). Presentation Attack Detection for Face Recognition using Light Field Camera. *IEEE Transactions on Image Processing,* 24(3), 1060-1075.

Rahmani, H., Mian, A., and Shah, M. (2018). Learning a Deep Model for Human Action Recognition from Novel Viewpoints. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 40(3), 667-681.

Russell, A. and Zou, J.J. (2013). Moving Shadow Detection on Spatial-Temporal Constancy. *7th International Conference on Signal Processing and Communication Systems (ICSPCS),* 16-18 January, Carrara, Vic, Australia, 1-6.

Russell, M., Zou, J.J. and Fang, G. (2015). Real-Time Vehicle Shadow Detection. *IEEE Electronic Letters*, 51(16), 1253-1255.

Salvador, E., Cavallaro, A. and Ebrahimi, T. (2004). Cast Shadow Segmentation Using Invariant Color Features. *Computer Vision and Image Understanding*, 95(2), 238–259.

Sanin, A., Sanderson, C. and Lovell, B.C. (2012). Shadow Detection: A Survey and Comparative Evaluation of Recent Methods. *Pattern Recognition*, 45(4), 1684-1695.

Sanin, A.C. and Lovell, B.C. (2010). Improved Shadow Removal for Robust Person Tracking in Surveillance Scenarios. *20th International Conference on Pattern Recognition (ICPR)*, 23-26 August, Istanbul, Turkey, 141-144.

Schindler, K. and David S. (2008). Object Detection by Global Contour Shape. *Pattern Recognition*, 41(12), December, 3736-3748.

ShabaniNia, E. and Naghsh-Nilchi, A.R. (2013). Robust Watershed Segmentation of Moving Shadows using Wavelets. *8th Iranian Conference on Machine Vision and Image Processing (MVIP)*, 10-12 September, Zanjan, Iran, 381-386.

Shan, Y. and Wang, R. (2005). Improved Algorithms for Motion Detection and Tracking. *Optical Engineering*, 45(6), 067201-067201.

Shi, L, Lefei, Z., Jie, Y., Liangpei Z. and Pingxiang, L. (2013). Supervised Graph Embedding for Polarimetric SAR Image Classification. *IEEE Geoscience and Remote Sensing Letters*, 10(2), June, 216-220.

Shoaib, M., Dragon, R., and Ostermann, J. (2009). Shadow Detection for Moving Humans using Gradient-based Background Subtraction. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009)*, 19-24 April, Taipei, 773-776.

Sobral, A. and Vacavant, A. (2014). A Comprehensive Review of Background Subtraction Algorithms Evaluated with Synthetic and Real Videos. *Computer Vision and Image Understanding*, 122(1), 4-21.

Song, K.T. and Tai, J.C. (2007). Image-based Traffic Monitoring with Shadow Suppression. *Proceedings of the IEEE*, 95(2), 413–426.

Stauder, J., Mech, R. and Ostermann, J. (1999). Detection of Moving Cast Shadows for Object Segmentation. *IEEE Transactions on Multimedia*. 1(1), 65-76.

Stauffer, C. and Grimson, W.E.L. (2000). Learning Patterns of Activity using Real-Time Tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), 2246–2252.

Su, Y., Li, A., Cai, Y., Feng, G., and Jin, G. (2014). Moving Shadow Detection with Multifeature Joint Histogram. *Journal Electronic Imaging*, 23(5), October, 053-065.

Sun, B. and Li, S. (2010). Moving Cast Shadow Detection of Vehicle Using Combined Color Models. *Proceeding of the Chinese Conference on Pattern Recognition (CCPR 2010)*, 21-23 October, Chongqing, China, 503-507.

Sun, Q., Liu, H., and Harada, T. (2017). Online Growing Neural Gas for Anomaly Detection in Changing Surveillance Scenes. *Pattern Recognition*, 64(1), 187-201.

Tang, C., Ahmad, M.O., and Wang, C. (2013). An Efficient Method of Cast Shadow Removal Using Multiple Features. *Signal Image and Video Processing*. 695-703.

Tian, B., Morris, B.T., Tang, M., Liu, Y., Yao, Y., Gou, C., Shen, D. and Tang, S. (2015). Hierarchical and Networked Vehicle Surveillance in ITS: A Survey. *IEEE Transactions on Intelligent Transportation Systems*, 16(2), 557-580.

