MOVING OBJECT DETECTION USING IMAGE REGISTRATION FOR A
MOVING CAMERA PLATFORM

SEYED ALI CHERAGHI

UNIVERSITI TEKNOLOGI MALAYSIA

MOVING OBJECT DETECTION USING IMAGE REGISTRATION FOR A MOVING CAMERA PLATFORM

SEYED ALI CHERAGHI

A project report submitted in partial fulfilment of the
requirements for the award of the degree of
master of Engineering (Electrical - Electronics & Telecommunications)

Faculty of Electrical Engineering
Universiti Teknologi Malaysia

JANUARY 2012

To My Loving and Caring Parents ...

ACKNOWLEDGEMENT

My first thanks is for my supervisor, Dr. Usman Ullah Sheikh, whose constant support, patience and unbounded enthuasiasm were of invaluable help. His devotion to the needs of the students and the encouragements have made working with him a true delight. Thanks for helping me to kickstart this research by providing insights and his work as reference.

My sincere appreciation to my fellow collegues in the Computer Vision, Video and Image Processing (CVVIP) Research Group, sharing the similar research interests. I value the camaraderie we share as well as the time they spent to share with me enriching ideas, as well as their concern.

My sincerest thanks to all those who have helped to make this thesis possible. Warmest regards to my parents and brothers for their seamless caring encouragement and moral support that has made this journey possible.

Seyed Ali Cheraghi

ABSTRACT

In this research work, an accurate and fast moving object detector that can detect all the moving objects from Unmanned Aerial Platform (UAV) is proposed. Because of the distance of the UAV to the objects and the movement of the platform, object detection is a challenging task. In order to achieve best results with low error, at first the camera motion has to be estimated so, by using the Rosten and Drummond technique the corners is detected and then by using the corners the camera motion is compensated. After motion compensation, by subtracting the registered frame from the reference frame all the moving objects are detected and extracted.

ABSTRAK

Penyelidikan ini mencadangkan algoritmauntuk mengesan objek bergerak dari platform udara tanpa pemandu (UAV) denganlebih pantas dan tepat. Disebabkan pergerakan platform serta jarak platform UAVdari objek bergerak, tugas mengesan adalah amat mencabar. Untuk mendapatkankeputusan pengesanan yang tepat dengan kadar ralat yang rendah, pertama sekali,pergerakan kamera hendaklah ditentukan. Teknik pengesanan bucu Rosten danDrummond digunakan bagi mengesan bucu-bucu dalam gambar dan seterusnya untukmembetulkan pergerakan kamera. Selepas operasi ini dilakukan, kerangka gambaryang telah dibetulkan ditolak dari kerangka gambar rujukan untuk mendapatkansemua objek bergerak.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Surveillance systems are one of the most important topics in security. Surveillance systems include many criteria which include monitoring. Monitoring consists of considering the situations that can happen in the area that is being monitored. For example in battle field, by monitoring the area all the movements of the enemys troops are captured and decisions can be made. Video surveillance systems have been used for a long time to monitor important places such as malls or museums. Video surveillance systems have three main generations.

First generation is based on analog sub systems which try to extend human eyes. In this generation the monitoring system just has to capture the videos and sent them to the displays in a control room and the decisions are made by humans. The main drawback of these systems is that they are based on the humans with limited abilities as the operators.

In the second generation the analog subsystems are combined with digital ones. So by using some of the improvements in digital video processing the accuracy of the systems have been increased. In the second generation, most of the works are concentrated on real time event detection.

Unlike the previous generations, in third generation the surveillance systems take the main control of monitoring and the humans just help to solve special circumstances. In this generation most of the decisions are made online and with high accuracy.

Different surveillance systems require different analyzing method. For example for surveillance systems that are used to monitor indoor environments such as malls, human detection is the main task for video analyzing, or, for traffic surveillance system, car detection is the first and one of the most important parts.

This research proposes a method that can be used in area monitoring from aerial videos. As it can be seen from Figure 1, the first step for automatic aerial video surveillance is moving object detection. In this step all the moving objects will be extracted. This step contains background modelling and foreground detection. The next step which can improve the result of detection part is object tracking, which create correspondence among detected objects in consecutive frames. Object classification is the third step which categorizes detected objects into various classes like human, vehicle, animal, etc.

