MOVING VEHICLES DETECTION AND TRACKING APPROACH VIA MODIFIED BACKGROUND SUBTRACTION AND INNOVATIVE ADAPTIVE SEARCH WINDOW

RAAD AHMED HADI

A thesis submitted in fulfilment of the requirements for the award of the degree of Doctor of Philosophy (Computer Science)

Faculty of Computing
Universiti Teknologi Malaysia

OCTOBER 2015
To my lovely parents and my beloved brothers.
ACKNOWLEDGEMENT

First and foremost, “Praise be to Allah, the cherisher and the sustainer of the world”, “praise be to Him He who taught by the pen, taught man, that which he did not know”

First of all, my sincere thanks to God, who endowed me to complete this PhD thesis. I would like to thank my supervisor, Prof. Dr. Ghazali Bin Sulong, for all your guidance, support, brilliant ideas, numerous hours of discussions, patience, and the opportunities you have presented me. Your managerial skills and uncompromising quest for excellence always motivated me to present the best of what I can.

I would like to express my heartfelt gratitude to my supervisor Assoc. Prof. Dr. Siti Zaiton Bt. Mohd Hashim. She has been helpful, understanding and generous throughout the study. She has truly been a mentor and I owe here my deepest thanks.

I would like to say special thanks my external supervisor Prof. Dr. Loay Edwar George, whose precious guidance, support and encouragement were pivotal establishing my self-confidence in this endeavour.

And most importantly, I would like to express my deep gratitude and thank to my beloved family for their love, patience and support, especially my mother, my father and my dearest brothers for their continuous support and supplication. In addition, I would like to thank my friends in Iraq for their support and prayers. Finally, I have made many friends during my time at UTM and I would like to thank all my friends and colleagues for their support and help.
ABSTRACT

Efficient and accurate video-based vehicles supervision systems for traffic surveillance are ever-demanding. Vehicle tracking is an action of detecting and maintaining the correspondence between moving vehicles instances across every frame of video for proper management of highway traffic. However, some effects such as change of weather conditions, tree leaves waving and light fluctuations that hamper detection accuracy of vehicles need to be addressed. Developing an accurate detection and tracking algorithm by incorporating the effects of shadowing of vehicles or obstruction/occlusion by visual obstacles including road signs, trees or other vehicles is a challenging task. These effects severely influence the detection and tracking performance of moving vehicles and cause gradual deterioration towards efficient and accurate traffic image analysis. Superior highway traffic surveillance systems with enhanced performance in terms of both detection accuracy and processing time are far from being achieved. This thesis proposes a novel approach by integrating modified background subtraction method with new adaptive search window for precise detection and tracking of highway traffic. Modified background subtraction detects and highlights the motion of moving object via morphological operations to refine the foreground binary image. Hence, the proposed new adaptive search window reduces searching space to predict and allocate moving vehicles positions across video frames. Furthermore, chromatic based analysis detects the shadow of moving vehicle and the partial occlusion between vehicles by estimating vehicle sizes, which is used as criteria to separate them during the moving vehicles tracking. The performance of the proposed approach is benchmarked using standard highway dataset of ChangeDetection.NET (CDNET) having 1231 frames. An accuracy rate as much as 99% and processing time of 29 millisecond per frame were achieved. The proposed approach provides significant improvement on the performance of highway traffic surveillance compared to previous works. A time complexity degree of $O(n^2)$ is demonstrated as the running time of algorithm analysis.
ABSTRAK

# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DECLARATION</td>
<td>ii</td>
</tr>
<tr>
<td></td>
<td>DEDICATION</td>
<td>iii</td>
</tr>
<tr>
<td></td>
<td>ACKNOWLEDGMENT</td>
<td>iv</td>
</tr>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td></td>
<td>ABSTRAK</td>
<td>vi</td>
</tr>
<tr>
<td></td>
<td>TABLE OF CONTENTS</td>
<td>vii</td>
</tr>
<tr>
<td></td>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td></td>
<td>LIST OF ABBREVIATIONS</td>
<td>xvi</td>
</tr>
<tr>
<td></td>
<td>LIST OF APPENDICES</td>
<td>xviii</td>
</tr>
</tbody>
</table>

1 INTRODUCTION 1

1.1 Introduction 1
1.2 Problem Background 3
1.3 Problem Statement 10
1.4 Research Questions 11
1.5 Research Goal and Objectives 12
1.6 Research Scope 12
1.7 Research Significance 13
1.8 Thesis Organization 13
1.9 Summary 14
2 LITERATURE REVIEW

2.1 Introduction 15

2.2 Moving Vehicle Detection and Segmentation Approaches 17
  2.2.1 Background Subtraction Method 18
  2.2.2 Feature Based Methods 25
  2.2.3 Frame Differencing and Motion Based Methods 29

2.3 Camera Calibration Approaches 31

2.4 Shadow Detection Approaches 33

2.5 Occlusion Handling Approaches 35

2.6 Vehicle Tracking Approaches 37
  2.6.1 Region-Based Tracking Method 38
  2.6.2 Contour Tracking Method 39
  2.6.3 3D Model-Based Tracking Method 41
  2.6.4 Feature-Based Tracking Method 41
  2.6.5 Color and Pattern-Based Method 42

