UNIVERSITI PUTRA MALAYSIA

MOVING OBJECTS DETECTION FROM UAV CAPTURED VIDEOS USING TRAJECTORIES OF MATCHED REGIONAL ADJACENCY GRAPHS

BAHAREH KALANTAR GHORASHI HARANDI

FK 2018 24

MOVING OBJECTS DETECTION FROM UAV CAPTURED VIDEOS USING TRAJECTORIES OF MATCHED REGIONAL ADJACENCY GRAPHS

By

BAHAREH KALANTAR GHORASHI HARANDI

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfillment of the Requirements for the Degree of Doctor of Philosophy

December 2017

DEDICATION

To my loving parents

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

MOVING OBJECTS DETECTION FROM UAV CAPTURED VIDEOS USING TRAJECTORIES OF MATCHED REGIONAL ADJACENCY GRAPHS

By

BAHAREH KALANTAR GHORASHI HARANDI

December 2017

Chairman : Professor Shattri Bin Mansor, PhD
Faculty : Engineering

Videos captured using cameras from unmanned aerial vehicles (UAV) normally produce dynamic footage that commonly contains unstable camera motion with multiple moving objects. These objects are sometimes occluded by vegetation or even other objects, which presents a challenging environment for higher level video processing and analysis. This thesis deals with the topic of moving object detection (MOD) whose intention is to identify and detect single or multiple moving objects from video. In the past, MOD was mainly tackled using image registration, which discovers correspondences between consecutive frames using pair-wise grayscale spatial visual appearance matching under rigid and affine transformations. However, traditional image registration is unsuitable for UAV captured videos since distance-based grayscale similarity fails to cater for the dynamic spatio-temporal differences of moving objects. Registration is also ineffective when dealing with object occlusion. This thesis therefore proposes a framework to address these issues through a two-step approach involving region matching and region labeling. Specifically, the objectives of this thesis are (i) to develop an image registration technique based on multigraph matching, (ii) to detect occluded objects through exploration of candidate object correspondences in longer frame sequences, and (iii) to develop a robust graph coloring algorithm for multiple moving object detection under different transformations.

In general, each frame of the footage will firstly be segmented into superpixel regions where appearance and geometrical features are calculated. Trajectory information is also considered across multiple frames taking into account many types of transformations. Specifically, each frame is modeled/represented as a regional adjacency graph (RAG). Then, instead of pair-wise spatial matching as with image registration, correspondences between video frames are discovered through multigraph matching of robust spatio-temporal features of each region. Since more than two frames are considered at one time, this step is able to discover better region correspondences as well as caters for object(s) occlusion. The second step of region labeling relies on the assumption that background and foreground moving objects exhibit different motions properties when in a sequence. Therefore, their spatial difference is expected to drastically differ over time. Banking on this, region labeling assigns the labels of either background or foreground region based on a proposed graph coloring algorithm, which considers trajectory-based features. Overall, the framework consisting of these two steps is termed as Motion Differences of Matched Region-based Features (MDMRBF). MDMRBF has been evaluated against two datasets namely the (i) Defense Advanced Research Projects Agency (DARPA) Video Verification of Identity (VIVID) dataset and (ii) two self-captured videos using a mounted camera on a UAV. Precision and recall are used as the criteria to quantitatively evaluate and validate overall MOD performance. Furthermore, both are computed with respect to the ground-truth data which are manually annotated for the video sequences. The proposed framework has also been compared with existing state-of-the-art detection algorithms. Experimental results show that MDMRBF outperforms these algorithms with precision and recall being 94% and 89%, respectively. These results can be attributed to the integration of appearance and geometrical constraints for region matching using the multigraph structure. Moreover, the consideration of longer trajectories on multiple frames and taking into account all the transformations also facilitated in resolving occlusion. With regards to time, the proposed approach could detect moving objects within one minute for a 30-second sequence, which means that it is efficient in practice. In conclusion, the multiple moving object detection technique proposed in this study is robust to unknown transformations, with significant improvements in overall precision and recall compared to existing methods. The proposed algorithm is designed in order to tackle many limitations of the existing algorithms such as handle inevitable occlusions, model different motions from multiple moving objects, and consider the spatial information.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

PENGESANAN OBJEK BERGERAK DARI RAKAMAN VIDEO DARI UAV MENGGUNAKAN TRAJEKTORI-TRAJEKTORI TERPADAN GRAF RANTAUAN BERSEBELAHAN

Oleh

BAHAREH KALANTAR GHORASHI HARANDI

Disember 2017

Pengerusi : Profesor Shattri Bin Mansor, PhD
Fakulti : Kejuruteraan

Video yang dirakam menggunakan kamera dari kenderaan tanpa-manusia udara (UAV) biasanya menghasilkan rakaman dinamik yang sering mempunyai pergerakan kamera tidak stabil dan beberapa objek bergerak. Objek-objek ini kadang kala dihalang oleh tumbuhan atau objek lain, di mana ini mewujudkan satu persekitaran mencabar bagi pemprosesan dan analisis video peringkat tinggi. Tesis ini berkait dengan tajuk pengesenan objek bergerak (MOD) bertujuan untuk mengenal pasti dan mengesan satu atau beberapa objek bergerak dari video. Pada kebiasannya, MOD kebanyakan kali diatasi menggunakan pendaftaran imej, di mana kesamaan dicari dan ditemukan antara kerangka berturutan berdasarkan persamaan penampilan visual skala-kelabu spatial di bawah transformasi tegar dan afine. Walau bagaimanpun, pendaftaran tidak sesuai untuk rakaman video UAV kerana persamaan ukuran jarak skala-kelabu gagal untuk mengambil kira perbezaan spatio-temporal objek-objek bergerak. Pendafataran imej juga tidak berkesan apabila berhadapan dengan masalah objek terhalang. Oleh itu, tesis ini mencadangkan satu rangka kerja untuk menangani masalah-masalah yang dinyatakan melalui pendekatan dua peringkat, yang melibatkan pemadanan rantau dan penglabelan rantau. Secara spesifik, objektif-objektif tesis ini adalah (i) untuk membangunkan teknik pendaftaran imej berdasarkan padanan graf pelbagai, (ii) untuk mengesan objek-objek terhalang melalui penerokaan koresponden calon objek dalam urutan kerangka yang lebih panjang, dan (iii) untuk membangunkan satu algoritma pewarnaan graf yang robus bagi objek-objek bergerak pelbagai di bawah transformasi berbeza.