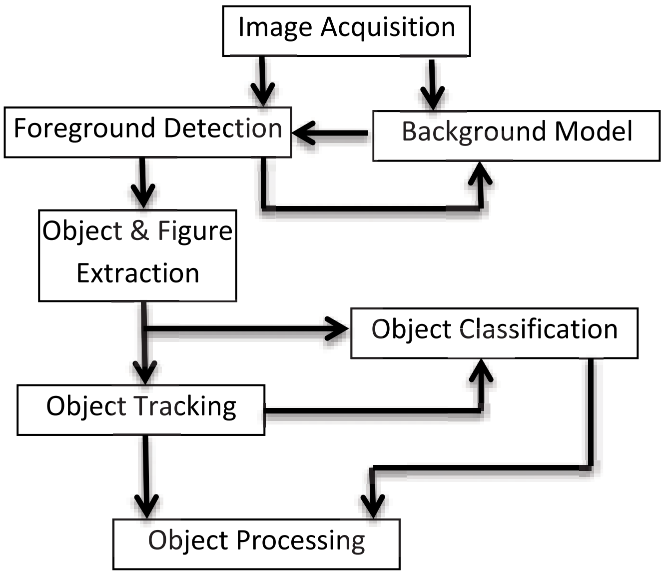

Figure 1: Automatic aerial video surveillance system.

Smart target detection, tracking and classification algorithms can be used in various applications and they are not limited to video surveillance only. Some examples are virtual reality, video compression, human machine interface and augmented reality. Some scenarios that smart surveillance systems and algorithms can be used are as follows:

Public and commercial security:

- i monitoring different places for crime prevention and detection
- ii patrolling critical places such as highways and railways for accident detection
- iii surveillance of properties and forests for fire detection and access control

Smart video data mining:

- i measuring traffic flow
- ii pedestrian congestion and athletic performance
- iii compiling consumer demographics in shopping centers and amusement parks
- iv extracting statistics from sport activities
- v counting endangered species

Law enforcement:

- i measuring speed of vehicles
- ii detecting red light crossings and unnecessary lane occupation Military security
- iii patrolling national borders
- iv measuring flow of refugees
- v monitoring peace treaties

As was mentioned earlier this research focuses on the public and commercial security. Aerial video processing is one of the topics in this area.It can be used for remote sensing, surveillance systems, military areas and many other places. In order to take the aerial videos, different platforms such as satellite, UAV and airplane can be used. UAV is one of the most important platforms which is used to take the aerial videos and these video sensors acts as the eyes of the system. Detecting the moving objects is one of the tasks that are used for aerial videos so by analyzing them various goals such as monitoring sensitive areas can be achieved.

Moving Target Detection (MTD) is one of the areas in computer vision that many researchers have investigated various approaches to improve it. MTD from stable platform has been well researched and numerous algorithms have been developed.

Many techniques such as Mixture of Gaussian (MOG), Optical Flow technique and Background Subtraction are used for detecting and tracking targets. In order to detect the moving objects from UAV the motion of the camera must be estimated and compensated. Then by using the detection algorithms the moving objects are detected.

0.1 Objectives

As was mentioned before, the first step for analysis of video is moving object detection. Moving object detection is an important part in motion perception of a mobile observatory system. It is very important for surveillance applications, smart moving objects tracking, automatic target recognition (ATR) and for many other applications [1, 2]. There are different reasons which change the scenes, such as the motion of the camera (ego motion), the object s movement or illumination changes. According to these reasons three possibilities for the camera and the objects can be defined:

i Stationary camera, moving objects

ii Moving camera, stationary objects

iii Moving camera, moving objects

The main goal of this project is to develop a system which is capable of detecting moving objects from the frames captured by a non-stationary camera. It attempts to make use of image registration technique in order to accurately detect all the moving objects.

0.2 Project Scope

In this project we aim to propose a detection algorithm which can detect all the moving objects in video streams which are captured from a moving camera. The propose algorithm does not detect the objects which do not have any distinction in color, texture or intensity besides, it will not consider if there is a rapid change in the background and also all the frames should be taken from the same camera continuously. The system that will be developed will be an off-line system.

0.3 Problem Statement

Problems concerning about this system is motion detection. In an image acquisition system on a moving platform, the entire scene which is captured is no longer static and this makes the moving object detection more difficult. Besides, after mounting the image acquisition on the moving platform the moving platform causes instabilities in image acquisition due to reasons such as disturbances which affect the actual motion of the moving objects.