2.7 Other Related Approaches for Moving Vehicles Detection and Tracking 44

2.8 Trends and Directions 47

2.9 Summary 49

3 RESEARCH METHODOLOGY

3.1 Introduction 50

3.2 Research Framework 51

3.3 Performance Evaluation 55
  3.3.1 Detection Accuracy 55
  3.3.2 Processing Time 57
  3.3.3 Time Complexity 57

3.4 Dataset 59

3.5 Experimental Setup 61

3.6 Research Planning and Implementation Map 62

3.7 Summary 65
## 4 DESIGN AND IMPLEMENTATION OF THE PROPOSED APPROACH

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>67</td>
</tr>
<tr>
<td>4.2</td>
<td>Pre-processing Stage</td>
<td>67</td>
</tr>
<tr>
<td>4.3</td>
<td>Gray-scale Conversion</td>
<td>72</td>
</tr>
<tr>
<td>4.4</td>
<td>Modified Background Subtraction Method for Vehicle Detection</td>
<td>74</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Foreground Detection</td>
<td>77</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Morphological Binary Operations</td>
<td>80</td>
</tr>
<tr>
<td>4.4.2.1</td>
<td>Dilation and Erosion Operations</td>
<td>82</td>
</tr>
<tr>
<td>4.5</td>
<td>Vehicle Counting and Labeling Technique</td>
<td>85</td>
</tr>
<tr>
<td>4.6</td>
<td>Vehicle Region Segmenting</td>
<td>93</td>
</tr>
<tr>
<td>4.7</td>
<td>Shadow Detection</td>
<td>99</td>
</tr>
<tr>
<td>4.7.1</td>
<td>Shadow Searching Window</td>
<td>103</td>
</tr>
<tr>
<td>4.7.2</td>
<td>Shadow Region Segregation</td>
<td>104</td>
</tr>
<tr>
<td>4.8</td>
<td>Occlusion Handling</td>
<td>107</td>
</tr>
<tr>
<td>4.9</td>
<td>Innovative Adaptive Search Window Method for Vehicle Tracking</td>
<td>114</td>
</tr>
<tr>
<td>4.10</td>
<td>Summary</td>
<td>124</td>
</tr>
</tbody>
</table>

## 5 RESULTS AND DISCUSSION

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>125</td>
</tr>
<tr>
<td>5.2</td>
<td>Results of Pre-processing Stage</td>
<td>126</td>
</tr>
<tr>
<td>5.3</td>
<td>Results of Gray-Scale Conversion</td>
<td>131</td>
</tr>
<tr>
<td>5.4</td>
<td>Results of Modified Background Subtraction Method for Vehicle Detection</td>
<td>132</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Foreground Detection Analysis</td>
<td>132</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Result of Dilation and Erosion Operations</td>
<td>137</td>
</tr>
<tr>
<td>5.5</td>
<td>Result for Vehicle Counting and Labeling</td>
<td>139</td>
</tr>
<tr>
<td>5.6</td>
<td>Results of Vehicle Region Segmentation</td>
<td>142</td>
</tr>
<tr>
<td>5.7</td>
<td>Shadow Detection Analysis</td>
<td>145</td>
</tr>
<tr>
<td>5.8</td>
<td>Results for Occlusion Handling</td>
<td>148</td>
</tr>
</tbody>
</table>
5.9 Results of Innovative Adaptive Search Window Method for Vehicle Tracking 156

