THREE DIMENSIONAL INFORMATION ESTIMATION AND TRACKING FOR MOVING OBJECTS DETECTION USING TWO CAMERAS FRAMEWORK

GOH KAM MENG

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy (Electrical Engineering)

Faculty of Electrical Engineering
Universiti Teknologi Malaysia

MARCH 2015

Specially dedicated to *Mum and Dad*

I love you both.

ACKNOWLEDGEMENT

First and foremost, I would like to express my deepest gratitude to my supervisor Assoc. Prof. Syed Abdul Rahman for his continuous and immense support, financial aid, supervision, and encouragement for the entire duration of my study. I would also like to express my gratitude to Dr Musa for his constructive comments and suggestions during this research work.

Special indebtedness goes to my friends in Universiti Teknologi Malaysia (UTM), especially my colleagues at the CVVIP lab for their assistance, understanding, and moral support.

I wish to thank Ministry of Science, Technology and Environment (MOSTI), Ministry of Higher Education (MOHE), Research Management Centre (RMC) and UTM for financial supports and facilities.

Last but not least, I am also grateful to my family members for their support, patience, and continuous helps.

ABSTRACT

Calibration, matching and tracking are major concerns to obtain 3D information consisting of depth, direction and velocity. In finding depth, camera parameters and matched points are two necessary inputs. Depth, direction and matched points can be achieved accurately if cameras are well calibrated using manual traditional calibration. However, most of the manual traditional calibration methods are inconvenient to use because markers or real size of an object in the real world must be provided or known. Self-calibration can solve the traditional calibration limitation, but not on depth and matched points. Other approaches attempted to match corresponding object using 2D visual information without calibration, but they suffer low matching accuracy under huge perspective distortion. This research focuses on achieving 3D information using self-calibrated tracking system. In this system, matching and tracking are done under self-calibrated condition. There are three contributions introduced in this research to achieve the objectives. Firstly, orientation correction is introduced to obtain better relationship matrices for matching purpose during tracking. Secondly, after having relationship matrices another post-processing method, which is status based matching, is introduced for improving object matching result. This proposed matching algorithm is able to achieve almost 90% of matching rate. Depth is estimated after the status based matching. Thirdly, tracking is done based on *x-y* coordinates and the estimated depth under self-calibrated condition. Results show that the proposed self-calibrated tracking system successfully differentiates the location of objects even under occlusion in the field of view, and is able to determine the direction and the velocity of multiple moving objects.

ABSTRAK

Penentukuran, pemadanan dan pengesanan adalah faktor utama untuk mendapatkan maklumat 3D yang terdiri daripada kedalaman, arah dan halaju. Untuk mendapatkan kedalaman, parameter kamera dan pemadanan objek adalah dua input yang diperlukan. Kedalaman, arahan dan objek berpadan boleh dicapai dengan tepat jika kamera ditentukur dengan baik menggunakan penentukuran tradisional manual. Walau bagaimanapun, kebanyakan kaedah penentukuran tradisional manual adalah sukar untuk digunakan kerana penanda atau saiz sebenar sesuatu objek dalam dunia sebenar mesti disediakan atau dikenali. Penentukuran diri boleh menyelesaikan had penentukuran tradisional, tetapi tidak sesuai untuk memadankan objek. Cara-cara yang lain telah cuba untuk memadankan objek menggunakan maklumat visual 2D tanpa penentukuran, tetapi cara-cara itu mengalami ketepatan padanan yang rendah di bawah herotan perspektif yang besar. Kajian ini memberi tumpuan kepada pencapaian maklumat 3D di bawah penentukuran diri. Dalam sistem ini, pemadanan objek dan pengesanan dijalankan di bawah keadaan penentukuran diri. Tiga sumbangan diperkenalkan dalam kajian ini untuk mencapai objektif. Pertama, pembetulan orientasi diperkenalkan untuk mendapatkan matriks hubungan yang lebih baik untuk pemadanan objek semasa pengesanan. Kedua, selepas matriks hubungan satu lagi kaedah pasca-pemprosesan, pemadanan objek menggunakan status, diperkenalkan untuk meningkatkan pencapaian ketepatan. Algoritma yang dicadangkan mampu mencapai kadar sepadan hampir 90%. Kedalaman dianggarkan selepas pemadanan objek menggunakan status. Ketiga, pengesanan dilakukan berdasarkan koordinat *xy* dan kedalaman dianggarkan di bawah keadaan penentukuran diri. Keputusan menunjukkan bahawa sistem pengesanan yang dicadangkan berjaya membezakan lokasi objek walaupun dalam keadaan halangan dalam bidang pandangan, dan mampu untuk menentukan arah dan halaju objek bergerak.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3D	-	3 Dimension
2D	-	2 Dimension
SURF	-	Speeded-Up Robust Feature
SIFT	-	Scale-Invariant Feature Transform
ASIFT	-	Affine Scale-Invariant Feature Transform
ASURF	-	Affine Speeded-Up Robust Feature
MSER	-	Maximal Stable Extremal Regions
IBR	-	Intensity extrema-based detector
EBR	-	Edge based detector
MM-SIFT	-	Multi-resolution MSERs and SIFT
SUSAN	-	Smallest Univalue Segment Assimilating Nucleus
FAST	-	Features from accelerated segment test
FAST-ER	-	Features from accelerated segment test- Enhanced repeatability
RANSAC	-	Random Sample Consensus
LMedS	-	Least Median of Squares
LTS	-	Least Trimmed Squares
MLESAC	-	Maximum Likelihood Estimation SAmple Consensus
EMD	-	Earth Movers Distance
MAP	-	Maximum A Posterior
HT	-	Hough Transform
OC	-	Orientation Correction

LIST OF SYMBOLS

π	-	Pi
$\leq$	-	Less-than or equal to
$\geq$	-	Greater than or equal to
$^\circ$	-	Degree

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Surveillance systems have been widely used especially in the security fields such as access control in restricted areas, person-specific identification, anomaly detection, and for alarm systems [1]. This system can detect, monitor, and also analyse moving object behaviour in the field of view even under occlusive conditions. In addition, the object's velocity and direction can also be estimated easily for applications such as crime prevention and traffic incident detection. Today's surveillance system can be found everywhere in the cities, either in indoors or outdoors such as shopping centres, banks, outdoor car park areas, airports, or even in the streets. Since early 1980s, surveillance systems have been installed widely in public spaces for crime prevention in developed countries such as UK, USA and Australia. In Malaysia, the first surveillance camera was installed in 1966 [2]. In 1993, a directive was issued by the government to install surveillance cameras in all the car parks of public buildings [2]. In following years, the Ministry of Housing and Local Government initiated a Safe City Programme to install CCTV cameras for crime prevention in Kuala Lumpur (KL) under Strategy 2 of Target Hardening[2]. According to Malaysian Communications and Multimedia Commission (MCMC) report, snatch-theft cases dropped by 50% in Kuala Lumpur after the installation [2]. In 2012, Automatic Enforcement System (AES) was introduced to detect speeding vehicles and record traffic offenders [3].