Secara am, setiap kerangka rakaman pada mulanya akan dibahagi kepada rantau-rantau superpiksel di mana fitur penampilan dan geometri akan dihitung. Maklumat trakejtori juga dipertimbang merentasi kerangka mengambil kira pelbagai

transformasi. Secara spesifik, setiap kerangka dimodel/diwakilkan oleh garf rantauan bersebelahan (RAG). Kemudian, selain menggunakan padanan spatial rawak antara dua kerangka sepertimana dalam pendaftaran imej, kesamaan di antara kerangka-kerangka video dikenal pasti melalui padanan pelbagai graf berasaskan fitur-fitur spatio-temporal robust setiap rantauan. Oleh kerana lebih dari dua kerangka diambil kira setiap masa, langkah ini lebih mampu mengenalpasti kesamaan antara rantauan dan boleh juga menangani halangan objek. Langkah kedua iaitu penglabelan rantauan bergantung kepada andaian bahawa objek bergerak pada latar belakang dan latar hadapan masing-masing menunjukkan sifat pergerakan berbeza apabila dikaji dalam turutan kerangka. Oleh itu, perbezaan spatial antara kedua jenis objek ini dijangka amat berbeza dari masa ke semasa. Berdasarkan andaian ini, proses penglabelan rantauan menetapkan sama ada label latar belakang atau latar hadapan diberi kepada sesuatu rantauan berdasarkan algoritma pewarnaan graf yang dicadangkan, yang mana ianya mengambil kira fitur-fitur trajektori. Secara keseluruhan, rangka kerja yang merangkumi kedua-dua langkah ini diberi nama Motion Differences of Matched Region-based Features (MDMRBF). MDMRBF telah dinilai menggunakan dua dataset iaitu (i) dataset Defense Advanced Research Projects Agency (DARPA) Video Verification of Identity (VIVID), dan (ii) dua video dari kamera UAV yang dirakam sendiri. Precision dan Recall digunakan sebagai kriteria kuantitatif untuk menilai dan mengesahkan prestasi pengesanan objek bergerak. Tambahan lagi, kedua-dua ukuran ini dikira berkenaan dengan data ground-truth yang dilabel secara manual untuk turutan video yang digunapakai. Rangka kerja yang dicadangkan juga telah dibandingkan dengan algoritma pengesan terkini. Keputusan eksperimen menunjukkan MDMRBF mampu mengatasi algoritma-algoritma berkenaan dengan ukuran precision dan recall masing-masing pada 94% dan 89%. Keputusan ini boleh dikaitkan dengan integrasi kekangan-kekangan penampilan dan geometrikal bagi pemadanan rantau menggunakan struktur multigraf. Tambahan lagi, pertimbangan trajektori-trajektori yang lebih panjang pada berbilang kerangka dan mengambil kira kesemua transformasi juga telah membantu dalam menyelesaikan masalah objek terhalang. Berkaitan masa, pendekatan yang dicadangkan mampu mengesan objek bergerak dalam masa satu minit bagi satu urutan 30-saat, yang membawa makna ianya eficien secara praktikal. Secara kesimpulan, teknik pengesanan objek pelbagai yang dicadangkan robus terhadap transformasi yang tidak diketahui, dengan penambah baikan yang penting di dalam kepersisan dan *recall* secara keseluruhan, berbanding kaedah-kaedah lain. Kaedah ini juga direka bentuk untuk menangani kekurangan kaedah-kaedah sedia ada seperti menangani objek terhalang, permodelan pergerakan berbeza bagi objek-objek bergerak berlainan, dan mengambil kira informosi reruang.

ACKNOWLEDGEMENTS

In the name of God, the most gracious the most merciful and praise be to Allah the cherisher and sustainer of the worlds. I am indeed grateful to him the Almighty for all the abundant blessings bestowed on me to successfully fulfill my purpose to travel out in search of knowledge.

I would like to express my over whelming gratitude to my great supervisor, Dato' Professor. Dr. Shattri Bin Mansor for making my dream of acquiring higher education a reality. You have been a tremendous mentor for me. I would like to thank you for encouraging my research and for allowing me to grow as an independent researcher during my PhD period. Thank you Dato' Prof. Shattri, you are indeed a full of wisdom, respect and empathy.

I am indeed very grateful to Dr. Alfian Abdul Halin and Associate. Prof. Dr. Helmi Zulhaidi M. Shafri for being part of my supervisory committee. I also would like to express the deepest thanks to Professor Dr. Biswajeet Pradhan for his support and encourages.

Million thanks go to my beloved family, father and mother and humble brother and sisters and their families for their relentless support during my candidature. My regards and appreciation also go to my colleagues and staff in Geospatial Information Science Research Center (GISRC), Department of Civil Engineering, Faculty of Engineering for their supports and encourages.

Last but not the least; I would like to present my special appreciations to Universiti Putra Malaysia for providing a beautiful, peaceful, and calm academic environment for research and study.

Bahareh Kalantar Ghorashi Harandi

This thesis was submitted to the Senate of the Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Shattri Bin Mansor, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Helmi Zulhaidi M. Shafri, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Alfian Abdul Halin, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

ROBIAH BINTI YUNUS, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date :

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software

Signature: ______________________________ Date: ______________________

Name and Matric No : Bahareh Kalantar Ghorashi Harandi, GS43158

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) were adhered to.

Signature: ______________________

Name of Chairman of Supervisory Committee: Professor Dr. Shattri Bin Mansor

Signature: ______________________

Name of Member of Supervisory Committee: Associate Professor Dr. Helmi Zulhaidi M. Shafri

Signature: ______________________

Name of Member of Supervisory Committee: Dr. Alfian Abdul Halin

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

8PP	Eight-Parameter Projective Model
ABM	Area-Based Methods
aMODERs	an Automatic Moving Object Detection, Extraction and Recognition System
bLPS-HOG	boosting Light and Pyramid Sampling Histogram Of Oriented Gradients
CC	Correlation Coefficient
CFOT	Consensus Foreground Object Template
CLIF	Columbus Large Image Format
DARPA	Defense Advanced Research Projects Agency
DATMO	Detect And Track Moving Objects
DCD	Dominant Color Descriptor
DECOLOR	DEtecting Contiguous Outliers in the LOw-rank Representation
DR	Detection Rate
EFD	Elliptic Fourier Descriptor
EM	Expectation–Maximization
FAR	False Alarm Rate
FBM	Feature-Based Methods
FGM	Factorized Graph Matching
FPR	False Positive Rate
GANs	General Adaptive Neighborhoods
GM	Graph Matching
GSD	Ground Sampling Distance
GTM	Graph Transformation Matching
ICP	Iterative Closest Point
KLT	Kanade–Lucas–Tomasi
LLT	Locally Linear Transforming
MDMRBF	Motion Differences of Matched Region-Based Features
MHI	Motion History Image

MI	Mutual Information
MIHT	Modified Iterated Hough Transform
MODAT	Moving Objects Detection And Tracking
MOD	Moving Object Detection
MSG	Motion Similarity Graph
PD	Probability of Detection
PM	Probability-based Matching
RAGs	Regional Adjacency Graphs
RANSAC	RANdom SAmple Consensus
RKHS	Reproducing Kernel Hilbert Space
SIFT	Scale-Invariant Feature Transform
SLIC	Simple Linear Iterative Clustering
SMAC	Spectral Matching with Affine Constraints
STARS	Spatio-Temporal Appearance-Related Similarity
SURF	Speeded Up Robust Features
SVM	Support Vector Machine
TLS	Total Least Squares
TPS	Thin-Plate Spline
UAV	Unmanned Aerial Vehicle
UPM	Universiti Putra Malaysia
VIS	Visual Optical
VIVID	VIdeo Verification of IDentity

CHAPTER 1

INTRODUCTION

1.1 Background

Due to the increasing demand for intelligent surveillance and limitations of stationary cameras, video surveillance on an airborne vehicle such as unmanned aerial vehicle (UAV) has been gaining increasing attention in recent years (Huang et al., 2010). The high mobility, fast deployment, and large surveillance scope in UAV videos offer a wide range of applications, such as moving object detection (MOD) (Zhou et al., 2013; Tian et al., 2011), object tracking (Hu et al., 2015;Cao et al., 2012; Babenko et al., 2011; Xiao et al., 2008), motion segmentation (Lian et al., 2013; Hu et al., 2012), object classification and identification (Somasundaram et al., 2013), event detection (Wang et al., 2008), and behavior understanding and description (Borges et al., 2013). Essentially, accurate MOD is a fundamental task for establishing robust intelligent video surveillance systems. UAVs are low-cost platforms that can provide efficient data acquisition mechanisms for intelligent systems. As they fly at low altitudes, videos are captured with very high spatial resolution (Zhou et al., 2015). Hence, the analysis and interpretation of these video sequences are active research fields.