5.10 Evaluation of The Proposed Approach 162

5.10.1 Evaluation of Detection Accuracy 162

5.10.2 Processing Time 168

5.10.3 Evaluation of Time Complexity 170

5.11 Summary 174

6 CONCLUSION AND FUTURE WORK 176

6.1 Overview 176

6.2 Thesis Summary 177

6.3 Contributions 179

6.4 Future Work 180

REFERENCES 182

Appendices A – C 201-215
<table>
<thead>
<tr>
<th>TABLE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Performance comparison among the on-road technologies</td>
<td>16</td>
</tr>
<tr>
<td>2.2</td>
<td>Summary of related works</td>
<td>48</td>
</tr>
<tr>
<td>3.1</td>
<td>An overview of research planning and implementation</td>
<td>62</td>
</tr>
<tr>
<td>5.1</td>
<td>Set of parameters representing the vehicle’s sizes with its bounding box coordinates for the CDNET dataset</td>
<td>150</td>
</tr>
<tr>
<td>5.2</td>
<td>Set of parameters indicating vehicle’s sizes with its bounding box coordinates for the second data set</td>
<td>151</td>
</tr>
<tr>
<td>5.3</td>
<td>Results of four measurements of the proposed detection method</td>
<td>162</td>
</tr>
<tr>
<td>5.4</td>
<td>A comparison between the proposed detection method and previous works for the four measurements</td>
<td>164</td>
</tr>
<tr>
<td>5.5</td>
<td>Results for the performance metrics of the proposed detection method</td>
<td>166</td>
</tr>
<tr>
<td>5.6</td>
<td>A comparison of detection accuracy performance of the proposed detection method with previous works</td>
<td>166</td>
</tr>
<tr>
<td>5.7</td>
<td>Computational time cost analysis</td>
<td>168</td>
</tr>
<tr>
<td>5.8</td>
<td>Time complexity of the proposed approach for six elements</td>
<td>174</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>FIGURE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>TSS applications</td>
<td>2</td>
</tr>
<tr>
<td>2.1</td>
<td>(a) A simple image (b) the modeling neuronal map</td>
<td>21</td>
</tr>
<tr>
<td>2.2</td>
<td>Revealing the underneath shadow (a) the measurement area and the lowest horizontal line (b) A region of vehicles and their shadows</td>
<td>22</td>
</tr>
<tr>
<td>2.3</td>
<td>Results of filtering swinging trees with (a) Original image; (b) absolute background of (a); (c) difference frame between (a) and (b); (d) results of motion compensation of (c); (e) results of filtering swinging trees</td>
<td>23</td>
</tr>
<tr>
<td>2.4</td>
<td>Detection and classification result With (a) Original image (b) the detected vehicles (c) the image generated via morphological operations</td>
<td>24</td>
</tr>
<tr>
<td>2.5</td>
<td>Discriminator description</td>
<td>24</td>
</tr>
<tr>
<td>2.6</td>
<td>PCA weight-vector distributions</td>
<td>27</td>
</tr>
<tr>
<td>2.7</td>
<td>Eigen-window method</td>
<td>28</td>
</tr>
<tr>
<td>2.8</td>
<td>The vehicle detection and tracking for nighttime traffic surveillance</td>
<td>30</td>
</tr>
<tr>
<td>2.9</td>
<td>Presence of common geometric primitives in the traffic scenes</td>
<td>31</td>
</tr>
<tr>
<td>2.10</td>
<td>The camera setting of automatic traffic system via 2D spatio-temporal images</td>
<td>32</td>
</tr>
<tr>
<td>2.11</td>
<td>The experimental result: (a) and (e) are the original frames; (b) and (f) are the results of using another method;</td>
<td>34</td>
</tr>
</tbody>
</table>
(c) and (g) are the detected shadows by using the proposed method; (d) and (h) are the final results of the proposed method

2.12 Intra-frame level occlusion handling. (a) separated occluded vehicles using “cutting line.” (b) convex image difference of occluded vehicles. (c) convex hull (depicted in red) of CIDmax (in green). (d) selected “cutting point,” and (e) Possible “cutting lines” (in magenta) and the selected “cutting line” (in green)

35

2.13 Inter-frame level occlusion handling. (a) two occluded vehicles (b) binary mask of occluded vehicles (c) motion vectors of occluded vehicles (d) results with the intersecting edge depicted in green

36

2.14 An illustration of partial occlusion handling

37

2.15 Moving regions detection

39

2.16 (a) Comparison between two motorcycles (bold white contour) and cars (white contour) with (b) the original figure (c) comparison between a bus (bold white contour) and cars (white contour) with (d) its original figure

40

2.17 Tracking Vehicle with mean-shift

43

3.1 An overview of the proposed approach

53

3.2 A block diagram displaying the pre-processing steps

54

3.3 Samples of the CDNET dataset

60

3.4 Samples of the another dataset for additional testing

61

4.1 The process for background image construction

69

4.2 Flowchart for gray-scale conversion

74

4.3 Foreground detection operation steps

78

4.4 A 3x3 square structuring element displaying eight-pixel connectivity

81

4.5 The dilation and erosion operations on foreground binary mask

83

4.6 SPST circuit design

88

4.7 Design of 2D model to count and label the vehicles

89
4.8 Vehicle counting steps using 2D model: (a) vehicle entering the 2D model (b) newly numbering the vehicle (c) counting is progressing (d) vehicle departing the 2D model

4.9 Region growing segmentation procedure (a) beginning of the scanning process (b) collection of white pixels in the first region (c) collection of white pixels in the second region