1.2 Problem Statement

Generally, surveillance systems are used in recognizing objects, tracking objects from different views, and identifying 3D information of objects. Surveillance systems may come with a single camera or more. The multiple camera systems normally involve several cameras positioned at different angles looking at certain overlapping areas. Some systems can only provide 2D space information (x-y coordinates) and thus not capable to provide 3D information of an object. The system is further upgraded during research growth in these years. For many surveillance applications, 3D information, i.e. depth, direction, and velocity are important parameters [4] (such as location detection or crowd behaviour detection). As a consequence, much recent research has been focused on tracking using the 3D location of the targeted objects [5-8]. By using 3D information, more accurate results can be obtained and at the same time occlusion problems can be solved. In order to extract 3D information, calibration, matching and tracking are the major concern in the surveillance system and much research have been conducted to improve the traditional system.

The key to the acquiring 3D information is calibration. 3D information can only be estimated accurately if all cameras are calibrated (i.e. Intrinsic and extrinsic parameters of the camera are extracted) from which the 3D space or world coordinates can be computed. Some methods use single camera calibration, while others use multi camera calibration. Calibration techniques can be grouped into either traditional calibration or self-calibration. In traditional calibration, both intrinsic and extrinsic parameters are extracted. The relationship between world coordinates and pixel coordinates is established from the parameters. The corresponding object can then be matched easily even under large perspective distortion since in the traditional calibration, all cameras are connected with a single world coordinate system. Likewise, spatial matching using alignment can be done easily under the traditional calibration. However, most of the traditional calibration techniques are very inconvenient to use because manual labelling and the size of the object in the real world are needed as inputs. To overcome this limitation, a self-calibration technique has been developed. This process depends only on images captured by the camera using image 2D space *x-y* coordinates. However, the currently available self-calibration is only able to estimate the intrinsic parameters such as the focal length and the performance can still be improved. Since the extrinsic parameters cannot be extracted, 3D information cannot be found and the spatio-temporal feature between cameras cannot be matched.

There are several methods commonly used in matching corresponding objects in 2D space based on visual information without using any calibration or self-calibration [9-12]. However, these state-of-the-art techniques lack matching accuracy under large perspective distortion. Some researchers introduced a method to match the object in 2D space with large perspective distortion, but this requires longitude and latitude values as input, which can only be determined experimentally and inconveniently [13, 14]. Some other methods have been introduced using spatial information for matching. However, these methods require traditional calibration or manually selected matched points as input [15, 16]. Overall, corresponding identified objects from different views and intrinsic parameters are necessary inputs to estimate the depth of the object. In estimating the depth of the objects based on multiple images only from different views with large perspective distortion without using complex calibration, feature matching between cameras is essentially important

A more accurate tracking can be performed higher with the presence of 3D information [5-8]. Previous work shows that 3D tracker can yield 50% less error compared to 2D tracker [6]. However, most current surveillance systems are not able to estimate the 3D information of the moving object without traditional calibration. Thus, a 3D surveillance tracking method that estimates the depth, direction and velocity of the moving object based on self-calibration approach is equally important. Additionally, such a system requires a good matching method under large perspective distortion to determine the depth, direction and velocity.

Therefore, a system that is able to estimate distances of moving objects from the camera using self-calibration and feature matching should be addressed. This system should be able to find the corresponding objects from multiple scenes without any traditional calibration. Also, this system should be able to estimate directions and velocities of the moving objects based on videos.

1.3 Research Objectives

Based on the problem statement, the aims of this research are given as follows:

i. To estimate 3D information which is the depth of moving object based on 2D matching and self-calibration.
ii. To track and to estimate directions and velocities of multiple moving objects based on the estimated 3D information.

1.4 Research Scopes and Assumptions

Many researchers focus on different aspects of surveillance. In this thesis, the focus is in calibration, matching and tracking. Therefore, several scopes and assumptions have been established for this research.

1.4.1 Scopes

- The focus is on the tracking of multiple moving objects (human and vehicles)
- Two static cameras are used.

- At least 50% overlapping region of images in multiple cameras are considered.
- The 3D information considered are depth, direction and velocity.

1.4.2 Assumptions

- All the cameras are assumed to be located vertically above the moving objects.
- Baseline of cameras is assumed to be known.
- The system should be based only on the video frame without knowing any real world information such as the real size of the objects.

1.5 Research Contributions

To extract 3D information, focal length and corresponding points are needed. Based on these two key points, the contributions of this research are as follows:

i. The tracking system is established based on the *x-y* coordinates and estimated depth using linear prediction that can solve the occlusion problem. In this, the locations of multiple moving objects can be distinguished even if there is occlusion. Directions of the moving objects are estimated by comparing the ratio of left and right depth value while the velocity is estimated based on 2D *x-y* coordinates and estimated depths.

ii. A depth estimation system is developed based on a new corresponding points matching algorithm and an object matching process during tracking. The new algorithm is established by combining rectification, speeded-up robust feature (SURF), orientation correction, epipolar geometry, and also status based matching so that the matched objects can be found even under large perspective distortion. Depth is estimated from the matched objects with self-calibration.

iii. An orientation correction method is proposed to increase the number of correct matched points between two images during interest point matching. This algorithm is established based on the relative rotational angle between two images.

1.6 Research Methodology

To find the depth in the uncalibrated or self-calibrated condition, this research assumes that all cameras are on the same baseline, i.e. the distance between two cameras at the same level of position. Before the depth can be estimated, the relationship between each camera must also be established for the purpose of finding a corresponding object. To find the corresponding objects, the system must be able to overcome the affine transformation problem. The following is the flow of proposed system of this research:

- Images from different views must be rectified to become undistorted images. If the affine transform no longer exists in the image, the matching between images can be obtained.

- SURF is used to find the corresponding points between images. Since better matched points can produce a better fundamental matrix, orientation correction is introduced in this thesis to increase the number of correctly matched points. The orientation correction is computed based on the hypothesis that all features are rotated at the same angle.

- With a set of correctly matched points which is evenly distributed on the entire image, fundamental matrix can be generated for computing depth.

Since the depth can only be estimated if all cameras are on the same baseline, both images must be aligned so that they are on the same view plane.

- After the fundamental matrix is established and the images are aligned, the 3D information depth can be estimated with the presence of focal length from self-calibration using vanishing points.