In contrast to the stationary cameras, detecting moving objects from videos captured by moving cameras is significantly more challenging. This is due to the camera motion which is independent of moving objects' motions. This task is even more challenging when surveillance cameras are mounted on airborne vehicles such as UAVs. Typically, UAVs fly at low altitudes, and render high mobility, fast deployment, and large surveillance scope (Zhou et al., 2015). Therefore, the existence of multiple moving objects, large object displacements due to the high velocity, very small object movement, objects leaving the field of view, changes in illumination, and occlusion caused by terrain features are common in UAV-captured videos. Several approaches have been proposed in the literature (Ingersoll et al., 2015; Lee et al., 2015; Zhou et al., 2015; Bhaskar et al., 2014; Cao et al., 2012; Rodríguez-Canosa et al., 2012; Bhattacharya et al., 2011; Miller et al., 2008), which can be generally divided into three main categories, i.e., temporal differencing, background subtraction, and optical flow (Lu et al., 2008).

Temporal differencing is interested in determining the differences between consecutive frames in order to classify background and foreground pixels. However, this technique is unable to detect all relevant pixels and complete shapes of foreground objects. Moreover, small changes in object movements or stopping objects can cause temporal differencing to fail. Unlike temporal differencing, background subtraction uses a model of the background based on the pixel distribution. Moving objects are then detected by subtracting the current frame from the background model (Liu et al., 2009). A priori knowledge about object shape, motion characteristics, and the spatial dependency among neighboring background pixels are not taken into account in background subtraction techniques. Although these methods are flexible and fast, background scenes need to be consistent while the camera should also be fixed. Otherwise, any changes can potentially be treated as moving objects. This issue however, can be resolved by employing image stabilization and registration methods, where two images of the same scene taken at different times are geometrically overlaid (Jackson and Goshtasby, 2010).

Feature detection and feature matching are two fundamental stages in the majority of the registration methods (Zitova and Flusser, 2003). A popular solution for feature detection is to represent images as bags of features and find correspondences using appearances (Yu et al., 2014). However, spatial feature layouts and correlations between pixels are ignored using such order-less sets of local descriptors. The feature matching problem is complicated due to the lack of one-to-one correspondences and the presence of noise. Moreover, illumination changes in consecutive frames make correspondences between feature points unreliable. More importantly, the transformations that align points are simply assumed to be parametric (e.g., rigid, and affine), which is not the case in real life situations. Hence, motions of the foreground objects are estimated once background motion is eliminated by image registration techniques. However, these techniques are not mature enough to truly eliminate different kinds of motions. Moreover, estimating multiple foreground motions is still problematic.

Alternatively, moving objects can be detected using optical flow. This is based on local image motion approximations where pixel movements between consecutive frames are specified. However, optical flow-based methods are sensitive to illumination changes, and only partial edge shapes of moving objects are detected (Lu et al., 2008). Although most existing algorithms are successful in many situations, MOD in aerial videos is still very challenging. This thesis suggests to geometrically overlay consecutive images and use motion information to label background and foreground objects. Thus, matching and labeling are two main steps in the proposed approach. The first step is similar to the registration techniques while it can address their problems. The second step can relieve multiple foreground detection under different transformations.

In this thesis, both appearance similarity and geometrical constraints are imposed on region-based features. We believe that this combination makes for a representation that is more robust to local variations. If images are seen as a set of connected regions, they can hence be represented using regional adjacency graphs (RAGs). This allows the spatial relationships between pixels to also be incorporated at a higher level, making it more robust to illumination and intensity across frame sequences. Representing images as graphs of regions also allows the utilization of graph matching algorithms in order to find visual correspondences. Therefore, both unary node-to-node, and pairwise edge-to-edge relationships can be integrated into the model, which is different from point-based matching or registration approaches such as Random Sample Consensus (RANSAC) (Fischler and Bolles, 1981) and Iterative Closest Point (ICP) (Liu, 2004). Specifically, the correspondences between regions of consecutive frames can be obtained using both appearance structural similarities and region coordinates as geometrical information. By using both appearances and geometrical constraints to find correspondences, the proposed method can be more robust to deformation, missing data, and outliers. Therefore, regions are only matched with corresponding regions having similar visual and geometrical descriptors within context.

Once one-to-one correspondences between regions in successive frames are established, motion models can be constructed to distinguish between background and foreground regions. Since different objects (whether belonging to background or foreground) in a sequence move independently, their spatial distances must be varying with time. Inspired by this, spatial distances between adjacent regions are observed in the sequence. Any detected change is interpreted as their motion difference. More specifically, a graph is constructed from all detected image regions which their correspondences in the trajectory have been successfully established. Edge weights in this graph are assigned according to the motion similarity between adjacent regions. The objective is to distinguish between background and foreground regions based on the motion features. We formulate this as a graph labeling problem which its aim is to label graph nodes as background or foreground. This can be efficiently achieved using the proposed graph coloring algorithm. This is very surprising since using temporal motion information of the neighboring regions in a sequence, we do not explicitly estimate their individual transformations. Thus, moving objects can take different transformations, either being parametric (e.g., rigid, affine) or nonparametric (e.g., nonrigid). We refer to this method as MDMRBF (motion differences of matched region-based features). Therefore, this thesis attempts to examine whether using trajectories of matched regional adjacency graphs can detect moving objects or not.

1.2 Problem Statement

Unlawful transgressive and boundary interference by immigrants are a major obstacle demeaned against the border security. Several techniques and scientific methods have been proposed over the past decade to ensure protection of the boundaries. The traditional border monitoring by artificial monitoring ways include standing guard, lookout, patrol, video camera, ground sensors, physical barriers, land vehicles and manned aircraft. These methods are unable to meet the needs of border monitoring well due to the complex border areas and small surveillance scope. Moreover, it becomes cumbersome for human operators to monitor for long durations.