4.10 Construction of shadow

4.11 Types of shadow

4.12 Schematics of shadow detection module

4.13 Features of shadow searching window

4.14 An illustration of occlusion handling steps

4.15 Estimation of the increasing proportion of vehicle size

4.16 Different steps of ASW performance

4.17 An illustration of ASW method

4.18 Implementation on non-occluded and occluded vehicles tracking

5.1 Histogram of video frames showing the most frequent pixel color values

5.2 Illustration of histogram for newly constructed background image

5.3 Background image construction of the CDNET dataset

5.4 Background image construction of the second dataset

5.5 Conversion of RGB to gray-scale image

5.6 Image subtraction process the CDNET dataset

5.7 Image subtraction process the second dataset

5.8 Histogram and the threshold of the difference image for the CDNET data set

5.9 Histogram and the threshold of the difference image for the second data set

5.10 The result of dilation process

5.11 The result of erosion process
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.12</td>
<td>The result of vehicle counting for the CDNET dataset</td>
</tr>
<tr>
<td>5.13</td>
<td>The result of vehicle counting for the second dataset</td>
</tr>
<tr>
<td>5.14</td>
<td>The result of uncounted moving vehicles in 2D model</td>
</tr>
<tr>
<td>5.15</td>
<td>Region growing algorithm generated segmentation from CDNET dataset</td>
</tr>
<tr>
<td>5.16</td>
<td>Region growing algorithm generated segmentation from second dataset</td>
</tr>
<tr>
<td>5.17</td>
<td>A step by step illustration of the shadow detection process: (a) original Image; (b) bounding box area (SSW); (c) shadow identification; (d) shadow segregation</td>
</tr>
<tr>
<td>5.18</td>
<td>Shadow detection in the CDNET dataset</td>
</tr>
<tr>
<td>5.19</td>
<td>Shadow detection in the second dataset</td>
</tr>
<tr>
<td>5.20</td>
<td>The partial occlusion occurrence in the CDNET dataset</td>
</tr>
<tr>
<td>5.21</td>
<td>The partial occlusion occurrence in the second dataset</td>
</tr>
<tr>
<td>5.22</td>
<td>The result of ASW size varying in the CDNET dataset</td>
</tr>
<tr>
<td>5.23</td>
<td>The result of variable ASW size for the second dataset</td>
</tr>
<tr>
<td>5.24</td>
<td>The result of ASW size fixing in the CDNET dataset</td>
</tr>
<tr>
<td>5.25</td>
<td>Multi-vehicle tracking process for the CDNET dataset: (a) without partial occlusion handling and shadow detecting (b) with partial occlusion handling and shadow detecting</td>
</tr>
<tr>
<td>5.26</td>
<td>Multi-vehicle tracking processes for the second dataset (a) without partial occlusion handling and shadow detection (b) with partial occlusion handling and shadow detection</td>
</tr>
<tr>
<td>5.27</td>
<td>Typical sampling of the proposed detection method’s FG image</td>
</tr>
<tr>
<td>5.28</td>
<td>A comparison between the proposed detection method and previous works for the four measurements</td>
</tr>
<tr>
<td>5.29</td>
<td>A comparison of detection accuracy of the proposed detection method with other previous works</td>
</tr>
<tr>
<td>5.30</td>
<td>Processing time required by the proposed approach for various processes</td>
</tr>
</tbody>
</table>
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASW</td>
<td>Adaptive Search Window</td>
</tr>
<tr>
<td>BDWT</td>
<td>Binary Discrete Transforms</td>
</tr>
<tr>
<td>BNECC</td>
<td>Bayesian Network Enhanced Cascades Classification</td>
</tr>
<tr>
<td>BS</td>
<td>Background Subtraction</td>
</tr>
<tr>
<td>CDNET</td>
<td>ChangeDetection.Net</td>
</tr>
<tr>
<td>CG</td>
<td>Computer Graphics</td>
</tr>
<tr>
<td>CoBE</td>
<td>Cascade of Boosted Ensembles</td>
</tr>
<tr>
<td>CPU</td>
<td>Central Processing Unit</td>
</tr>
<tr>
<td>DDMCMC</td>
<td>Data Driven Markov Chain Monte Carlo</td>
</tr>
<tr>
<td>DWT</td>
<td>Discrete Wavelet Transform</td>
</tr>
<tr>
<td>EM</td>
<td>Expectation Maximization</td>
</tr>
<tr>
<td>EPI</td>
<td>EPI-Polar Plane</td>
</tr>
<tr>
<td>FG</td>
<td>Foreground Binary Mask</td>
</tr>
<tr>
<td>FN</td>
<td>False Negative</td>
</tr>
<tr>
<td>FP</td>
<td>False Positive</td>
</tr>
<tr>
<td>FV</td>
<td>Feature Vector</td>
</tr>
<tr>
<td>HGV</td>
<td>Heavy Goods Vehicle</td>
</tr>
<tr>
<td>HOS</td>
<td>High Order Statistics</td>
</tr>
<tr>
<td>ICA</td>
<td>Independent Component Analysis</td>
</tr>
<tr>
<td>LBSP</td>
<td>Local Binary Similarity Patterns</td>
</tr>
<tr>
<td>MAP</td>
<td>Maximizing a Posterior Probability</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov Chain Monte Carlo</td>
</tr>
<tr>
<td>MTP</td>
<td>Multimodal Temporal Panorama</td>
</tr>
<tr>
<td>MRFs</td>
<td>Markov Random Fields</td>
</tr>
<tr>
<td>NN</td>
<td>Neural Network</td>
</tr>
<tr>
<td>OIRDS</td>
<td>Overhead Imaging Research Dataset</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>PLP</td>
<td>Plump Line Projection</td>
</tr>
<tr>
<td>RAM</td>
<td>Random Access Memory</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of Interest</td>
</tr>
<tr>
<td>SIFT</td>
<td>Scale Invariant Feature Transform</td>
</tr>
<tr>
<td>SPST</td>
<td>Single Pole Single Throw</td>
</tr>
<tr>
<td>SSW</td>
<td>Shadow Search Window</td>
</tr>
<tr>
<td>SURF</td>
<td>Speed-Up Robust Features</td>
</tr>
<tr>
<td>SUV</td>
<td>Sport Utility Vehicle</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>SW</td>
<td>Searching Window</td>
</tr>
<tr>
<td>TN</td>
<td>True Negative</td>
</tr>
<tr>
<td>TPs</td>
<td>Toggle Points</td>
</tr>
<tr>
<td>TP</td>
<td>True Positive</td>
</tr>
<tr>
<td>TSS</td>
<td>Traffic Surveillance System</td>
</tr>
</tbody>
</table>
## LIST OF APPENDICES

<table>
<thead>
<tr>
<th>APPENDIX</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Samples of data set’s original and ground truth images</td>
<td>201</td>
</tr>
<tr>
<td>B</td>
<td>Samples of results for the CDNET dataset</td>
<td>206</td>
</tr>
<tr>
<td>C</td>
<td>Samples of results for the second dataset</td>
<td>211</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Introduction