- The object is tracked using 2D + depth linear prediction along with the estimated 3D information, and in this way the direction and velocity can be estimated.

1.7 Structure of Thesis

This thesis is organized as follows: Chapter One presents the introduction. Chapter Two discusses all the literature reviews related to the surveillance system. State-of-the-art techniques for all stages in the surveillance are discussed in this chapter. Chapter Three highlights the details of all the stages of the proposed technique. The experimental results based on the matching and tracking on the standard datasets are presented in Chapter Four. Last but not least, Chapter Five concludes the thesis along with suggestions for future work.

REFERENCES

[1] Weiming, H., Tieniu, T., Liang, W., and Maybank, S. A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*. 2004. 34(3): p.334-352.

[2] Malaysia, C. S. K. d. M. (2008). *Video Surveillance in Public Spaces* [Report]. Available: http://www.skmm.gov.my/skmmgovmy/files/attachments/Video_Surveillance_Public_Spaces.pdf

[3] Jamil, H. M., Shabadin, A., and Rahim, S. A. S. M. R. (2014). *The Effectiveness of Automated Enforcement System in Reducing Red Light Running Violations in Malaysia: Pilot Locations*. Available: http://www.miros.gov.my/web/guest/reports?p_p_id=101_INSTANCE_ssIa&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=rightbar&p_p_col_pos=1&p_p_col_count=2&_101_INSTANCE_ssIa_struts_action=%2Ftagged_content%2Fview_content&_101_INSTANCE_ssIa_redirect=%2Fweb%2Fguest%2Freports&_101_INSTANCE_ssIa_assetId=744668

[4] Rodrigues de Almeida, I. and Rosito Jung, C. Change Detection in Human Crowds. *Proceeding 2013 26th SIBGRAPI Conference on of Graphics, Patterns and Images (SIBGRAPI)*. 5-8 Aug. 2013. p.63-69.

[5] Lou, J., Yang, H., Hu, W. M., and Tan, T. Visual vehicle tracking using an improved EKF. *Proceeding of Asian Conference of Computer Vision*. 2002. p.296-301.

[6] Taghirad, H. D., Atashzar, S. F., and Shahbazi, M. Robust solution to three-dimensional pose estimation using composite extended Kalman observer and Kalman filter. *Computer Vision, IET*. 2012. 6(2): p.140-152.

[7] Tyagi, A., Keck, M., Davis, J. W., and Potamianos, G. (2006). *A Method for 3D Tracking Using Multiple Cameras* [Report]. Available: ftp://ftp.cse.ohio-state.edu/pub/tech-report/2006/TR79.pdf

[8] Salih, Y. and Malik, A. S. 3d tracking using particle filters. *2011 IEEE Instrumentation and Measurement Technology Conference (I2MTC)*. Hangzhou, China: IEEE. 2011. p.1-4.

[9] Mikolajczyk, K. and Schmid, C. Scale & Affine Invariant Interest Point Detectors. *International Journal of Computer Vision*. 2004. 60(1): p.63-86.

[10] Matas, J., Chum, O., Urban, M., and Pajdla, T. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*. 2004. 22(10): p.761-767.

[11] Tuytelaars, T. and Van Gool, L. Matching Widely Separated Views Based on Affine Invariant Regions. *International Journal of Computer Vision*. 2004. 59(1): p.61-85.

[12] Kadir, T., Zisserman, A., and Brady, M. An Affine Invariant Salient Region Detector. *8th European Conference on Computer Vision*. 11 May. Prague, Czech Republic: Springer Berlin Heidelberg. 2004. p.228-241.

[13] Morel, J.-M. and Yu, G. ASIFT: A New Framework for Fully Affine Invariant Image Comparison. *SIAM Journal on Imaging Sciences*. 2009. 2(2): p.438-469.

[14] Pang, Y., Li, W., Yuan, Y., and Pan, J. Fully affine invariant SURF for image matching. *Neurocomputing*. 2012. 85( ): p.6-10.

[15] Jens, P. PTZ camera network calibration from moving people in sports broadcasts. *Proceeding of 2012 IEEE Workshop on Applications of Computer Vision (WACV)*. Breckenridge, Colorado: IEEE. 2012. p.25-32.

[16] Zhou, Q. and Aggarwal, J. K. Object tracking in an outdoor environment using fusion of features and cameras. *Image and Vision Computing*. 2006. 24(11): p.1244-1255.

[17] Wang, X. Intelligent multi-camera video surveillance: A review. *Pattern recognition letters*. 2013. 34(1): p.3-19.

[18] Adel, M., Moussaoui, A., Rasigni, M., Bourennane, S., and Hamami, L. Statistical-Based Tracking Technique for Linear Structures Detection: Application to Vessel Segmentation in Medical Images. *IEEE Signal Processing Letters*. 2010. 17(6): p.555-558.

[19] Sung-Woo, S. and Kang-Hyun, J. 3D mapping and estimation from moving direction of indoor mobile robot using vanishing points. *Proceeding of 2009 ICCAS-SICE* 18-21 Aug. 2009. Fukuoka, Japan 2009. p.3504-3508.

[20] Nakabo, Y., Ishi, I., and Ishikawa, M. 3D tracking using two high-speed vision systems. *Proceeding of 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2002. Lausanne, Switzerland. 2002. p.360-365 vol.1.

[21] Wang, Y., *Performance analysis of 3-Dimensional Fingerprint Scan System*, University of Kengtucky, 2008.

[22] Chang, R., Yue, W., and Leman, K. Robust unmanned aerial vehicle camera self-calibration for surveillance applications. *Proceeding of 3rd Conference of Sensor Signal Processing for Defence (SSPD 2012)*. 25-27 Sept. London: IET. 2012. p.1-5.

[23] Rahim, H. A., Ahmad, R. B., Zain, A. S. M., and Sheikh, U. U. An adapted point based tracking for vehicle speed estimation in linear spacing. *2010 International Conference on Computer and Communication Engineering (ICCCE)* 11-12 May Kuala Lumpur, Malaysia. 2010. p.1-4.

[24] Li, R., Lewis, J. H., Jia, X., Gu, X., Folkerts, M., Men, C., Song, W. Y., and Jiang, S. B. 3D tumor localization through real-time volumetric x-ray imaging for lung cancer radiotherapy. *Medical physics*. 2011. 38(5): p.2783-2794.

[25] Rahim, H., Sheikh, U., Ahmad, R., and Zain, A. Vehicle velocity estimation for traffic surveillance system. *World academy of science*. 69 p.772-775.