The rapid development of technology leads to intelligent monitoring system development using UAV. Automatic MOD by UAV replaces traditional border monitoring. UAV can supply images even on cloudy days, monitor over long distance with a low cost, fly flexibly across broad spatial and temporal scales, and carry various types of sensors to collect abundant data. However, detecting moving objects in UAV-captured video sequences is very challenging as two types of motions need to be considered. They are the camera motion and also the motions of the moving objects. In addition, other issues make detection even more challenging. For instance, the existence of multiple moving objects, large object displacements due to the high velocity, very small object movement, objects leaving the field of view, changes in illumination, and occlusion caused by terrain features are common in UAV videos. This thesis proposes a novel method to effectively address these challenges. Hence, it is not only applicable to any surveillance system, but also considers the especial difficulties in UAV-captured videos.

Though many algorithms have been proposed for different situations, developing a unified framework for remote sensing image registration is still challenging. This is due to the specific circumstances of remote sensing imaging. For instance, unavoidable noise, occlusions, repeated structures, and nonrigid transformation are prevalent in these images. Hence, simple parametric models (e.g., rigid or affine transformation) cannot be applied for images captured under different imaging viewpoint variations and distortions (Ma et al., 2015c). Consequently, an accurate matching must take into account the nonrigid transformations and outlier removal while accommodate to environmental changes.

In this thesis, a novel approach is introduced, which can efficiently leverage spatial context at the pixel level, appearance similarity and geometrical constraints at the region level, and outlier removal and occlusion detection on the sequence level; all to establish the correspondences between regions in frame sequences from UAV videos. This is achieved by the proposed multigraph matching technique on RAGs of successive frames. Moreover, temporal motion information of the neighboring regions in a sequence is used to label graph nodes as background or foreground regions.

1.3 Research Objectives

The overarching aim of this research is to develop a multiple moving object detection approach, using region trajectories in matched regional adjacency graphs, for UAV-captured videos. The three objectives of this thesis are:

1. To develop an image registration technique using a multigraph matching algorithm on RAGs of successive frames.
2. To detect occluded regions by exploring the correspondences in a longer sequence, since they can be visible and detectable at long-term trajectories.
3. To develop a robust graph coloring algorithm for multiple moving object detection under different transformations.

1.4 Research Questions

This thesis comprehensively addresses the following research questions:

1. Which image features should be used for motion compensation?
2. What kind of methodology is most suitable and more accurate for small moving object detection?
3. Which method can use for detection moving object in dynamic environment?
4. How should the motion, appearance, and shape of the object be modeled?
5. How to label background and foreground objects?

1.5 Scope of Study

In this thesis, a multiple moving object detection approach from a UAV video based on trajectories in matched RAGs will be designed. Trajectories in a matched region provide rich information in motion estimation. As the motions are estimated independently for each pair of regions in the trajectories, the most common geometric distortions found in remotely sensed imagery can be effectively handled. Moreover, the motion information in the region trajectories are used not only to differentiate between background and foreground regions, but also to locate and detect multiple moving objects. More importantly, using a longer sequence of multiple frames has another advantage, which is to detect occluded regions because the occluded regions which are not visible in the current regions can be detectable in long-term trajectories.

The videos that are used in the experimental purposes are public standard color imagery datasets and our home-made videos. While public standard dataset covers about $0.5km^2$ with an image size of 640×480 pixels and a frame rate of 30 Hz, our home-made videos have an image size of 1280×720 and 1920×1080 pixels and a frame rate of 30 Hz. The videos that are chosen to be tested in this study are: EgTest01, EgTest02, EgTest03, EgTest04, EgTest05 selected from a popular dataset, the DARPA VIVID videos dataset (Defense Advanced Research Projects Agency (DARPA) Video Verification of Identity (VIVID)) (Collins et al., 2005) and our home-made videos collected from the campus of Universiti Putra Malaysia (UPM). We address vehicles as moving object in this research. In the standard dataset, similar vehicles move on a runway, speed up, and pass by each other or two groups of vehicles pass by each other on a runway. The scale is changed as the camera circles the scene. The vehicles are occluded by each other or by trees, and there are illumination changes. Some frames are duplicated as the camera fails to record these frames. Thus, there is no motion followed by a sudden discontinuity in the sequence. On the other hand, our home-made videos, vehicles are tracked along the outdoor campus environment. It contains appearance variations and cluttered scenes.

To quantitatively evaluate the performance of the proposed approach, precision and recall are utilized. Precision is calculated as the percentage of correctly detected object pixels over the total number of detected object pixels. Recall indicates the ratio of correctly detected object pixels to the number of actual object pixels. The developed algorithm is successfully applied to the data, and the verification result is promising.

The proposed framework can be used for border protection, reconnaissance, or homeland security. Illegal trespassing and border encroachment by immigrants is a huge predicament against border security and homeland security. The main application of this research is to assist the human operators, by implementing intelligent visual surveillance systems which help in detecting suspicious or unusual events in the video sequence.

1.6 Thesis Organization

The remainder of this thesis is prepared as follows:

Chapter two presents a description about object representation, and object detection from videos using popular algorithms such as segmentation, background subtracting and modeling, temporal differencing, and optical flow. Furthermore, this chapter reviews some of the major limitations and challenges of MOD using UAV-captured videos. This is followed by a review of previous studies include image registration, general object detection, and UAV-based object detection. Finally, the gap of research is mentioned according to the previous studies.

Chapter three presents the methodology and the algorithms developed for this thesis. A novel approach is proposed for multiple moving object detection from a nonstationary camera, based on region trajectories in matched regional adjacency graphs. This chapter describes the segmentation algorithm followed by a graph construction technique to generate the graphs of the regions. Then a region merging process is proposed. Afterward, a multigraph matching algorithm is presented to find the correspondence between segmented regions in the frame sequence. Thus, the output of the region matching stage is a set of all the matched regions in the trajectory of specific length. Later, detection occluded regions are explored by using long-term analyzing of multiple frames. A motion similarity graph is constructed using corresponding regions in the trajectory, in order to discriminate the foreground and background regions. Subsequently, a graph coloring algorithm is proposed to assign appropriate labels to background and moving object. Finally, a computational complexity analysis and validation of the proposed framework is presented.

Chapter four provides the results and discussion of the proposed algorithm. In this chapter, first, the dataset and implementation and development environment are explained. Second, the results of image segmentation and region merging on different standard publicly available benchmark datasets and our home-made videos are discussed. Then, the results of region matching are explained. This is followed by an evaluation of our algorithm for region matching task on different sequences, and this is compared with other methods. Afterward, the occlusion detection results using long-term trajectories are explained. Then, MOD results are explained followed by investigation the effect of UAV parameters on the proposed algorithm. In order to quantitatively evaluate the performance of the proposed method, well-known precision, and recall techniques are utilized. Finally, the performance of our approach on different datasets are discussed and compared with the best state-of-the-art MOD techniques.

Lastly, Chapter five summarizes the contributions of this thesis to the advancement of our knowledge in this field, and provides conclusion for this thesis, and suggests some future research directions.

REFERENCES

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods.*IEEE transactions on pattern analysis and machine intelligence*, *34*(11), 2274-2282.

Aeschliman, C., Park, J., & Kak, A. C. (2014). Tracking vehicles through shadows and occlusions in wide-area aerial video. *IEEE Transactions on Aerospace and Electronic Systems*, *50*(1), 429-444.