In this chapter, the problem statement, research questions, objectives, scope of studies and research significance are highlighted. The traffic surveillance system (TSS) is known to be one of the most significant tools for video-based supervision systems. It is a machine-based vision organization system for detecting, recognizing, classifying and tracking moving objects/vehicles on the road. The system is comprised of a static positioned camera setup that captures traffic image in a sequence of computer vision frames and perform image processing operations including detection, tracking, etc. TSS provides and facilitates zone selection via camera’s view to detect vehicles within the specified zone. It collects image signals of vehicles and send them to a television monitor as video signals containing traffic data sequence (Hatae et al., 2002; Mayeaux and Stamps, 1992).

Over the years, dedicated efforts are made to design TSS for transportation planning and architecture. Many traffic engineering applications are developed to extract useful and precise passage information for traffic image analysis and flow control. These include vehicle detection, vehicle counting, vehicle trajectory
prediction, vehicle tracking, vehicle flow, vehicle classification, traffic density assessment, vehicle velocity measurement, traffic lane changes observation and license plate recognition as shown in Figure 1.1, to cite a few (Chung-Lin and Wen-Chieh, 2004; Hsieh et al., 2006; Kanhere, 2008; Kanhere and Birchfield, 2008; Wei et al., 2008). Despite many efforts an efficient and precise TSS is far from being achieved.

Vehicle image may differ in size, shape, colour and can change depending on its position, the presence of neighbourhood objects, the cluttered background and the varying outdoor scene especially the lighting conditions (Gupte et al., 2002; Zheng et al., 2006). Currently, the existing vehicle detection and tracking systems are limited by their poor image analysis and obscure recognition, which effect on the performance of TSS in terms of accuracy of detection and processing time (CPU time) (Jodoin et al., 2012; Maddalena and Petrosino, 2012; St-Charles et al., 2014). These shortcomings originate from the vehicles obstruction by other vehicles or visual obstacles such as road signs, trees, and weather conditions etc. Consequently, the performance accuracy of TSS for detection, classification and tracking of the vehicles is decided by its reliable and precise traffic image analysis (Feng and Shuguang, 2009; Jodoin, et al., 2012; Shih-Shinh et al., 2007).

**Figure 1.1** TSS applications
1.2 Problem Background

With rapid pace in advanced technology, the cost of video recording cameras and digital media storage has increased and became affordable. In recent years, there is a remarkable increase in recording video data in the world. In order to analysis and process video data, there is a growing demand in analysis and management this huge video data. One of video contents understanding and analysis process is visual object tracking, it is represented an essence step of any video surveillance system. Visual object tracking is the process of finding, localizing and dynamic configuration of the position of moving objects in every frame of video sequences (Alsaqre and Yuan, 2004; Jalal and Singh, 2012; Ramasubramanian et al., 2014).

Video surveillance system has a broad range of applications to detect and track the moving objects such as airplanes, micro-organisms (red or white cells), vehicles, animals, etc. Traffic Surveillance System (TSS) is one of hot topics of video surveillance that is interested the researchers. TSS is the process of extracting the motion of moving vehicles and analysis the dynamic behaviors of these vehicles across every scene of video sequences (Hu et al., 2013; Jalal and Singh, 2012). In general, TSS is object tracking process which consists of detection, segmentation and modeling of interesting moving vehicle, predicting the motion and position of candidate vehicle in each frame, and localization the vehicles by using the similarity measure in feature space of interesting vehicle (Jalal and Singh, 2012; (Ramasubramanian et al., 2014).

However, the TSS is a challenge task because of the change of object appearance by complex real environment changes between scenes like illumination changes, motion variations like camera motion (jitter), abrupt object motion, shadows and reflections, noise in frame of video sequences and occlusion. These issues have significant effects on the performance of detection and tracking of moving vehicles (Bin et al., 2014; Bing-Fei et al., 2013). Various proposed methods
have distinguished based on the manner to deal with these issues it will mentioned later.

The detection of moving vehicles represents the primary important step in object tracking system. It can improve the efficiency and reliability of the system by increasing the detection accuracy, reducing the search space and processing time requirements for extracting the moving objects/vehicles (Jodoin, et al., 2012). In fact, there is real-world environments variations such as motion variations (like tree branches and leaves waving) and illumination changes (weather conditions, light fluctuations and shadow) leads to misdetection (false detection) and losing of tracked foreground objects (vehicles). As a result, the moving vehicles detection and tracking process become more difficult and complex (Goyette et al., 2012; Kastrinaki et al., 2003; Porikli and Yilmaz, 2012; Shih-Shinh, et al., 2007; Tanaka et al., 2010; Zehang et al., 2006). Many detection methods are summarized below with its pros and cons such as thresholding, frame differencing, Background Subtraction (BS) and etc.

In the past, despite low accuracy and simple to implement thresholding was used as a part of the very first automatic observation systems (Mahmassani et al., 1999; Sahoo et al., 1988). Correspondingly, Vehicle edge detection approach based on sobel operator was proposed (Weihua, 2009). The main advantage of the edge-based method is scale invariant, but they suffer from noise, inaccurate and it can’t deal with congested scenes (D. Koller et al., 1994; Weihua, 2009; Xiao Xu and Grimson, 2005).