[26] Barron, J. and Thacker, N., Tutorial: Computing 2D and 3D optical flow, Medical School, University of Manchester, Tutorial [Report], 2005.

[27] Czuba, T. B., Rokers, B., Huk, A. C., and Cormack, L. K. Speed and eccentricity tuning reveal a central role for the velocity-based cue to 3D visual motion. *Journal of neurophysiology*. 2010. 104(5): p.2886-2899.

[28] van der Hulst, A. E., Westenberg, J. J., Kroft, L. J., Bax, J. J., Blom, N. A., de Roos, A., and Roest, A. A. Tetralogy of Fallot: 3D Velocity-encoded MR

Imaging for Evaluation of Right Ventricular Valve Flow and Diastolic Function in Patients after Correction 1. *Radiology*. 2010. 256(3): p.724-734.

[29] Velipasalar, S. and Wolf, W. Multiple object tracking and occlusion handling by information exchange between uncalibrated cameras. *Proceeding of IEEE International Conference on Image Processing (ICIP)* 11-14 Sept. 2005. Genoa, Italy. 2005. p.II-418-21.

[30] Cai, Q., Sankaranarayanan, A., Zhang, Q., Zhang, Z., and Liu, Z. Real time head pose tracking from multiple cameras with a generic model. *Proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 13-18 June: IEEE. 2010. p.25-32.

[31] Zhihua, L., Xiang, T., Li, X., and Yaowu, C. Improved Object Classification and Tracking Based on Overlapping Cameras in Video Surveillance. *Proceeding of ISECS International Colloquium on Computing, Communication, Control, and Management (CCCM '08)* 3-4 Aug. . 2008. p.725-729.

[32] Martínez-del-Rincón, J., Herrero-Jaraba, E., Gómez, J. R., Orrite-Uruñuela, C., Medrano, C., and Montañés-Laborda, M. A. Multicamera sport player tracking with Bayesian estimation of measurements. *Optical Engineering*. 2009. 48(4): p.047201-047201-23.

[33] Qi, W., Li, F., and Zhenzhong, L. Review on camera calibration. *Proceeding of Chinese Control and Decision Conference (CCDC)*. IEEE. 2010. p.3354-3358.

[34] Shen, Z., Zhou, S., Miao, C., and Zhang, Y. Vehicle Speed Detection Based on Video at Urban Intersection. *Research Journal of Applied Sciences, Engineering and Technology*. 2013. 5(17): p.7.

[35] Boracchi, G., Caglioti, V., and Giusti, A. Single-image 3D reconstruction of ball velocity and spin from motion blur. *Proceeding of The 3rd International Conference on Computer Vision Theory and Applications*. 2008. p.22-29.

[36] Ab-Rahman, A., Sheikh, U., Maliki, M., Heriansyah, R., Singh, K., and Abu-Bakar, S. Vestro: Velocity estimation using stereoscopic vision. *Proceeding of 1st International Conference on Computers, Communications, & Signal Processing with Special Track on Biomedical Engineering, (CCSP)*. IEEE. 2005. p.120-124.

[37] Faugeras, O. D. and Hebert, M. The representation, recognition, and locating of 3-D objects. *The international journal of robotics research*. 1986. 5(3): p.27-52.

[38] Ganapathy, S. Decomposition of transformation matrices for robot vision. *Pattern Recognition Letters*. 1984. 2(6): p.401-412.

[39] Abdel-Aziz, Y. I. K. H. M. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. *Proceeding of Symposium on C lose-Rarige Phomgramrnerry*. VA: American Society of Photogrammetry. 1971.

[40] Faugueras, O. D. and Toscani, G., The Calibration Problem for Stereoscopic Vision, in *Sensor Devices and Systems for Robotics*. vol. 52, ed: Springer Berlin Heidelberg, 1989, pp. 195-213.

[41] Zhang, G., He, J., and Yang, X. Calibrating camera radial distortion with cross-ratio invariability. *Optics & Laser Technology*. 2003. 35(6): p.457-461.

[42] Tsai, R. Y. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation.* 1987. 3(4): p.323-344.

[43] Martins, H., Birk, J., and Kelley, R. Camera models based on data from two calibration planes. *Computer Graphics and Image Processing.* 1981. 17(2): p.173-180.

[44] Zhang, Z. Flexible camera calibration by viewing a plane from unknown orientations. *The Proceedings of the Seventh IEEE International Conference on Computer Vision* IEEE. 1999. p.666-673.

[45] Tuan Hue, T., Lu, S., and Zhang, J. Self-Calibration of Traffic Surveillance Camera using Motion Tracking. *The Proceeding of 11th International IEEE Conference on Intelligent Transportation Systems (ITSC).* 12-15 Oct. . Beijing, China: IEEE. 2008. p.304-309.

[46] Sung Chun, L. and Nevatia, R. Robust camera calibration tool for video surveillance camera in urban environment. *Proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).* 20-25 June 2011. Colorado, USA: IEEE. 2011. p.62-67.

[47] Sheikh, U. and Abu-Bakar, S. Three-dimensional vehicle pose estimation from two-dimensional monocular camera images for vehicle classification. *Proceeding of 6th WSEAS International Conference on Circuits, Systems, Electronics, Control & Signal Processing.* 1-3 November. Texas, USA. 2006. p.356-361.

[48] Pflugfelder, R. and Bischof, H. People tracking across two distant self-calibrated cameras. *Proceeding of IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS).* 5-7 Sept. London, UK. 2007. p.393-398.

[49] Kusakunniran, W., Hongdong, L., and Jian, Z. A Direct Method to Self-Calibrate a Surveillance Camera by Observing a Walking Pedestrian. *Digital Image Computing: Techniques and Applications, 2009. DICTA '09.* 1-3 Dec. . 2009. p.250-255.

[50] Gang-Qiang, Z., Ling, C., and Gen-Cai, C. A simple 3D face tracking method based on depth information. *Proceedings of International Conference on Machine Learning and Cybernetics.* 18-21 Aug. 2005. 2005. p.5022-5027

[51] Caprile, B. and Torre, V. Using Vanishing Points for Camera Calibration. *International Journal of Computer Vision.* 1990. 4(2): p.127-139.

[52] Guillou, E., Meneveaux, D., Maisel, E., and Bouatouch, K. Using vanishing points for camera calibration and coarse 3D reconstruction from a single image. *Visual Computer.* 2000. 16(7): p.396-410.

[53] Avinash, N. and Murali, S. Perspective geometry based single image camera calibration. *Journal of Mathematical Imaging and Vision.* 2008. 30(3): p.221-230.

[54] Frémont, V. and Chellali, R. Direct camera calibration using two concentric circles from a single view. *Proceeding of International Conference on Artificial Reality and Telexistence (ICAT).* Tokyo, Japan. 2002. p.93-98.