Aguilar, W., Frauel, Y., Escolano, F., Martinez-Perez, M. E., Espinosa-Romero, A., & Lozano, M. A. (2009). A robust graph transformation matching for non-rigid registration. *Image and Vision Computing*, *27*(7), 897-910.

Ali, S., & Shah, M. (2006, May). COCOA: tracking in aerial imagery. In*Defense and Security Symposium* (pp. 62090D-62090D). International Society for Optics and Photonics.

Alkanat, T., Tunali, E., & Öz, S. (2015). A Real-time, Automatic Target Detection and Tracking Method for Variable Number of Targets in Airborne Imagery. In *VISAPP (2)* (pp. 61-69).

Babenko, B., Yang, M. H., & Belongie, S. (2011). Robust object tracking with online multiple instance learning. *IEEE transactions on pattern analysis and machine intelligence*, *33*(8), 1619-1632.

Belongie, S., Malik, J., & Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE transactions on pattern analysis and machine intelligence*, *24*(4), 509-522.

Bentoutou, Y., Taleb, N., Kpalma, K., & Ronsin, J. (2005). An automatic image registration for applications in remote sensing. *IEEE transactions on geoscience and remote sensing*, *43*(9), 2127-2137.

Bhaskar, H., Dias, J., Seneviratne, L., & Al-Mualla, M. (2014, December). μ-UAV Based Dynamic Target Tracking for Surveillance and Exploration. In*International Symposium on Visual Computing* (pp. 445-456). Springer International Publishing.

Bhattacharya, S., Idrees, H., Saleemi, I., Ali, S., & Shah, M. (2011). Moving object detection and tracking in forward looking infra-red aerial imagery. In*Machine vision beyond visible spectrum* (pp. 221-252). Springer Berlin Heidelberg.

Borges, P. V. K., Conci, N., & Cavallaro, A. (2013). Video-based human behavior understanding: A survey. *IEEE transactions on circuits and systems for video technology*, *23*(11), 1993-2008.

Brox, T., & Malik, J. (2010). Object segmentation by long term analysis of point trajectories. *Computer Vision–ECCV 2010*, 282-295.

Bunke, H., Irniger, C., & Neuhaus, M. (2005). Graph matching–challenges and potential solutions. *Image Analysis and Processing–ICIAP 2005*, 1-10.

Cao, X., Lan, J., Yan, P., & Li, X. (2012). Vehicle detection and tracking in airborne videos by multi-motion layer analysis. *Mach. Vis. Appl.*, *23*(5), 921-935.

Cao, X., Wu, C., Lan, J., Yan, P., & Li, X. (2011). Vehicle detection and motion analysis in low-altitude airborne video under urban environment. *IEEE Transactions on Circuits and Systems for Video Technology*, *21*(10), 1522-1533.

Cho, M., Lee, J., & Lee, K. (2010). Reweighted random walks for graph matching. *Computer Vision–ECCV 2010*, 492-505.

Chung, K. L., Lin, Y. R., & Huang, Y. H. (2009). Efficient shadow detection of color aerial images based on successive thresholding scheme. *IEEE Transactions on Geoscience and Remote sensing*, *47*(2), 671-682.

Cohen, I., & Medioni, G. (1999). Detecting and tracking moving objects for video surveillance. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.* (Vol. 2, pp. 319-325). IEEE.

Collins, R. T., Lipton, A. J., Kanade, T., Fujiyoshi, H., Duggins, D., Tsin, Y., ... & Wixson, L. (2000). A system for video surveillance and monitoring.*VSAM final report*, 1-68.

Collins, R., Zhou, X., & Teh, S. K. (2005, January). An open source tracking testbed and evaluation web site. In *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS 2005)* (Vol. 2, p. 35).

Comaniciu, D., Meer, P., & Foran, D. J. (1999). Image-guided decision support system for pathology. *Machine Vision and Applications*, *11*(4), 213-224.

Commowick, O., Wiest-Daesslé, N., & Prima, S. (2012, May). Block-matching strategies for rigid registration of multimodal medical images. In*Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on* (pp. 700-703). IEEE.

Cour, T., Srinivasan, P., & Shi, J. (2007). Balanced graph matching. In*Advances in Neural Information Processing Systems* (pp. 313-320).

Debayle, J., & Pinoli, J. C. (2006). General adaptive neighborhood image processing. *Journal of Mathematical Imaging and Vision*, *25*(2), 245-266.

Debayle, J., & Presles, B. (2016). Rigid image registration by General Adaptive Neighborhood matching. *Pattern Recognition*, *55*, 45-57.

Deng, K., Tian, J., Zheng, J., Zhang, X., Dai, X., & Xu, M. (2010). Retinal fundus image registration via vascular structure graph matching. *Journal of Biomedical Imaging*, *2010*, 14.

Dimitriou, N., & Delopoulos, A. (2015). Incorporating higher order models for occlusion resilient motion segmentation in streaming videos. *Image and Vision Computing*, *36*, 70-82.

Donné, S., Aelterman, J., Goossens, B., & Philips, W. (2015, October). Fast and Robust Variational Optical Flow for High-Resolution Images Using SLIC Superpixels. In *International Conference on Advanced Concepts for Intelligent Vision Systems* (pp. 205-216). Springer International Publishing.

Dou, J., & Li, J. (2013). Moving object detection based on improved VIBE and graph cut optimization. *Optik-International Journal for Light and Electron Optics*, *124*(23), 6081-6088.

Du, Q., Raksuntorn, N., Orduyilmaz, A., & Bruce, L. M. (2008). Automatic registration and mosaicking for airborne multispectral image sequences.*Photogrammetric Engineering & Remote Sensing*, *74*(2), 169-181.

Fan, B., Du, Y., & Tang, Y. (2011, June). Efficient registration algorithm for UAV image sequence. In *Information and Automation (ICIA), 2011 IEEE International Conference on* (pp. 111-116). IEEE

Felzenszwalb, P. F., & Huttenlocher, D. P. (2004). Efficient graph-based image segmentation. *International journal of computer vision*, *59*(2), 167-181.

Finlayson, G. D., Hordley, S. D., Lu, C., & Drew, M. S. (2006). On the removal of shadows from images. *IEEE transactions on pattern analysis and machine intelligence*, *28*(1), 59-68.

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, *24*(6), 381-395.

Fonseca, L. M., & Manjunath, B. S. (1996). Registration techniques for multisensor remotely sensed imagery. *PE & RS- Photogrammetric Engineering & Remote Sensing*, *62*(9), 1049-1056.

Fowers, S. G., Lee, D. J., Ventura, D. A., & Archibald, J. K. (2013). The nature-inspired BASIS feature descriptor for UAV imagery and its hardware implementation. *IEEE Transactions on Circuits and Systems for Video Technology*, *23*(5), 756-768.

Friedman, N., & Russell, S. (1997, August). Image segmentation in video sequences: A probabilistic approach. In *Proceedings of the Thirteenth conference on Uncertainty in artificial intelligence* (pp. 175-181). Morgan Kaufmann Publishers Inc.