0.4 Thesis Outline

Chapter 2 gives an insight to the existing moving object detection algorithms from moving and stable platforms which have been developed by various researchers. It include stable platform algorithms because object detection from stable platform plays an important role in this project.

Chapter 3 prepares the methodology of the proposed detection and, provides a short explaination for each of the main steps in the developed detection system.

Chapter 4 concentrate over the Moving objects detection details. This chapter tries to explain each step of detection while it provides the algorithms and the parameters that are needed for implementing them.

Chapter 5 is mainly devoted for demonstrating the experimental results and performance of the proposed detection algorithm on some aerial videos.

Chapter 6 deals with the summary and conclusions of the research. Besides, some realistic extensions as well as possible enhancements for the research are proposed.

LITERATURE REVIEW

Moving object detection can be divided in two categories, moving object detection from stationary platform and moving platform. Different approaches including hybrid algorithms have been investigated for stable platform. One of the popular methods that are used to detect the moving objects from stationary platform is adaptive background subtraction. Some researches [3,4] tried to prepare different kinds of this algorithm with various updating rules. The main drawback of this algorithm is missing moving objects in the scene that is just starting to move.

Another method for moving object detection is statistical background modeling. In this method each pixel is updated according to the statistics, then foreground and background pixels statistics are compared to each other. Although in moving object detection from the stationary platform, the movement of the platform does not exist, illumination changes or camera instabilities can make detection very hard, therefore, in order to consider the various cases, a good model for detection that can be updated frequently is required.

Mixture of Gaussian (MoG) is one of the best approaches that were proposed by Grimson and Stauffer [5, 6]. In this approach, each pixel in image is modeled as a mixture of Gaussians with 3, 4 and 5 Gaussian distributions beside, by using an on-line approximation all the parameters would be updated. By evaluating the mixture model of each pixel they can be categorized as the foreground or background pixels. By regarding each pixel value over the time, MoG can detect both lighting changes and objects which are moving.

Another method which is developed by Li *et al.* [7] uses the Bayes decision framework to detect all the moving objects from real-time complex video. According to the Bayes decision rule, all the pixels can be classified as the foreground or background therefore a data structure is used to learn and maintain the statistics which are belong to different feature vectors.

Another technique which is used to detect the motion in both stationary and moving platform is optical flow technique [8]. Related to this technique, motion of objects is considered as vectors that start or terminate at a pixel. According to this technique, there is no change for intensity values of the region, but rather just shifting from one position to the other one. Optical flow techniques include different variations such as differential techniques that use different kind of image intensity derivatives or region-based, feature-based, energy based, and phase-based techniques. [9]

In this research we assume the camera is placed over a moving platform, so we have the frames at slightly different time, from slightly different viewpoints and this is very similar to the definition of image registration [10]. There are various techniques for image registration, such as wavelets, the Fourier transform, optical flow, correlation methods, and feature based on approaches. Related to the image registration, there are four steps to overlay two images of reference image and sensed image over each other, which are feature detection, feature matching, transform model estimation and image resampling and transformation [10]. Barbara Zitova and Jan Flusser [11] prepared a comprehensive survey about image registration methods.

0.5 Background modelling

According to Toyama *et al.* [12] background modeling modules should follow a set of principles. Because background modeling is used as a part of a larger system so it should not try to extract the semantics of foreground objects on its own. The adaption of background model to sudden and gradual changes is very critical. Most background modelling techniques operate at the pixel-level. Toyama *et al.* [12] and Javed *et al.* [13] process images at the pixel, region, and frame levels. Cristani *et al.* [14] proposed model representation to represent the background, model initialization as the initialization of this model, and model adaptation for adapting the model to the background changes (e.g. illumination changes).

We have different situations for background modeling. In the first case the camera is fixed and the background stays relatively constant. We model the background as a single static image that can be identified and ignored. In another situation if the background is not actually constant, then modeling both the mean intensity at a pixel and its variance gives an adaptive tolerance for some variation in the background.

In the case that a scene contains motion as the part of the background, more tolerant models are required. This means that, a single multivariate Gaussian distribution should be measured with the mean and covariance as the parameters of this model and if a single Gaussian is insufficient to model the distribution of pixel values, a finite mixture of Gaussians (MoG) may be used instead.