[55] Chen, Q., Wu, H., and Wada, T. Camera calibration with two arbitrary coplanar circles. *Proceeding of 8th European Conference on Computer Vision (ECCV).* 11-14 May. Prague, Czech Republic: Springer. 2004. p.521-532.

[56] Colombo, C., Comanducci, D., and Del Bimbo, A. Camera calibration with two arbitrary coaxial circles. *The Proceeding of 8th European Conference on*

Computer Vision (ECCV). 7 - 13 May Graz, Austria Springer. 2006. p.265-276.

[57] Wang, G., Tsui, H.-T., Hu, Z., and Wu, F. Camera calibration and 3D reconstruction from a single view based on scene constraints. *Image and Vision Computing*. 2005. 23(3): p.311-323.

[58] Hong, W., Yang, A. Y., Huang, K., and Ma, Y. On symmetry and multiple-view geometry: Structure, pose, and calibration from a single image. *International Journal of Computer Vision*. 2004. 60(3): p.241-265.

[59] Park, J. Quaternion-Based Camera Calibration and 3D Scene Reconstruction. *Proceeding of 4th International Conference on Computer Graphics, Imaging and Visualisation (CGIV)*. 14 - 16 August. Bangkok, Thailand. 2007. p.89-92.

[60] Whitehead, A. and Roth, G. Estimating intrinsic camera parameters from the fundamental matrix using an evolutionary approach. *EURASIP Journal on Advances in Signal Processing*. 2004. 2004(8): p.1113-1124.

[61] Ze-Tao, J., Wenhuan, W., and Min, W. Camera autocalibration from Kruppa's equations using particle swarm optimization. *Computer Science and Software Engineering, 2008 International Conference on*. IEEE. 2008. p.1032-1034.

[62] Hartley, R. I. Kruppa's equations derived from the fundamental matrix. *IEEE Transactions on pattern analysis and machine intelligence*. 1997. 19(2): p.133-135.

[63] Beynon, M. D., Van Hook, D. J., Seibert, M., Peacock, A., and Dudgeon, D. Detecting abandoned packages in a multi-camera video surveillance system. *Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS)*. 21-22 July Miami, FL, USA. 2003. p.221-228.

[64] Liu, R. J. Automatic surveillance camera calibration without pedestrian tracking. *Proceedings of the British Machine Vision Conference (BMCV)*. Sept. 2011. p.117-1.

[65] Davis, J. and Chen, X. Calibrating pan-tilt cameras in wide-area surveillance networks. *Proceedings of Ninth IEEE International Conference on Computer Vision (ICCV)*. 14-17 Oct. Nice, France: IEEE. 2003. p.144-149.

[66] Feng, G. Plane Rectification Using a Circle and Points from a Single View. *Proceeding of 18th International Conference on Pattern Recognition (ICPR)*. 20-24 Aug. Hong Kong, China: IEEE. 2006. p.9-12.

[67] Lingfeng, X., Au, O. C., Wenxiu, S., Yujun, L., Sung-Him, C., and Chun-Wing, K. Image rectification for single camera stereo system. *Proceedings of 18th IEEE International Conference on Image Processing (ICIP)*. 11-14 Sept. Brussels, Belgium. 2011. p.977-980.

[68] Zhang, Z. and He, L.-w., Whiteboard scanning and image enhancement, [Report], 2003.

[69] Manchikalapudi, V. Skew Correction and Localisation of Number Plate Using Hough Rectangular Transform. *International Journal of Computer Science and Technology*. 2011. 2

[70] Hartley, R. and Zisserman, A., *Multiple View Geometry in Computer Vision*: Cambridge University Press, 2003.

[71] Banks, J., Electrical, Q. U. o. T. S. o., Engineering, E. S., and Navigation, S. C. f. S., *A Taxonomy of Image Matching Techniques for Stereo Vision*: Space Centre for Satellite Navigation, School of Electrical and Electronic Systems Engineering, Queensland University of Technology, 1997.

[72] Porrill, J. and Pollard, S. Curve matching and stereo calibration. *Image and Vision Computing.* 1991. 9(1): p.45-50.

[73] Abbasi-Dezfouli, M., Freeman, T. G., Heipke, C., and Eder, K. Patch matching in stereo images based on shape. *Proceedings of ISPRS Commission III Symposium: Spatial Information from Digital Photogrammetry and Computer Vision.* Aug. Munich, Federal Republic of Germany. 1994. p.1-8.

[74] Lowe, D. G. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision.* 2004. 60(2): p.91-110.

[75] Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding.* 2008. 110(3): p.346-359.

[76] Harris, C. and Stephens, M. A Combined Corner and Edge Detection. *Proceedings of The Fourth Alvey Vision Conference.* 31 Aug -2 Sept. Manchester. 1988. p.147-151.

[77] Mikolajczyk, K. and Schmid, C. An Affine Invariant Interest Point Detector. *Proceedings of the 7th European Conference on Computer Vision-Part I.* London, UK: Springer-Verlag. 2002. p.128-142.

[78] Tao, C., Tan, Y., Cai, H., and Tian, J. Airport Detection From Large IKONOS Images Using Clustered SIFT Keypoints and Region Information. *IEEE Geoscience and Remote Sensing Letters.* 2011. 8(1): p.128-132.

[79] Liu, C., Yuen, J., and Torralba, A. SIFT flow: dense correspondence across scenes and its applications. *IEEE Transaction on Pattern Analysis and Machine Intelligence.* 2011. 33(5): p.978-94.

[80] Ling, H., Cheng, H., Ma, Q., Zou, F., and Yan, W. Efficient Image Copy Detection Using Multiscale Fingerprints. *IEEE Multimedia.* 2012. 19(1): p.60-69.

[81] Brox, T., Rosenhahn, B., Gall, J., and Cremers, D. Combined region and motion-based 3D tracking of rigid and articulated objects. *IEEE Transaction on Pattern Analysis and Machine Intelligence.* 2010. 32(3): p.402-15.

[82] Hasanuzzaman, F. M., Yang, X., and Tian, Y. Robust and Effective Component-Based Banknote Recognition for the Blind. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews).* 2012. 42(6): p.1021-1030.

[83] Soyel, H. and Demirel, H. Facial expression recognition based on discriminative scale invariant feature transform. *Electronics Letters.* 2010. 46(5): p.343.

[84] Juan, L. and Gwon, O. A Comparison of SIFT, PCA-SIFT and SURF. *International Journal of Image Processing (IJIP).* 2009. 3(4): p.143-152.