Gandhi, T. L., Devadiga, S., Kasturi, R., & Camps, O. I. (1996, December). Detection of obstacles on runway using ego-motion compensation and tracking of significant features. In *Applications of Computer Vision, 1996. WACV'96., Proceedings 3rd IEEE Workshop on* (pp. 168-173). IEEE.

Gaszczak, A., Breckon, T. P., & Han, J. (2011, January). Real-time people and vehicle detection from UAV imagery. In *IS&T/SPIE Electronic Imaging* (pp. 78780B-78780B). International Society for Optics and Photonics.

Gold, S., & Rangarajan, A. (1996). A graduated assignment algorithm for graph matching. *IEEE Transactions on pattern analysis and machine intelligence*, *18*(4), 377-388

Goshtasby, A., & Stockman, G. C. (1985). Point pattern matching using convex hull edges. *IEEE Transactions on Systems, Man, and Cybernetics*, (5), 631-637.

Habib, A., Asmamaw, A., Kelley, D., & May, M. (2000). *Linear features in photogrammetry*. Ohio State University. Division of Geodetic Science.

Habib, A., Morgan, M., Kim, E. M., & Cheng, R. (2004, July). Linear features in photogrammetric activities. In *ISPRS Congress, Istanbul, Turkey* (p. 610).

Heikkilä, J., & Silvén, O. (2004). A real-time system for monitoring of cyclists and pedestrians. *Image and Vision Computing*, *22*(7), 563-570.

Heikkila, M., & Pietikainen, M. (2006). A texture-based method for modeling the background and detecting moving objects. *IEEE transactions on pattern analysis and machine intelligence*, *28*(4), 657-662.

Hong, G., & Zhang, Y. (2007, July). Combination of feature-based and area-based image registration technique for high resolution remote sensing image. In *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International* (pp. 377-380). IEEE.

Hu, M. K. (1962). Visual pattern recognition by moment invariants. *IRE transactions on information theory*, *8*(2), 179-187.

Hu, W. C., Chen, C. H., Chen, T. Y., Huang, D. Y., & Wu, Z. C. (2015). Moving object detection and tracking from video captured by moving camera. *Journal of Visual Communication and Image Representation*, *30*, 164-180.

Hu, W. C., Chen, C. H., Huang, D. Y., & Ye, Y. T. (2012). Video object segmentation in rainy situations based on difference scheme with object structure and color analysis. *Journal of Visual Communication and Image Representation*, *23*(2), 303-312.

Huang, C. H., Wu, Y. T., Kao, J. H., Shih, M. Y., & Chou, C. C. (2010). A hybrid moving object detection method for aerial images. *Advances in Multimedia Information Processing-PCM 2010*, 357-368.

Ibrahim, A. W. N., Ching, P. W., Seet, G. G., Lau, W. M., & Czajewski, W. (2010, November). Moving objects detection and tracking framework for UAV-based surveillance. In *Image and Video Technology (PSIVT), 2010 Fourth Pacific-Rim Symposium on* (pp. 456-461). IEEE.

Ingersoll, K., Niedfeldt, P. C., & Beard, R. W. (2015, June). Multiple target tracking and stationary object detection in video with Recursive-RANSAC and tracker-sensor feedback. In *Unmanned Aircraft Systems (ICUAS), 2015 International Conference on* (pp. 1320-1329). IEEE.

Jackson, B. P., & Goshtasby, A. A. (2010). Registering aerial video images using the projective constraint. *IEEE Transactions on image processing*,*19*(3), 795-804.

Jaggi, M. (2013, June). Revisiting Frank-Wolfe: Projection-Free Sparse Convex Optimization. In *ICML (1)* (pp. 427-435).

Johnson, H. J., & Christensen, G. E. (2002). Consistent landmark and intensity-based image registration. *IEEE transactions on medical imaging*,*21*(5), 450-461.

KaewTraKulPong, P., & Bowden, R. (2002). An improved adaptive background mixture model for real-time tracking with shadow detection. In*Video-based surveillance systems* (pp. 135-144). Springer US.

Kamate, S., & Yilmazer, N. (2015). Application of Object Detection and Tracking Techniques for Unmanned Aerial Vehicles. *Procedia Computer Science*, *61*, 436-441.

Karasulu, B. (2010). Review and Evaluation of Well-Known Methods for Moving Object Detection and Tracking in Videos. *Journal of aeronautics and space technologies*, *4*(4), 11-22.

Khashman, A. (2008). Automatic Detection, Extraction and Recognition of moving objects. *International Journal of Systems Applications, Engineering & Development*, *2*(1), 43-51.

Khotanzad, A., & Hong, Y. H. (1990). Invariant image recognition by Zernike moments. *IEEE Transactions on pattern analysis and machine intelligence*,*12*(5), 489-497.

Kuhl, F. P., & Giardina, C. R. (1982). Elliptic Fourier features of a closed contour. *Computer graphics and image processing*, *18*(3), 236-258.

Le Moigne, J., Campbell, W. J., & Cromp, R. F. (2002). An automated parallel image registration technique based on the correlation of wavelet features. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(8), 1849-1864.

Lee, S., Kim, N., Jeong, K., Park, K., & Paik, J. (2015). Moving object detection using unstable camera for video surveillance systems. *Optik-International Journal for Light and Electron Optics*, *126*(20), 2436-2441.

Leordeanu, M., Hebert, M., & Sukthankar, R. (2009). An integer projected fixed point method for graph matching and map inference. In *Advances in neural information processing systems* (pp. 1114-1122).

Li, H., Zhang, L., & Shen, H. (2014). An adaptive nonlocal regularized shadow removal method for aerial remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(1), 106-120.

Lian, F. L., Lin, Y. C., Kuo, C. T., & Jean, J. H. (2013). Voting-based motion estimation for real-time video transmission in networked mobile camera systems. *IEEE transactions on industrial informatics*, *9*(1), 172-180.

Lin, C. C., & Wolf, M. (2010, August). Detecting moving objects using a camera on a moving platform. In *Pattern Recognition (ICPR), 2010 20th International Conference on* (pp. 460-463). IEEE.

Lipton, A. J., Fujiyoshi, H., & Patil, R. S. (1998, October). Moving target classification and tracking from real-time video. In *Applications of Computer Vision, 1998. WACV'98. Proceedings., Fourth IEEE Workshop on* (pp. 8-14). IEEE.

Liu, C., Yuen, P. C., & Qiu, G. (2009). Object motion detection using information theoretic spatio-temporal saliency. *Pattern Recognition*, *42*(11), 2897-2906.

Liu, Y. (2004). Improving ICP with easy implementation for free-form surface matching. *Pattern Recognition*, *37*(2), 211-226.

Liu, Z., An, J., & Jing, Y. (2012). A simple and robust feature point matching algorithm based on restricted spatial order constraints for aerial image registration. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(2), 514-527.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In*Computer vision, 1999. The proceedings of the seventh IEEE international conference on* (Vol. 2, pp. 1150-1157). Ieee.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, *60*(2), 91-110.

Lu, N., Wang, J., Wu, Q. H., & Yang, L. (2008). An Improved Motion Detection Method for Real-Time Surveillance. *IAENG International Journal of Computer Science*, *35*(1).