In order to obtain sensitive detection Elgammal *et al.* [15] estimate the density function of pixel s distribution at any moment of time, considering only recent history information. In this model only the last sample of intensity value for a pixel is considered, therefore, the probability density function for the case that one pixel has a certain intensity value at a certain time can be estimated by using the kernel estimator [16].

In the case that a particular distribution of spatio-temporal image derivatives is emerged, points follow a constant optical flow. As the result, the image derivatives should fit the optic flow constraint equation: $I_{xu} + I_{yv} + I_t = 0$, for an optic flow vector (u, v) that is constant for background pixels. If there is no difference between the motion fields of the foreground and background, motion-based approaches cannot be successful [17]. Base on the time history of intensity at a particular pixel Toyama *et al.* [12] proposed to use Wiener filter as a linear predictor to explain periodic variations of pixel intensity. Based on the current frame intensity and the recent intensity values at a particular pixel the measurements are done.

Mittal and Paragios [18] classified the background adoption methods to either predictive or non-predictive. In predictive methods the scene is considered as a time series so after creating a dynamical model, the current input can be modeled based on past observations, whereas in non-predictive methods the order of the input observations is not considered, and this method usually is based on the probabilistic representation (pdf) of the observations at a particular pixel. One of the approaches for background detection is Adaptive Background detection. In Adaptive Background detection all the pixels in a video sequence are classified into either foreground or background. A broad classification of background subtraction techniques is given in [19]. Cheung and Kamath [20] divide these techniques in two categories, recursive or non-recursive.

A nonrecursive technique uses buffer for the previous N video frames, so it estimates the background image based on the buffer data for each pixel. Unlike nonrecursive technique, the recursive techniques do not maintain a buffer for background estimation. Recursive techniques update the models based on each of the frames that it takes so, by doing that the previous frames has its effects over the current frame and background. We can mention Frame differencing, Average filter, Median filtering, and some other techniques as the Non-recursive Techniques and Approximated Median Filter, Single Gaussian, Kalman filter and some others as the recursive Techniques.

0.5.1 Non-recursive Techniques

Simpler methods such as frame differencing which use a sliding window to estimate the background belongs to this category. In the first step of this approach, the previous frames are stored in the buffer and then related to the temporal variation of each pixel within the buffer the background would be estimated. Two methods which are belong to this category are:

0.5.1.1 Frame differencing

The commonly used technique for motion segmentation in static scene is background subtraction. In this technique by subtracting the current image pixel-by-pixel from a reference image and using the threshold parameter the foreground and the background are defined. After subtracting and thresholding some noises appear which can be reduced using morphological operations such as dilation, erosion, and closing.

As was mentioned before, the simplest technique for background estimation is frame differencing. In this method the previous frame is stored in the buffer and then according to this frame the background and foreground is detected. In this approach a pixel is defined as the foreground if:

$$|Frame_i - Frame_{i-1}| > Th \tag{1}$$

Which Th (Threshold) is a fixed value. Frame differencing depends on the object structure, speed, frame rate and global threshold, and it is very sensitive to the threshold.

0.5.1.2 Average or Median Filtering (MF)

This technique defines the background to be the median of the previous n frames [21, 22]. In this approach the background is updated according to:

$$B_{(x,y,t)} = Median\left\{I(x,y,t-i)\right\} \tag{2}$$

$$\Downarrow$$

$$|I_{(x,y,t)} - Median\{I(x,y,t-i)\}| > Th \tag{3}$$

where $I\epsilon\{0,1,2,...,n-1\}$

Median is rather fast, but consumes a lot of memory. In the average method the background update can be achieved with the following formula:

$$B_{t+1} = \alpha I_t + (1-\alpha)B_t \tag{4}$$

where α is adaptive learning rate and typically is 0.05. In this method two background corrections are applied:

a If a pixel is marked as foreground for more than m of the last M frames, then the background is updated as $B_{t+1} = I_t$. This correction is designed to compensate for sudden illumination changes and the appearance of static new objects.

b If a pixel changes state from foreground to background frequently, it is masked out from inclusion in the foreground. This is designed to compensate for fluctuating illumination, such as swinging branches.

0.5.2 Recursive Techniques

By building the probabilistic representation of the observations at each pixel on the frames, a background can be modeled (e.g. Kalman filtering or a Mixture of Gaussians based methods). In these methods, from each input frame one or several background models are created and updated.