Tian, Y.L., Feris, R.S., Liu, H., Hampapur, A. and Sun, M.T. (2011). Robust Detection of Abandoned and Removed Objects in Complex Surveillance Videos. *IEEE Transactions on Systems, Man, and Cybernetics*, 41(5), 565-577.

Toth, D., Stuke, I., Wagner, A. and Aach, T. (2004). Detection of Moving Shadows using Mean Shift Clustering and A Significance Test. *Proceedings of the 17th International Conference on Pattern Recognition*, 4(1), 260–263.

Toyama, K., and Krumm, J., Brumitt, B. and Meyers, B. (1999). Wallflower: Principles and Practice of Background Maintenance. *7th IEEE International Conference on Computer Vision (ICCV)*, Corfu, Greece, 255–261.

Tsai, D.M. and Lin, C.T. (2003). Fast Normalized Cross Correlation for Defect Detection. *Pattern Recognition Letters*. 24(15), 2625-2631.

Tsai, V.J.D. (2006). A Comparative Study on Shadow Compensation of Color Aerial Images in Invariant Color Models. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6), 1661-1671.

Wang, H., Li, A., Cui, Z., Fang, H., and Shen, D. (2016). Moving Shadow Detection based on Multi-Feature Fusion. *2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 27-28 Aug., Hangzhou, 138-141.

Wang, J.M., Chung, Y.C., Chang, C.L. and Chen, S.W. (2004). Shadow Detection and Removal for Traffic Images. *Proceedings of IEEE International Conference on Networking, Sensing and Control*, 21-23 March, Taipei, Taiwan, 1, 649–654.

Wang, K., Liu, Y., Gou, C. and Wang, F.Y. (2016). A Multi-View Learning Approach to Foreground Detection for Traffic Surveillance Applications. *IEEE Transactions on Vehicular Technology*, 65(6), 4144-4159.

Weinland, D., Ronfard, R. and Boyer, E. (2011). A Survey of Vision-based Methods for Action Representation, Segmentation and Recognition. *Computer Vision and Image Understanding*, 115(2), 224-241.

Wu, C., Jiemi, Z., Ozan, S., Bart, S., Silvio, S. and Ashutosh, S. (2018). Watch-n-patch: Unsupervised Learning of Actions and Relations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(2), March, 467-481.

Xiang, J., Fan, H., Liao, H., Xu, J., Sun, W. and Yu, S. (2014). Moving Object Detection and Shadow Removing under Changing Illumination Condition. *Mathematical Problems in Engineering*, February 2014, 1 – 10.

Xiang, T. and Gong, S. (2008). Video Behaviour Profiling for Anomaly Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5), 893-908.

Xiao, M., Han, C.Z. and Zhang L. (2005). Moving shadow Detection and Removal for Traffic Sequences. *International Journal of Automation and Computing*, 4(1), 38-46.

Yang, M.T., Lo, K.H., Chiang, C.C. and Tai, W.K. (2008). Moving Cast Shadow Detection by Exploiting Multiple Cues. *IET Image Processing*, 2(2), April, 95-104.

Yoneyama, A., Yeh, C.H. and Kuo, C.C.J. (2005). Robust Vehicle and Traffic Information Extraction for Highway Surveillance. *EURASIP Journal on Applied Signal Processing*, 14(1), 2305–2321.

Zang, Q. and Klette, R. (2006). Parameter Analysis for Mixture of Gaussians Model. *CITR*, The University of Auckland, New Zealand.

Zhang, W., Fang, X.Z., Yang, X.K. and Wu, Q.M.J. (2007). Moving Cast Shadow Detection using Ratio Edge. *IEEE Transactions on Multimedia*, 9(6), 1202-1213.

Zhu, J., Samuel K.G.G., Masood S.Z. and Tappen, M.F. (2010). Learning to Recognize Shadows in Monochromatic Natural Images. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 13-18 June, San Francisco, CA, 223-230.

Zivkovic, Z. and Heijden, F.V.D. (2006). Efficient Adaptive Density Estimation Per Image Pixel for the Task of Background Subtraction. *Pattern Recognition Letters*, 27(7), 773-780.