Ma, J., Chan, J. C. W., & Canters, F. (2010). Fully automatic subpixel image registration of multiangle CHRIS/Proba data. *IEEE transactions on geoscience and remote sensing*, *48*(7), 2829-2839.

Ma, J., Ma, Y., Zhao, J., & Tian, J. (2015a). Image feature matching via progressive vector field consensus. *IEEE Signal Processing Letters*, *22*(6), 767-771.

Ma, J., Qiu, W., Zhao, J., Ma, Y., Yuille, A. L., & Tu, Z. (2015b). Robust L2E Estimation of Transformation for Non-Rigid Registration. *IEEE Trans. Signal Processing*, *63*(5), 1115-1129.

Ma, J., Zhao, J., Tian, J., Yuille, A. L., & Tu, Z. (2014). Robust point matching via vector field consensus. *IEEE Transactions on Image Processing*, *23*(4), 1706-1721.

Ma, J., Zhou, H., Zhao, J., Gao, Y., Jiang, J., & Tian, J. (2015c). Robust feature matching for remote sensing image registration via locally linear transforming. *IEEE Transactions on Geoscience and Remote Sensing*,*53*(12), 6469-6481.

Manjunath, B. S., & Ma, W. Y. (1996). Texture features for browsing and retrieval of image data. *IEEE Transactions on pattern analysis and machine intelligence*, *18*(8), 837-842.

Manjunath, B. S., Salembier, P., & Sikora, T. (2002). *Introduction to MPEG-7: multimedia content description interface* (Vol. 1). John Wiley & Sons.

Mao, H., Yang, C., Abousleman, G. P., & Si, J. (2010, April). Automated multiple target detection and tracking in UAV videos. In *SPIE Defense, Security, and Sensing* (pp. 76680J-76680J). International Society for Optics and Photonics.

Martinez, A., Garcia-Consuegra, J., & Abad, F. (1999). A correlation-symbolic approach to automatic remotely sensed image rectification. In*Geoscience and Remote Sensing Symposium, 1999. IGARSS'99 Proceedings. IEEE 1999 International* (Vol. 1, pp. 336-338). IEEE.

Máttyus, G., Benedek, C., & Szirányi, T. (2010). Multi target tracking on aerial videos.

Mayhew, C. A., & Mayhew, C. M. (2012, May). Parallax visualization of UAV FMV and WAMI imagery. In *SPIE Defense, Security, and Sensing* (pp. 83600H-83600H). International Society for Optics and Photonics.

McCartney, M. I., Zein-Sabatto, S., & Malkani, M. (2009, March). Image registration for sequence of visual images captured by uav. In *Computational Intelligence for Multimedia Signal and Vision Processing, 2009. CIMSVP'09. IEEE Symposium on* (pp. 91-97). IEEE.

Miller, A., Babenko, P., Hu, M., & Shah, M. (2008). Person tracking in UAV video. *Multimodal Technologies for Perception of Humans*, 215-220.

Nex, F., & Remondino, F. (2014). UAV for 3D mapping applications: a review.*Applied Geomatics*, *6*(1), 1-15.

Nock, R., & Nielsen, F. (2005). Semi-supervised statistical region refinement for color image segmentation. *Pattern Recognition*, *38*(6), 835-846.

Ourselin, S., Roche, A., Prima, S., & Ayache, N. (2000, October). Block matching: A general framework to improve robustness of rigid registration of medical images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 557-566). Springer Berlin Heidelberg.

Paragios, N., & Deriche, R. (2000). Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE Transactions on pattern analysis and machine intelligence*, *22*(3), 266-280.

Penatti, O. A., Valle, E., & Torres, R. D. S. (2012). Comparative study of global color and texture descriptors for web image retrieval. *Journal of Visual Communication and Image Representation*, *23*(2), 359-380.

Rao, S. R., Tron, R., Vidal, R., & Ma, Y. (2008, June). Motion segmentation via robust subspace separation in the presence of outlying, incomplete, or corrupted trajectories. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on* (pp. 1-8). IEEE.

Rashwan, H. A., García, M. A., & Puig, D. (2013). Variational optical flow estimation based on stick tensor voting. *IEEE Transactions on Image Processing*, *22*(7), 2589-2599.

Reddy, B. S., & Chatterji, B. N. (1996). An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE transactions on image processing*, *5*(8), 1266-1271.

Reilly, V., Idrees, H., & Shah, M. (2010). Detection and tracking of large number of targets in wide area surveillance. *Computer Vision–ECCV 2010*, 186-199.

Rodríguez-Canosa, G. R., Thomas, S., del Cerro, J., Barrientos, A., & MacDonald, B. (2012). A real-time method to detect and track moving objects (DATMO) from unmanned aerial vehicles (UAVs) using a single camera. *Remote Sensing*, *4*(4), 1090-1111.

Rodríguez-Canosa, G. R., Thomas, S., del Cerro, J., Barrientos, A., & MacDonald, B. (2012). A real-time method to detect and track moving objects (DATMO) from unmanned aerial vehicles (UAVs) using a single camera. *Remote Sensing*, *4*(4), 1090-1111.

Rosenbaum, D., Leitloff, J., Kurz, F., Meynberg, O., & Reize, T. (2010). Real-time image processing for road traffic data extraction from aerial images.

Saleemi, I., & Shah, M. (2013). Multiframe many–many point correspondence for vehicle tracking in high density wide area aerial videos.*International journal of computer vision*, *104*(2), 198-219.

Schellewald, C., & Schnörr, C. (2005, November). Probabilistic subgraph matching based on convex relaxation. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition* (pp. 171-186). Springer Berlin Heidelberg.

Shafique, K., & Shah, M. (2005). A noniterative greedy algorithm for multiframe point correspondence. *IEEE transactions on pattern analysis and machine intelligence*, *27*(1), 51-65.

Shen, D. (2007). Image registration by local histogram matching. *Pattern Recognition*, *40*(4), 1161-1172.

Shen, H., Li, S., Zhu, C., Chang, H., & Zhang, J. (2013). Moving object detection in aerial video based on spatiotemporal saliency. *Chinese Journal of Aeronautics*, *26*(5), 1211-1217.

Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence*, *22*(8), 888-905.

Siam, M., ElSayed, R., & ElHelw, M. (2012, December). On-board multiple target detection and tracking on camera-equipped aerial vehicles. In*Robotics and Biomimetics (ROBIO), 2012 IEEE International Conference on*(pp. 2399-2405). IEEE.

Solé-Ribalta, A., & Serratosa, F. (2011). Models and algorithms for computing the common labelling of a set of attributed graphs. *Computer Vision and Image Understanding*, *115*(7), 929-945.

Somasundaram, G., Sivalingam, R., Morellas, V., & Papanikolopoulos, N. (2013). Classification and counting of composite objects in traffic scenes using global and local image analysis. *IEEE Transactions on Intelligent Transportation Systems*, *14*(1), 69-81.

Song, E. Y., & Lee, J. J. (2004). Moving object detection using region tracking. *Artificial Life and Robotics*, *8*(1), 20-28.

Spagnolo, P., Leo, M., & Distante, A. (2006). Moving object segmentation by background subtraction and temporal analysis. *Image and Vision Computing*,*24*(5), 411-423.