[85] Chen, M., Shao, Z., Li, D., and Liu, J. Invariant matching method for different viewpoint angle images. *Applied Optic.* 2013. 52(1): p.96-104.

[86] Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., and Gool, L. V. A Comparison of Affine Region Detectors. *International Journal of Computer Vision.* 2005. 65(1-2): p.43-72.

[87] Smith, S. M. and Brady, J. M. SUSAN-A New Approach to Low Level Image Processing. *Internatiol Journal of Computer Vision.* 1997. 23(1): p.45-78.

[88] Rosten, E. and Drummond, T. Machine learning for high-speed corner detection. *Proceedings of the 9th European conference on Computer Vision - Volume Part I*. Graz, Austria: Springer-Verlag. 2006. 430-443.

[89] Rosten, E., Porter, R., and Drummond, T. Faster and Better: A Machine Learning Approach to Corner Detection. *IEEE Transaction on Pattern Analysis and Machine Intelligence*. 2010. 32(1): p.105-119.

[90] Tian, Q., Sebe, N., Lew, M. S., Loupias, E., and Huang, T. S. Image retrieval using wavelet-based salient points. *Journal of Electronic Imaging*. 2001. 10(4): p.835-849.

[91] Loupias, E. and Sebe, N. Wavelet-based salient points: Applications to image retrieval using color and texture features. *Proceedings of International Conference on Advances in Visual Information Systems*. Lyon, France. 2000. p.223-232.

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

[93] Rousseeuw, P. J. Least median of squares regression. *Journal of the American statistical association*. 1984. 79(388): p.871-880.

[94] Čížek, P. and Víšek, J. Á., *Least trimmed squares*: Springer, 2000.

[95] Torr, P. H. and Zisserman, A. MLESAC: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*. 2000. 78(1): p.138-156.

[96] Frahm, J.-M. and Pollefeys, M. RANSAC for (quasi-) degenerate data (QDEGSAC). *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. NY, USA: IEEE. 2006. p.453-460.

[97] Chum, O., Matas, J., and Obdrzalek, S. Enhancing RANSAC by generalized model optimization. *Proceedings of 6th Asian conference on computer vision*. Jeju, Korea. 2004. p.812-817.

[98] Choi, S., Kim, T., and Yu, W. Performance evaluation of RANSAC family. *Journal of Computer Vision*. 2009. 24(3): p.271-300.

[99] Khan, S. and Shah, M. Consistent labeling of tracked objects in multiple cameras with overlapping fields of view. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2003. 25(10): p.1355-1360.

[100] Black, J. and Ellis, T. Multi camera image tracking. *Image and Vision Computing*. 2006. 24(11): p.1256-1267.

[101] Tan, T. N., Sullivan, G. D., and Baker, K. D. Recognizing Objects on the Ground-Plane. *Image and Vision Computing*. 1994. 12(3): p.164-172.

[102] Stein, F. and Medioni, G. Map-Based Localization Using the Panoramic Horizon. *IEEE Transactions on Robotics and Automation*. 1995. 11(6): p.892-896.

[103] Thompson, W. B., Henderson, T. C., Colvin, T. L., Dick, L. B., and Valiquette, C. M. Vision-based localization. *DARPA Image Understanding Workshop*. Citeseer. 1993. p.491-498.

[104] Cozman, F. and Krotkov, E. Automatic mountain detection and pose estimation for teleoperation of lunar rovers. *Experimental Robotics V*. 1998. 232 p.207-215.

[105] Stein, G. P. Tracking from multiple view points: Self-calibration of space and time. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 1999. CO, USA. 1999. p.527 Vol. 1.

[106] Lee, L., Romano, R., and Stein, G. Monitoring activities from multiple video streams: Establishing a common coordinate frame. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2000. 22(8): p.758-767.

[107] Black, J., Ellis, T., and Rosin, P. Multi view image surveillance and tracking. *Proceedings of IEEE Workshop on Motion and Video Computing*. 5-6 Dec. 2002. Orlando, Florida 2002. p.169-174.

[108] Jing-Ying, C., Tzu-Heng, W., Shao-Yi, C., and Liang-Gee, C. Spatial-temporal consistent labeling for multi-camera multi-object surveillance systems. *Proceedings of IEEE International Symposium on Circuits and Systems*. 18-21 May 2008. Washington, USA: IEEE. 2008. p.3530-3533.

[109] Nunziati, W., Sclaroff, S., and Del Bimbo, A. An invariant representation for matching trajectories across uncalibrated video streams. *Proceedings of the 4th International Conference on Image and Video Retrieval*. Singapore: Springer. 2005.

[110] Reilly, V., Idrees, H., and Shah, M. Detection and tracking of large number of targets in wide area surveillance. *Proceedings of the 11th European conference on computer vision conference on Computer vision: Part III*. Heraklion, Crete, Greece: Springer-Verlag. 2010. 186-199.

[111] Perera, A. G. A., Srinivas, C., Hoogs, A., Brooksby, G., and Wensheng, H. Multi-Object Tracking Through Simultaneous Long Occlusions and Split-Merge Conditions. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 17-22 June New York, USA: IEEE. 2006. p.666-673.

[112] Qin, W. and Yaonan, W. Multiple Moving Objects Tracking under Complex Scenes. *Proceedings of 6th World Congress on Intelligent Control and Automation.* 0-0 0. Dalian, China: IEEE. 2006. p.9871-9875.

[113] Clark, A. J., Green, R. D., and Grant, R. N. Perspective correction for improved visual registration using natural features. *Proceedings of 23rd International Conference on Image and Vision Computing New Zealand (IVCNZ)*. 26-28 Nov. 2008. New Zealand. 2008. p.1-6.

[114] Stephen, T. B. Disparity Analysis of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 1980. 2(4): p.333-340.

[115] Sun, C. A Fast Stereo Matching Method. *In Digital Image Computing: Techniques and Applications*. 1997. 95-100.

[116] Nayar, S. K. Shape from focus system. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 15-18 June IEEE. 1992. p.302-308.

[117] Bove Jr, V. M. Entropy-based depth from focus. *Journal of the Optical Society of America A.* 1993. 10(4): p.561-566.

[118] Brown, M. Z., Burschka, D., and Hager, G. D. Advances in computational stereo. *IEEE Transaction on Pattern Analysis and Machine Intelligence.* 2003. 25(8): p.993-1008.

[119] Dhond, U. R. and Aggarwal, J. K. Structure from stereo-a review. *IEEE Transactions on Systems Man and Cybernetics.* 1989. 19(6): p.1489-1510.

[120] Pentland, A. P. A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 1987.(4): p.523-531.

[121] Rajagopalan, A. and Chaudhuri, S. A variational approach to recovering depth from defocused images. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 1997. 19(10): p.1158-1164.