Also, temporal or frame differencing is a method based on the orientations and speed of their movements from the background of the motion scene image sequence (Han et al., 2007; Dieter Koller et al., 1994; Z. Wei, et al., 2008; Zhang et al., 2010). It is can detect accurately the objects region of interest without requiring a background image. However, this approach cannot suitably handle realistic traffic situations when there is no vehicle movement in the image for a period of time.
Multi-resolution is a power tool used in the detection of objects which is based on scale space theory (Cheng et al., 2006; Freixenet et al., 2002; Hilario et al., 2005; Lindeberg, 1996). Nevertheless, this method is not accurate enough for traffic data problems because it is unable to deal with image perspective. Meanwhile, lane marks used to analyze and calibrate the road region coordinates depending on road features extraction and it used lane geometrical models to detect vehicles. This method possesses highly complex computational procedures and frequently affected by shadows and high contrast variations (Lim et al., 2009).

Background Subtraction (BS) is one of the most widely used method in vehicle regions detection and very well-suited in detecting foreground objects (Region of Interest (ROI)), tracing, and classifying processes for any vehicle supervision system (Ambardekar, 2007; Bing-Fei et al., 2014; Huwer and Niemann, 2000; Li et al., 2003). However, this method suffers from false detection which occurs due to the motion variations such as camera movement (jitter) and tree branches or leaves waving (Jodoin, et al., 2012; Tavakkoli et al., 2006). This often leads to inaccurate identification of foreground objects from the background image. The non-adaptive characteristic in the background subtraction method is another drawback. This arises due to the changes in the lighting and the climate situations which necessitates continuous updating (Al-garni and Abdennour, 2006; De Gregorio and Giordano, 2013; Jodoin, et al., 2012). In addition, the BS suffer from high computational time cost due to false detection and non-adaptive characteristic issues. Overcoming these problems require more computational efforts that gradually deteriorates the performance effectiveness of the supervision system in terms of detection accuracy and processing time factors (Feng and Shuguang, 2009; Mandellos et al., 2011; Tanaka et al., 2010; Xiaofeng et al., 2014).

In addition, the presence of shadow and occlusion in static images and frames of video sequences causes severe problems and affect the performance of many of computer vision applications including object detection, image segmentation, tracking, etc. However, it is very difficult to differentiate between real object and shadow which is caused due to complex interactions such as the variability of
moving objects in real outdoor daylight scenes, and non-stationary of background (weather conditions, albedo, illumination variations) (Hsieh et al., 2003; Jia-Bin and Chu-Song, 2009; Joshi and Papanikolopoulos, 2008). In fact, the existence of shadow forms and appears like another object related with interesting object (vehicle) which lead for detection errors and may be caused to lose the real target during the tracking.

Correspondingly, many new feature-based methods were proposed to detect accurately and efficiently the shadow using object’s feature (Sanin et al., 2012). The intensity of illumination features were used to detect the shadow (Zhang Wei et al., 2006). The intensity features were needed to enhance the shadow detection by combining it with another feature like color space of moving object caused by illumination changes.

The geometric features such as orientation, size, and shape were used to detect moving shadow without analyzing background image (Fang et al., 2008; Nicolas and Pinel, 2006; Yoneyama et al., 2003). This type of shadow detection methods were not well designed for accurate shade recognition from multiple objects.

Also, the color information (Chromaticity) of moving object such as HUV, LUV, RGB color space, etc are simple to implement and computationally inexpensive to detect the shadow (Cavallaro et al., 2005; Chun-Ting et al., 2010; Cucchiara et al., 2003; Salvador et al., 2004). Another efficient shadow detection method based on the textures features of objects was developed (Leone and Distante, 2007; Yang et al., 2006). This method being independent of object color is robust to illumination fluctuations, but computationally inefficient.

Undoubtedly, the object detection and tracking in the presence occlusion is a challenging task. These occlusions may originate from the overlapping, touching or
hiding by other objects. For example, recognizing a vehicle which is partially hidden by fluctuation of leaves (Huang et al., 2012). In most cases, the partially occluded vehicle means the way in which some parts of the vehicle are hidden by other objects including vehicle, road sign, tree, etc. In the last few years, several attempts are made to resolve the partial occlusion problem and to achieve enhanced detection and tracking performance. The feature points such as corners, eigenvalue, intensity variance, gradient of vehicles were used for handling the occluded vehicles (Goo et al., 2008; Wei et al., 2011). Despite computational complexity, this method handled the partial occlusion efficiently. However, it was affected by unrelated structures of objects in complex scenes which lead to loss of feature points.

In addition, the 3D world box models were used (Kanhere and Birchfield, 2008; Yung and Lai, 1998) to detect the feature points (similar to feature points method in its work which mentioned in previous section) of vehicles which were gathered together to segment the occluded vehicles. The 3D box models accurately handled the partial occlusion in spite of its computational expenses related to large number of varying vehicles shapes models and camera calibration calculation and pose.