0.5.2.1 Approximate Median Filter

Median filtering requires N previous frames for its calculation to achieve a background model and then by subtracting the result from the new frames the foregrounds are detected. The main drawback of this method is that it needs large amount of memory and many frames processing.

McFarlane and Schofield [23], proposed an effective recursive approximation of median filter. In this method, current frame s pixel is compared with the corresponding pixel in the background frame. If the pixel s value from current frame is larger than the background pixel s value then the background pixel is incremented by one while, if it has a lesser value, the background pixel is decremented by one. The performance of this algorithm is comparable with the performance of higher-complexity methods, besides, related to the computation and storage, this method has the same performance to the frame differencing method and it can adapt itself to the slow background changing. The disadvantage of this method is that it adapts slowly towards a large change in the background and it needs several frames to learn the new background.

0.5.2.2 Single Gaussian

This method is one of the easiest methods for estimating the background. In this method the average of the incoming new frames are calculated and then by subtracting the average result with the new incoming frames the result would be achieved. In this method in order to update the model, a single Gaussian model is used which by using a simple adaptive filter it would adapt itself to the slow changes. This method can detect all the changes that occurrs on the structure of the scene. The main feature of this method is that because it works on the distribution of pixels intensity, it does not consider the frames order or in other words, it ignores the observation order. The main drawback of this method is that, it does not adapt to changing background as its performance to sudden changes to illumination is very low.

0.5.2.3 Kalman Filter

Kalman filter is a group of mathematical equations that is used to minimize the squared error. This filter considers the estimations of the past, present and future so it can handle the cases that the nature of the modeled system is unknown. Cappellini and Karmann [24] proposed to use this filter to predict RGB values of background pixels based on previous frames. This widely-used recursive technique is used for tracking linear dynamic systems under Gaussian noise. In this approach by modeling the Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$ of each background pixels and updating the correspondence mean and variance, the changes in lighting or objects that are part of background can be accommodated.

0.5.2.4 Mixture of Gaussian

Mixture of Gaussian is one of the best approaches that were implemented by Grimson and Stauffer [5,6] to model the background. This approach does not consider any inter-dependencies between image pixels. This approach tries to model the background variation by using a number of Gaussian distribution. The characteristic of this approach is that this one is parametric so it can be updated without trying to store a large number of previous frames, therefore, it does not need a large buffer.

In this approach, each pixel in image is modeled as a mixture of Gaussians with 3, 4 and 5 Gaussian distributions beside, by using an on-line approximation all the parameters can be updated. By evaluating the mixture model of each pixel they can be categorized as the foreground or background pixels. By regarding each pixels value over the time, MoG can detect both lighting changes and objects which are moving. The drawback of this method is it is not robust to dynamic motion at the background.

0.5.2.5 Optical Flow

Optical flow (OF) is the distribution of apparent velocities of brightness movement in an image. Optical flow can arise when there is relative motion between the object and the observer. Optical Flow is an old concept and has been greatly exploited in the field of computer vision. In 1980, Horn and Schunck [25] formulated and computed Optical Flow for a sequence of images. This technique can give information on spatial arrangement of the object viewed and the rate of change of this arrangement [8].

Discontinuity in the OF enables the segmentation of the image to regions corresponding to different moving objects. A list of such techniques can be found in [25] but they all assume that the optical flow is already known. In the past the main limitation to such techniques was the high sensitivity to noise and high computational cost. These days, with the high processing speed of computers, OF is widely used. In 2007 [26], Xiao, Yang, Han and Cheng proposed an algorithm which couples a flow estimation process with background registration technique to generate a difference map.

Background and foreground moving objects are represented by different layers respectively. Each layer maintains an independent appearance and motion model. The authors have shown that the approach works well in vehicle tracking. However, for person tacking, the problem remains unsolved due to the small size of moving person, slow motion and low video contrast.

In 2007 [27], Zhang, Shi, Wang and Liu proposed a method to segment multiple rigid-body motions using Line Optical Flow. This algorithm can also work when the moving object has a homogeneous surface, provided that the object edges can be identified and used as straight line. The use of straight lines limits the approach in only identifying rigid motions. In this approach a K-means clustering method is used to build the final clusters. Although this method is able to deal with multiple rigid-body motions, it assumes that the number of moving objects (i.e, the number of clusters, K) is known a priority.