[122] Foix, S., Alenya, G., and Torras, C. Lock-in time-of-flight (ToF) cameras: a survey. *IEEE Sensors Journal.* 2011. 11(9): p.1917-1926.

[123] Weingarten, J. W., Gruener, G., and Siegwart, R. A state-of-the-art 3D sensor for robot navigation. *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Sendai, Japan: IEEE. 2004. p.2155-2160.

[124] Saxena, A., Schulte, J., and Ng, A. Y. Depth Estimation Using Monocular and Stereo Cues. *Proceedings of International Joint Conference on Artificial Intelligence*. Hydebrad, India. 2007.

[125] Tu, Q., Xu, Y., and Zhou, M. Robust vehicle tracking based on scale invariant feature transform. *International Conference on Information and Automation (ICIA)*. Changsha, China: IEEE. 2008. p.86-90.

[126] Lieberknecht, S., Benhimane, S., Meier, P., and Navab, N. A dataset and evaluation methodology for template-based tracking algorithms. *8th IEEE International Symposium on Mixed and Augmented Reality*. Florida, USA: IEEE. 2009. p.145-151.

[127] Yeoh, P. Y. and Abu-Bakar, S. A. R. Accurate real-time object tracking with linear prediction method. *Proceedings of International Conference on Image Processing*. Catalonia, Spain: IEEE. 2003. p.III-941-4 vol. 2.

[128] Comaniciu, D. and Meer, P. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2002. 24(5): p.603-619.

[129] Shan, C., Tan, T., and Wei, Y. Real-time hand tracking using a mean shift embedded particle filter. *Pattern Recognition.* 2007. 40(7): p.1958-1970.

[130] Xin, L., Kejun, W., Wei, W., and Yang, L. A multiple object tracking method using Kalman filter. *Proceedings of IEEE International Conference on Information and Automation (ICIA)*. 20-23 June 2010. Heilongjiang, China. 2010. p.1862-1866.

[131] Thrun, S. Particle filters in robotics. *Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc. 2002. p.511-518.

[132] Paletta, L., *et al.* Attention in mobile interactions: gaze recovery for large scale studies. *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. ACM. 2014. p.1717-1722.

[133] Park, S., Yu, S., Kim, J., Kim, S., and Lee, S. 3D hand tracking using Kalman filter in depth space. *Eurasip Journal on Advances in Signal Processing.* 2012.(1): p.1-18.

[134] Alamsyah, D. and Fanany, M. I. Particle filter for 3D fingertips tracking from color and depth images with occlusion handling. *Proceedings of International Conference on Advanced Computer Science and Information Systems (ICACSIS)*. 28-29 Sept. 2013. Kuta, Bali. 2013. p.445-449.

[135] Yang, H. and Sikdar, B. A protocol for tracking mobile targets using sensor networks. *Proceedings of the 1st IEEE International Workshop onSensor Network Protocols and Applications*. AK, USA: IEEE. 2003. p.71-81.

[136] Vaidyanathan, P. The theory of linear prediction. *Synthesis Lectures on Signal Processing.* 2007. 2(1): p.1-184.

[137] Mingzhong, L., Zhaozheng, Y., Thimgan, M. S., and Ruwen, Q. Track fast-moving tiny flies by adaptive LBP feature and cascaded data association. *Image Processing (ICIP), 2013 20th IEEE International Conference on*. 15-18 Sept. 2013. 2013. p.1172-1176.

[138] Geng, C. and Jiang, X. Face recognition using sift features. *16th IEEE International Conference on Image Processing (ICIP)*. Nov. Cairo, Egypt: IEEE. 2009. p.3313-3316.

[139] Su, J., Xu, Q., and Zhu, J. A scene matching algorithm based on SURF feature. *Proceedings of International Conference on Image Analysis and Signal Processing (IASP)*. 9-11 April. Zhejiang, China. 2010. p.434-437.

[140] Lu, X.-m., Wang, J.-b., and He, Z. An Improved Algorithm for Image Mosaic Based on Speeded-Up Robust Features. *Proceedings of International Conference on Management and Service Science (MASS)*. 24-26 Aug. Wuhan, China. 2010. p.1-4.

[141] Bing, H., Yongming, W., and Xiaozhi, J. Fast calculating feature point's main orientation in SURF algorithm. *Proceedings of International Conference on Computer, Mechatronics, Control and Electronic Engineering (CMCE)*. 24-26 Aug. 2010. p.165-168.

[142] Zhao, F., Huang, Q. M., Wang, H., and Gao, W. MOCC: A Fast and Robust Correlation-Based Method for Interest Point Matching under Large Scale Changes. *Eurasip Journal on Advances in Signal Processing*. 2010. 2010(1): p.1-16.

[143] Shang, S., Ding, R., Zheng, K., Jensen, C., Kalnis, P., and Zhou, X. Personalized trajectory matching in spatial networks. *The VLDB Journal*. 2014. 23(3): p.449-468.

[144] Sechidis, L. A., Patias, P., Tsioukas, V., . Low-level tracking of multiple objects. *Proceedings of The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Nov. 2002. p.6.

[145] *Second IEEE International Workshop on Performance Evaluation of Tracking and Surveillance* [Dataset]. Available: http://www.cvg.reading.ac.uk/PETS2001/pets2001-dataset.html

[146] *ACM Multimedia Grand Challenge 2010* [Dataset]. Available: http://www.cdvp.dcu.ie/tennisireland/TennisVideos/acm_mm_3dlife_grand_challenge/

[147] *IEEE International Workshops on Performance Evaluation of Tracking and Surveillance (PETS2009)* [Dataset]. Available: http://pets2009.net/

[148] Aksay, A., Kitanovski, V., Vaiapury, K., Onasoglou, E., Agapito, J. D. P. M., Daras, P., and Izquierdo, E. Robust 3d tracking in tennis videos. *Engage Summer School*. 2010.

[149] Nieto, R. M. and Sánchez, J. M. M. An automatic system for sports analytics in multi-camera tennis videos. *Proceedings of 10th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. Krakow, Poland: IEEE. 2013. p.438-442.

[150] Wang, H., Shen, J., Shen, J., and Chen, Z. Tracking Object by Logic Reasoning. *International Journal of Hybrid Information Technology*. 2012. 5(2):

[151] Leung, H. Joint estimation fusion and tracking of objects in a single camera using EM-EKF. *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics. 2013. p.885617-885617-10.

[152] Viola, P., Jones, M. J., and Snow, D. Detecting pedestrians using patterns of motion and appearance. *Proceedings of 9th IEEE International Conference on Computer Vision*. Nice, France: IEEE. 2003. p.734-741.