The object tracking is an action for detecting and maintaining the correspondence between moving vehicles (objects) instances through every frame of video sequence (Jalal and Singh, 2012). The tracking systems are consists of the motion and the matching parts. The motion process is to predict the position/location of a moving object being tracked in the next frame by defining a limited search region or space in which the moving object is likely to be found. Conversely, the matching process is connected to the location or detection of the moving object within designated search region that is identified in the next frame (Lu et al., 2014; Trucco and Plakas, 2006).
Generally, the positioning and matching of each vehicle across every frame improves the robustness of tracking performance and minimizes the search space. Furthermore, it can speed up the performance of whole processes for Traffic Surveillance System (TSS) (Jalal and Singh, 2012; Shih-Shinh, et al., 2007). In fact, the real-world happenings degrade the performance of tracking systems such as sudden object motion, changing of object pattern's appearance (shapes), shadows, tree branches or leaves waving and occlusions and etc. Accordingly, many efforts have been suggested to cope with these problems.

The Mean-shift is an efficient and accurate method to locate and match the moving object by using color histogram. However, this method has drawbacks, it is sensitive to illumination variations, incapable in handling the partial or full occlusion problem and the multi-target tracking facility (Fernando and Cooray, 2013; Jalal and Singh, 2012; Ming et al., 2012; Wei et al., 2014; Xia et al., 2009).

The other tracking approach called Kalman filter is widely explored in computer vision. It is an efficient linear predictive method against noise during the tracking process (Djalalov et al., 2010; Welch and Bishop, 1995). There are major disadvantageous of this method, it uses single hypothesis (moving object) and simultaneous multi-hypothesis (multi-target) employment is infeasible. Besides, the manual initialization of moving object state variables is required as a previous knowledge for tracking process and cannot deal with too intricate motion patterns (Pan et al., 2006; Zhaoxiang et al., 2010).

The most popular non-linear probabilistic algorithm for multi-hypothesis tracking is Particle filter. (Bardet et al., 2009; Maggio and Cavallaro, 2010; Subhash Challa and Evans, 2011). Despite its good performance in tracking field it suffers from high computational complexity. The other problem is related to the on-line tracking which gets deteriorated due to the inefficiency in sampling process (degeneracy and loss of diversity problem) (Bardet, et al., 2009; Jalal and Singh, 2012; Mihaylova et al., 2007).
A time-frequency-transform technique called Wavelet transform such as Discrete Wavelet Transform (DWT) has gained wide popularity in object tracking. It deals with illumination and object appearance variations, handles the image noise and camera jittering. In spite of wavelet efficiency this method is computationally expensive, unable to handle the object occlusion and suffers from shift-sensitivity problems (Chang et al., 2007; Fang et al., 2008; Jalal and Singh, 2012; Nishidha and Janardhanan, 2013; Selesnick et al., 2005). Lately, template matching method is a robust tracking method which is based on predefined models for matching the moving objects across the frames of video. However, it suffered from computational burdens (Mandellos et al., 2011).

Various methods are proposed for TSS to achieve superior detection accuracy and less computational time cost for vehicle detection and tracking systems. Despite all these efforts which are previously mentioned, several issues are far from being solved. In this view, the present study intends to overcome the problem associated with the false detection by developing and implementing an efficient improved Background Subtraction (BS) method in terms of accuracy of detection. Moreover, a new method is proposed by combining BS with Adaptive Search Window (ASW) to reduce the search space cost (processing time) for efficient tracking moving vehicle across consecutive frames. Yet, an attempt is taken to resolve the shadow and partial occlusion related issues of moving vehicle in terms of computational time cost. Those can be accurately detected and handled via the newly introduced enhanced technique. Finally, this innovative approach is expected to be highly useful for reducing the search space cost, counting and labelling moving vehicles on traffic highways.
1.3 Problem Statement

In the information technology era, the digital media storage and video camera technologies are in rapid progression in computing sphere. In recent years, huge traffic video data are captured and stored. These are remarkably growing worldwide. However, the requirement of processing, analyses and thorough understanding of these traffic video data led the computer vision society to create state-of-the-art algorithms. Ironically, TSS being the most attractive and active area in computer vision applications processes massive traffic video data contents. In fact, some following notable problems need to be solved prior to traffic supervision applications:

1. Real-world environments changes such as tree branches and leaves waving have declined and gradually deteriorated the performance of the detecting and tracking of moving vehicles applications. Despite great effort by earlier researches focused on enhancing and increasing the performance of traffic surveillance systems in terms *accuracy of detection*, but the enhancement of accuracy of detection stills wide open (De Gregorio and Giordano, 2013; Jodoin, *et al.*, 2012; Maddalena and Petrosino, 2012; St-Charles, *et al.*, 2014).

2. One of the important methods in moving objects detection and tracking is the BS. Despite its efficiency in detecting the moving object the time requirement in tracking is too long. This is due to searching steps on finding moving objects pixel regions through the whole image (search space) over time in every frame from video sequence, which lead to computational processing time consuming (Jodoin, *et al.*, 2012; Maddalena and Petrosino, 2012; Mandellos, *et al.*, 2011; Salvi, 2012). In fact, improvement of the tracking process can speed up the performance of TSS in terms of *processing time* (Shih-Shinh, *et al.*, 2007). Additionally, the variable-size search window surrounding the moving vehicles (bounding box) is required for more accurate tracking. However, any method containing fixed-size search window often affects the robustness of tracking process (Fernando and Cooray, 2013).
3. The shadow and partial occlusion are important factors that need to be tackled elegantly. Otherwise, their presence inhibits efficient analysis of the traffic data. In fact, the shadow and partial occlusion issues influence the detection accuracy by missing the real target during the tracking process. Accordingly, these problems influence, complicate and degrade the performance of moving vehicle detection and tracking processes over time (Jalal and Singh, 2012; Meher and Murty, 2013; F. Wei, et al., 2011).