0.6 Image Registration

"Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. [10]" Given two images of a scene, the following steps are usually taken to register the images.

0.6.1 Preprocessing

In this step the images that are supposed to register over each other would be prepared. In other words, all the processing that contains noise removal, image enhancement and image segmentation. should be done in this step. This preprocessing can include changing the scale of one image to be appropriate for another image for doing registration. If the images have noises, in this step it is tried to remove these noises, or if it needs to segment images and fix them for feature extraction they have to be done here.

0.6.2 Feature Selection

In order to register two images over each other, correspondence between the images must be established. From the correspondence a transformation function will be created which can warp the sensed image over the reference image. The features used in image registration are corners, lines, curves, templates, regions, and patches.

Based on different situations different features can be used. For example because satellite images contains contours and regions the type of features that should be selected is different from the scenes that are taken from the image acquisition that is placed over a car. Barbara Zitov and Jan Flusser [11] prepared a comprehensive research about the features that can be selected in different situations.

0.6.3 Feature Correspondence

Feature correspondence can be done in two ways. In the first approach some features in one image are selected and it tries to find those features in the other image. Then it tries to find correspondence between them. In the second approach the images are searched for finding the features separately and after finding these features the correspondence between them would be made. If the features contain considerable information, the first method should be used unless, the second method is more appropriate.

0.6.4 Transformation Function and Resampling

After defining the correspondence between the two images, a transformation function is determined for resampling two images over each other. Transformation functions relate to the geometric difference between two images. Knowing the transformation function, the sensed image is resampled to the geometry of the reference image.

REFERENCES

1. Irani, M. and Anandan, P. A unified approach to moving object detection in 2d and 3d scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998. 20.
2. Lourakis, M., Argyros, A., Orphanoudakis, S. and Yz, S. C. O. Independent 3D Motion Detection Using Residual Parallax Normal Flow Fields, 1998.
3. Boult, T. E., Micheals, R., Gao, X., Lewis, P., Power, C., Yin, W. and Erkan, A. Frame-rate omnidirectional surveillance and tracking of camouflaged and occluded targets. *in Second IEEE International Workshop on Visual Surveillance*. 1999. 48–55.
4. A. Liu, A. H. and Guangyou, X. Moving object detection and tracking based on background subtraction. *Proceedings of SPIE, Wuhan, China*, 2001. 4554.
5. Stauffer, C. and Grimson, W. E. L. Adaptive Background Mixture Models for Real-Time Tracking. *CVPR*. 1999. 2246–2252.
6. Stauffer, C. and Grimson, W. E. L. Learning Patterns of Activity Using Real-Time Tracking. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2000. 22(8): 747–757.
7. Li, L., Huang, W., Gu, I. Y. H. and Tian, Q. Foreground object detection from videos containing complex background. *Proceedings of the eleventh ACM international conference on Multimedia*. New York, NY, USA: ACM. 2003, MULTIMEDIA 03. ISBN 1-58113-722-2. 2–10.
8. Barron, J. L., Fleet, D. J. and Beauchemin, S. S. Performance of optical flow techniques. *International Journal of Computer Vision*, 1994. 12: 43–77.
9. Liu, H., Hong, T.-H., Herman, M. and Chellappa, R. Accuracy vs. Efficiency Trade-offs in Optical Flow Algorithms. *Computer Vision and Image Understanding*. Academic Press. 1996. 271–286.
10. Goshtasby, A. A. *2-D and 3-D Image Registration: for Medical, Remote Sensing, and Industrial Applications*. Wiley-Interscience. 2005. ISBN 0471649546.

11. Zitov, B. and Flusser, J. Image registration methods: a survey. *Image and Vision Computing*, 2003. 21: 977–1000.

12. Toyama, K., Krumm, J., Brumitt, B. and Meyers, B. Wallflower: principles and practice of background maintenance. *Proceedings of the Seventh IEEE International Conference on Computer Vision*, 1999. 00: 255–261 vol.1.

13. Javed, O., Shafique, K. and Shah, M. A hierarchical approach to robust background subtraction using color and gradient information. *Workshop on Motion and Video Computing 2002 Proceedings*, 2002. pp2227: 22–27.