[153] Khan, S., Javed, O., and Shah, M. Tracking in uncalibrated cameras with overlapping field of view. *Proceedings of 2nd IEEE Workshop on Performance Evaluation of Tracking and Surveillance*. Dec. Kauai. 2001.

[154] Hu, W., Hu, M., Zhou, X., Tan, T., Lou, J., and Maybank, S. Principal axis-based correspondence between multiple cameras for people tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2006. 28(4): p.663-671.

[155] Yun, Y., Gu, I. Y.-H., and Aghajan, H. Maximum-likelihood object tracking from multi-view video by combining homography and epipolar constraints. *Proceedings of Sixth International Conference on Distributed Smart Cameras (ICDSC)* 30 Oct - 2 Nov. Hong Kong, China: IEEE. 2012. p.1-6.

[156] Khan, M. H., Valstar, M. F., and Pridmore, T. P. A Multiple Motion Model Tracker Handling Occlusion and Rapid Motion Variation. *Proceedings of British Machine Vision Conference*. Bristol. 2013.

[157] *Affine Covariant Feature* [Dataset]. Available: http://www.robots.ox.ac.uk/~vgg/research/affine

[158] Codreanu, V., Feng, D., Baoquan, L., Roerdink, J. B. T. M., Williams, D., Po, Y., and Yasar, B. GPU-ASIFT: A fast fully affine-invariant feature extraction algorithm. *Proceedings of International Conference on High Performance Computing and Simulation (HPCS)*. 1-5 July. Helsinki, Finland 2013. p.474-481.

[159] Zhou, H., Pan, Y., and Zhang, Z. A speeded-up affine invariant detector. *Proceedings of 5th International Congress on Image and Signal Processing (CISP)*. Sichuan, China: IEEE. 2012. p.401-406.

[160] Mishkin, D., Perdoch, M., and Matas, J. Two-view matching with view synthesis revisited. *Proceedings of 28th International Conference of Image and Vision Computing New Zealand (IVCNZ)*. 27-29 Nov. Wellington 2013. p.436 - 441

[161] Heinly. *Semper Dataset* [Dataset]. Available: http://www.cs.unc.edu/~jheinly/feature-evaluation/datasets.html

[162] Seo, J.-K., Hong, H.-K., Jho, C.-W., and Choi, M.-H. Two quantitative measures of inlier distributions for precise fundamental matrix estimation. *Pattern recognition letters*. 2004. 25(6): p.733-741.

[163] Moulon, P., Monasse, P., and Marlet, R. Adaptive Structure from Motion with a contrario model estimation. *Proceedings of The 11th Asian Conference on Computer Vision* 5-9 Nov. Daejeon, Korea: Springer. 2013. p.257-270.

[164] Jain, P. K. and Jawahar, C. Homography estimation from planar contours. *Proceedings of Third International Symposium on 3D Data Processing, Visualization, and Transmission*. 14-16 June. Chapel Hill, NC IEEE. 2006. p.877-884.

[165] Khmanee, C. and Nguyen, D. On the design of 2D Gabor filtering of fingerprint images. *Proceedings of First IEEE Consumer Communications and Networking Conference (CCNC)*. 5-8 Jan. Las Vegas, NV, USA IEEE. 2004. p.430-435.

[166] Huang, Z. and Leng, J. Analysis of Hu's moment invariants on image scaling and rotation. *Proceedings of 2nd International Conference on Computer Engineering and Technology (ICCET)*. 16-18 April. Chengdu IEEE. 2010. p.V7-476-V7-480.

[167] Rao, A. R., *A taxonomy for texture description and identification*: Springer Publishing Company, Incorporated, 2012.

[168] Shi, G., Xu, X., and Dai, Y. SIFT Feature Point Matching Based on Improved RANSAC Algorithm. *Proceedings of 5th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*. 26-27 Aug.: IEEE. 2013. p.474-477.

[169] Motai, Y., Kumar Jha, S., and Kruse, D. Human tracking from a mobile agent: optical flow and Kalman filter arbitration. *Signal Processing: Image Communication*. 2012. 27(1): p.83-95.

[170] Berclaz, J., Fleuret, F., Turetken, E., and Fua, P. Multiple Object Tracking Using K-Shortest Paths Optimization. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. 2011. 33(9): p.1806-1819.

[171] Hu, C., Chen, W., Chen, Y., and Liu, D. Adaptive Kalman filtering for vehicle navigation. *Journal of Global Positioning System*. 2003. 1(04): p.0.

[172] Marron, M., Garcia, J. C., Sotelo, M. A., Cabello, M., Pizarro, D., Huerta, F., and Cerro, J. Comparing a Kalman Filter and a Particle Filter in a Multiple Objects Tracking Application. *Intelligent Signal Processing, 2007. WISP 2007. IEEE International Symposium on*. 3-5 Oct. 2007. 2007. p.1-6.

[173] Magee, D. R. Tracking multiple vehicles using foreground, background and motion models. *Image and vision Computing*. 2004. 22(2): p.143-155.

[174] Chen, Z. Epipole Estimation under Pure Camera Translation. *Proceedings of 7th International Conference on Digital Image Computing: Techniques and Applications*. Dec. 2003. p.849-858.

[175] Luong, Q.-T. and Faugeras, O. D. On the determination of epipoles using cross-ratios. *Computer Vision and Image Understanding*. 1998. 71(1): p.1-18.

[176] Zhong, H. and Hung, Y. Conjugate epipole-based self-calibration of camera under circular motion. *Proceedings of 10th IEEE Conference on Mechatronics and Machine Vision in Practice*. 9-11 Dec. Hong Kong, China. 2003.

[177] Aguilera, D., Lahoz, J. G., and Codes, J. F. A new method for vanishing points detection in 3d reconstruction from a single view. *Proceedings of the ISPRS Working Group V/4 Workshop* 22-24 Aug. Venice, Italy: International Society of Photogrammetry and Remote Sensing (ISPRS). 2005.

[178] Kogecka, J. and Zhang, W. Efficient computation of vanishing points. *Proceedings of International Conference on Robotics and Automation (ICRA)*. IEEE. 2002. p.223-228.

[179] Baker, Z. L. H., Kurillo, G., and Bajcsy, R. Projective Epipolar Rectification for a Linear Multi-imager Array. *Proceedings of 5th International Symposium on 3D Data Processing, Visualization and Transmission*. Paris, France. 2010. p.17 - 20.

[180] Madrigal, F. and Hayet, J.-B. Multiple view, multiple target tracking with principal axis-based data association. *Proceedings of 8th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS)*. Klagenfurt, Austria: IEEE. 2011. p.185-190.