4. Only a few existing techniques provide the information of the computational time cost which is necessary in real time TSS applications (Maddalena and Petrosino, 2012; Mandellos, et al., 2011; Shih-Shinh, et al., 2007; Shimada et al., 2013; Tanaka, et al., 2010). Thus, it is fundamental to trade-off between accuracy of detection and computational time cost of any method for moving vehicles detecting and tracking.

1.4 Research Questions

The following research questions related to moving vehicles detection and tracking are framed in this thesis:

i. How to develop a new efficient approach for moving vehicles detection and tracking by fulfilling the requirements of high accuracy and computationally economic?

ii. How to combine the ASW method with BS to achieve better detection, labelling, and counting performance in reducing the searching space during the tracking process?

iii. How to detect and handle the shadow and partial occlusion of moving vehicles?
iv. Does the new approach achieve high detection accuracy, less time complexity and computational efficiency that determine effective tracking, labelling and counting of moving vehicles on highway traffic?

1.5 Research Goal and Objectives

The goal of this research is to develop an efficient and accurate method for detecting and tracking moving vehicles in the presence of the shadow and partial occlusion. The newly proposed technique is capable of creating a real time image sequence with superior accuracy and less time complexity.

The objectives of this research are the following:

1) To modify a background subtraction (binary morphological operations) method for accurate detection of vehicles.
2) To enhance shadow detection and introduce new partial occlusion methods between moving vehicles with low computational time cost.
3) To propose an adaptive search window (flexible window) method for efficient tracking of vehicles.

1.6 Research Scope

To achieve the goals in fulfilling the proposed objectives for accurate moving vehicles detection and efficient tracking the following scopes of the research are emerged:
1. Color traffic video standard dataset with RGB video images (320 x 240) is acquired from ChangeDetection.NET (CDNET) http://www.changedetection.net/ and gray-scale AVI video.
2. Fixed position camera and front view of moving vehicles are used.
3. This study focuses on moving vehicles detection and tracking in the presence of shadow and partial occlusion matters.
4. Detection accuracy and processing time of the proposed method will be used to evaluate and benchmark the performance efficiency of the newly proposed method for highway traffic detection and tracking.

1.7 Research Significance

The capability of accurate detection and tracking of moving vehicles over time in a complex environment is a challenging and tough task. Several earlier researches are unable to achieve highly accurate detection system with low computation time for traffic surveillance application. The proposed detection and tracking approach of moving vehicles on highway traffic is expected to achieve enhanced accurate detection with efficient minimal computational time cost and less time complexity. Furthermore, the new approach will be able to detect the shadow and handle the partial occlusion between moving vehicles that will precisely count and label the moving vehicles.

1.8 Thesis Organization

This thesis is comprised of five chapters. Chapter 1 provides a brief introduction of the study to set the research rationale, overviews some previous
relevant works concerning moving objects detection and tracking for TSS applications. In addition, the problem statement, objectives, scope of the research together with the significance are discussed. The detailed literature review relevant for the cited research topic is underscored in Chapter 2. The state-of-the-art methods, techniques, and approaches in the area of traffic surveillance systems (TSS) are systematically emphasized. An overview of research methodology is explained in Chapter 3, this chapter is clearly examined the principles and procedures of the research. Chapter 4 deals with detailed design and implementation of the proposed methodology for the achievements of research goal. Here, different methods and techniques related to BS and ASW are explained to demonstrate their effective employment in the present work. Experimental results in terms of testing, evaluation, and benchmarking are presented, analyzed, and discussed in Chapter 5. Chapter 6 concludes the thesis with future directions in this newly evolving and significant research domain.

1.9 Summary

This chapter described the basis of the undertaken research work. It briefly examined the current problems or limitations connected to the existing traffic surveillance system (TSS) and its applications. The strength and weakness of Background Subtraction (BS) method for moving vehicle detection and tracking is discussed. The advantages and disadvantageous of the existing systems are emphasized. Accordingly, the problem background, research question, research objectives, research scope, significance of the study are justified and highlighted. It argued why the research in the field of moving vehicle detection and tracking is essential. The organization of the thesis is presented.
REFERENCES


Bing-Fei, W., Chih-Chung, K., Cheng-Lung, J., Yen-Feng, L., Ying-Han, C. and Jhy-Hong, J. (2014). A Relative-Discriminative-Histogram-of-Oriented-
Gradients-Based Particle Filter Approach to Vehicle Occlusion Handling and Tracking. *IEEE Transactions on Industrial Electronics.* 61(8), 4228-4237. doi: 10.1109/TIE.2013.2284131


Han, B., Roberts, W., Wu, D. and Li, J. (2007). Motion-Segmentation-Based Change
Detection. SPIE Defence & Security Symposium 2007, 65680Q-65680Q.


Vasu, L. (2010). *An effective step to real-time implementation of accident detection system using image processing*. Master of Science, Oklahoma State University, USA.