14. Cristani, M., Bicego, M. and Murino, V. Multi-level background initialization using Hidden Markov Models. *First ACM SIGMM international workshop on Video surveillance*. New York, NY, USA: ACM. 2003, IWVS 03. ISBN 1-58113-780-X. 11–20. doi:http://doi.acm.org/10.1145/982452.982455. URL `http://doi.acm.org/10.1145/982452.982455`.

15. Elgammal, A., Harwood, D. and Davis, L. Non-parametric model for background subtraction. *Frame-Rate Workshop, IEEE*. 2000. 751–767.

16. Scott, D. *Multivariate density estimation: theory, practice, and visualization*. Wiley series in probability and mathematical statistics: Applied probability and statistics. Wiley. 1992. ISBN 9780471547709.

17. Zhong, J. and Sclaroff, S. Segmenting Foreground Objects from a Dynamic Textured Background via a Robust Kalman Filter. *Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume 2*. Washington, DC, USA: IEEE Computer Society. 2003, ICCV 03. ISBN 0-7695-1950-4. 44–. URL `http://dl.acm.org/citation.cfm?id=946247.946643`.

18. Mittal, A. and Paragios, N. Motion-Based Background Subtraction Using Adaptive Kernel Density Estimation. 2004. 302–309.

19. Aach, T., Kaup, A., Mester, R., Aach, T., Kaup, A. and Mester, R. Statistical modelbased change detection in moving video. *Signal Processing*, 1993: 165–180.

20. Cheung, S.-C. S. and Kamath, C. Robust background subtraction with foreground validation for urban traffic video. *EURASIP J. Appl. Signal Process.*, 2005. 2005: 2330–2340. ISSN 1110-8657. doi:http://dx.doi.org/10.1155/ASP.2005.2330. URL `http://dx.doi.org/10.1155/ASP.2005.2330`.

21. Cutler, R. View-Based Detection and Analysis of Periodic Motion. *In International Conference on Patern Recognition*. 1998. 1–4.

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

23. McFarlane, N. J. B. and Schofield, C. P. Segmentation and tracking of piglets in images. *Mach. Vis. Appl.*, 1995. 8(3): 187–193.

24. Cappellini V., K. K. and A., B. Moving object recognition using and adaptive background memory. *Elsevier Science Publishers B.V.*, 1990: 289–307.

25. Horn, B. K. P. and Schunck, B. G. Determining Optical Flow. *ARTIFICAL INTELLIGENCE*, 1981. 17: 185–203.

26. Xiao, J., Yang, C., Han, F. and Cheng, H. Vehicle and Person Tracking in Aerial Videos. *CLEAR*. 2007. 203–214.

27. Zhang, J., Shi, F., Wang, J. and Liu, Y. 3D motion segmentation from straight-line optical flow. *Proceedings of the 2007 international conference on Multimedia content analysis and mining*. Berlin, Heidelberg: Springer-Verlag. 2007, MCAM 07. ISBN 978-3-540-73416-1. 85–94. URL `http://dl.acm.org/citation.cfm?id=1776594.1776611`.

28. Rosten, E. and Drummond, T. Fusing Points and Lines for High Performance Tracking. *IN INTERNATIONAL CONFERENCE ON COMPUTER VISION*. Springer. 2005. 1508–1515.

29. Cohen, I. and Medioni, G. G. Detecting and Tracking Moving Objects for Video Surveillance. *CVPR*. 1999. 2319–2325.

30. Karsli*, F. and Dihkan, M. Determination of geometric deformations in image registration using geometric and radiometric Measurements. *Academic Journals*, 2010.

31. Shi, J. and Tomasi, C. Good Features to Track. *1994 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 94)*. 1994. 593 – 600.

32. Harris, C. and Stephens, M. A Combined Corner and Edge Detector. *Proceedings of the 4th Alvey Vision Conference*. 1988. 147–151.

33. Smith, S. M. and Brady, J. M. SUSAN - A New Approach to Low Level Image Processing. *International Journal of Computer Vision*, 1995. 23: 45–78.

34. Gonzalez, R. C. and Woods, R. E. *Digital image processing*. Upper Saddle River, N.J.: Prentice Hall. 2008. ISBN 9780131687288 013168728X 9780135052679 013505267X.

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

36. Chen, S. and Leung, H. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 2004. 13(1): 220.

37. VIVID database. URL `https://www.sdms.afrl.af.mil/`.