Stauffer, C., & Grimson, W. E. L. (1999). Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.* (Vol. 2, pp. 246-252). IEEE.

Su, B., Wang, H., Liang, X., & Ji, H. (2014). Moving Objects Detecting and Tracking for Unmanned Aerial Vehicle. In *Foundations and Practical Applications of Cognitive Systems and Information Processing* (pp. 317-333). Springer Berlin Heidelberg.

Sun, S. W., Wang, Y. C. F., Huang, F., & Liao, H. Y. M. (2013). Moving foreground object detection via robust SIFT trajectories. *Journal of Visual Communication and Image Representation*, *24*(3), 232-243.

Teutsch, M., & Krüger, W. (2012a, July). Spatio-temporal fusion of object segmentation approaches for moving distant targets. In *Information Fusion (FUSION), 2012 15th International Conference on* (pp. 1988-1995). IEEE.

Teutsch, M., & Krüger, W. (2012b, September). Detection, segmentation, and tracking of moving objects in UAV videos. In *Advanced Video and Signal-Based Surveillance (AVSS), 2012 IEEE Ninth International Conference on*(pp. 313-318). IEEE.

Tian, Y., Feris, R. S., Liu, H., Hampapur, A., & Sun, M. T. (2011). Robust detection of abandoned and removed objects in complex surveillance videos.*IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *41*(5), 565-576.

Torresani, L., Kolmogorov, V., & Rother, C. (2008). Feature correspondence via graph matching: Models and global optimization. *Computer Vision–ECCV 2008*, 596-609.

Trémeau, A., & Colantoni, P. (2000). Regions adjacency graph applied to color image segmentation. *IEEE Transactions on image processing*, *9*(4), 735-744.

Tsai, V. J. (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.

Tu, Z., Poppe, R., & Veltkamp, R. C. (2016). Weighted local intensity fusion method for variational optical flow estimation. *Pattern Recognition*, *50*, 223-232.

Vedaldi, A., & Soatto, S. (2008). Quick shift and kernel methods for mode seeking. *Computer vision–ECCV 2008*, 705-718.

Veksler, O., Boykov, Y., & Mehrani, P. (2010). Superpixels and supervoxels in an energy optimization framework. *Computer Vision–ECCV 2010*, 211-224.

Wang, F., Jiang, Y. G., & Ngo, C. W. (2008, October). Video event detection using motion relativity and visual relatedness. In *Proceedings of the 16th ACM international conference on Multimedia* (pp. 239-248). ACM.

Wang, J., Zhang, Y., Lu, J., & Xu, W. (2012). A framework for moving target detection, recognition and tracking in UAV videos. *Affective Computing and Intelligent Interaction*, 69-76.

Wang, L., Hu, W., & Tan, T. (2003). Recent developments in human motion analysis. *Pattern recognition*, *36*(3), 585-601.

Wang, S. (2011, August). Vehicle detection on aerial images by extracting corner features for rotational invariant shape matching. In *Computer and Information Technology (CIT), 2011 IEEE 11th International Conference on* (pp. 171-175). IEEE.

Wren, C. R., Azarbayejani, A., Darrell, T., & Pentland, A. P. (1997). Pfinder: Real-time tracking of the human body. *IEEE Transactions on pattern analysis and machine intelligence*, *19*(7), 780-785.

Xiao, J., Cheng, H., Sawhney, H., & Han, F. (2010, June). Vehicle detection and tracking in wide field-of-view aerial video. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on* (pp. 679-684). IEEE.

Xiao, J., Yang, C., Han, F., & Cheng, H. (2008). Vehicle and person tracking in aerial videos. In *Multimodal Technologies for Perception of Humans* (pp. 203-214). Springer Berlin Heidelberg.

Yan, J., Cho, M., Zha, H., Yang, X., & Chu, S. M. (2016). Multi-graph matching via affinity optimization with graduated consistency regularization.*IEEE transactions on pattern analysis and machine intelligence*, *38*(6), 1228-1242.

Yan, J., Tian, Y., Zha, H., Yang, X., Zhang, Y., & Chu, S. M. (2013). Joint optimization for consistent multiple graph matching. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1649-1656).

Yan, J., Wang, J., Zha, H., Yang, X., & Chu, S. (2015). Consistency-driven alternating optimization for multigraph matching: A unified approach. *IEEE Transactions on Image Processing*, *24*(3), 994-1009.

Yang, L., Meer, P., & Foran, D. J. (2007, June). Multiple class segmentation using a unified framework over mean-shift patches. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on* (pp. 1-8). IEEE.

Yang, N. C., Chang, W. H., Kuo, C. M., & Li, T. H. (2008). A fast MPEG-7 dominant color extraction with new similarity measure for image retrieval.*Journal of Visual Communication and Image Representation*, *19*(2), 92-105.

Yilmaz, A., Shafique, K., & Shah, M. (2003). Target tracking in airborne forward looking infrared imagery. *Image and Vision Computing*, *21*(7), 623-635.

Yu, H., Chang, Y., Lu, P., Xu, Z., Fu, C., & Wang, Y. (2014). Contour level object detection with top-down information. *Optik-International Journal for Light and Electron Optics*, *125*(11), 2708-2712.

Yu, Q., & Medioni, G. (2009, June). Motion pattern interpretation and detection for tracking moving vehicles in airborne video. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* (pp. 2671-2678). IEEE. doi:10.1109/CVPRW.2009.5206541

Zand, M., Doraisamy, S., Halin, A. A., & Mustaffa, M. R. (2015). Texture classification and discrimination for region-based image retrieval. *Journal of Visual Communication and Image Representation*, *26*, 305-316. doi:10.1016/j.jvcir.2014.10.005

Zang, W., Lin, J., Zhang, B., Tao, H., & Wang, Z. (2011, June). Line-based registration for UAV remote sensing imagery of wide-spanning river basin. In*Geoinformatics, 2011 19th International Conference on* (pp. 1-4). IEEE. doi:10.1109/GeoInformatics.2011.5980864

Zass, R., & Shashua, A. (2008, June). Probabilistic graph and hypergraph matching. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on* (pp. 1-8). IEEE. doi:10.1109/CVPR.2008.4587500

Zhou, F., & De la Torre, F. (2016). Factorized graph matching. *IEEE transactions on pattern analysis and machine intelligence*, *38*(9), 1774-1789. doi:10.1109/TPAMI.2015.2501802

Zhou, H., Kong, H., Wei, L., Creighton, D., & Nahavandi, S. (2015). Efficient road detection and tracking for unmanned aerial vehicle. *IEEE Transactions on Intelligent Transportation Systems*, *16*(1), 297-309.

Zhou, X., Yang, C., & Yu, W. (2013). Moving object detection by detecting contiguous outliers in the low-rank representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(3), 597-610.

Zitova, B., & Flusser, J. (2003). Image registration methods: a survey.*Image and vision computing*, *21*(11), 977-1000. doi:10.1016/S0262-8856(03)00137-9

Zitova, B., Flusser, J., & Sroubek, F. (2005). Image registration: a survey and recent advances. In *IEEE ICIP*.