FAST AND PRECISE BINARY IMAGE DESCRIPTOR FOR AUTONOMOUS VEHICLE VISUAL LOCALIZATION

By

Ahmed Zakaria Abd El Khalek Bibars

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
in
Electronics and Communications Engineering

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2019

FAST AND PRECISE BINARY IMAGE DESCRIPTOR FOR AUTONOMOUS VEHICLE VISUAL LOCALIZATION

By

Ahmed Zakaria Abd El Khalek Bibars

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
in
Electronics and Communications Engineering

Under the Supervision of

Prof. Dr. Magdi Fikri Ragaey
Professor of Communications
Electronics and Communications Eng. Dep.
Faculty of Engineering, Cairo University

Dr. Mohsen Mohamed Mahroos
Assistant Professor
Electronics and Communications Eng. Dep.
Faculty of Engineering, Cairo University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2019

FAST AND PRECISE BINARY IMAGE DESCRIPTOR FOR AUTONOMOUS VEHICLE VISUAL LOCALIZATION

By

Ahmed Zakaria Abd El Khalek Bibars

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
in
Electronics and Communications Engineering

Approved by the Examining Committee:

Prof. Dr. Magdi Fikri Ragaey, Thesis Main Advisor

Prof. Dr. Mohsen Abd El Razak Rashwan, Internal Examiner

Prof. Dr. Mohamed Ibrahim Aladawy, External Examiner
- Helwan University.

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2019

Engineer: Ahmed Zakaria Abd El Khalek Bibars
Date of Birth : 1 / 10 / 1984
Nationality : Egyptian
E-mail : ahmed.bibars@pg.cu.edu.eg
Phone. : 01000083714
Address : Haram, Giza, 12111

Registration Date : 1/ 10 / 2013
Awarding Date : / / 2019
Degree : Doctor of Philosophy
Department : Electronics and Communications Engineering

Supervisors : Prof. Dr. Magdi Fikri Ragaey
Dr. Mohsen Mohamed Mahroos

Examiners :

Prof. Dr. Mohamed Ibrahim Eladawy (External Examiner)
(Helwan University)
Prof. Dr. Mohsen Abd El Razak Rashwan (Internal Examiner)
Prof. Dr. Magdi Fikri Ragaey (Thesis Main Advisor)

Title of Thesis :

Fast and Precise Binary Image Descriptor for Autonomous Vehicle Visual Localization

Key Words :

Vehicle visual localization; Loop closer detection; Binary image descriptors; Multi hypothesis Markov filter; Autonomous driving vehicles

Summary :

Autonomous vehicle self-localization by scene matching under extreme environmental changes has been among the most challenging problems in robotics and computer vision in the last few years. Large dynamic illumination changes during day hours and appearance changes between year seasons are the major difficulties of this problem. This thesis presents: 1) a new binary image descriptor addressed as ”Extended Local Difference Binary” (ELDB), which is an extension to the state-of the-art Local Difference Binary (LDB) image descriptor, and 2) a new algorithm for vehicle visual localization under extreme environmental changes that uses Multi-Hypothesis Markov Localization (MHML) as a data fusion algorithm, and uses ELDB for image matching. Experimental results presented in the thesis show that ELDB has better image matching accuracy and computational efficiency than LDB, and that the proposed vehicle visual localization algorithm is faster and more accurate than other state-of-the-art algorithms.

Disclaimer

I hereby declare that this thesis is my own original work and that no part of it has been submitted for a degree qualification at any other university or institute.

I further declare that I have appropriately acknowledged all sources used and have cited them in the references section.

Name: Ahmed Zakaria Abd El Khalek Bibars Date:

Signature:

Dedication

This thesis is dedicated to my parents.

Acknowledgements

I would like to thank my supervisors Prof. Dr. Magdi Fikri Ragaey and Dr. Mohsen Mohamed Mahroos for their guidance, encouragement, support, and constant patience. It was a great honor to work with them.

I would like to give special thanks to my family; Thanks for your support and encouragement all the time.

Table of Contents

Disclaimer i

Dedication ii

Acknowledgements iii

Table of Contents iv

List of Tables vii

List of Figures viii

Nomenclature xi

Abstract xii

1 Introduction 1
- 1.1 Introduction to visual navigation . . . 1
- 1.2 Evolution of visual localization algorithms and their limitations . . . 2
- 1.3 Thesis contributions . . . 4
- 1.4 Thesis organization . . . 5

2 Literature Review 7
- 2.1 Image features . . . 7
  - 2.1.1 Handcraft image descriptors . . . 7
    - 2.1.1.1 Local image descriptors . . . 8
    - 2.1.1.2 Local vector based image descriptors . . . 9
    - 2.1.1.3 Local binary image descriptors . . . 10
    - 2.1.1.4 Global image descriptors . . . 12
    - 2.1.1.5 Global vector based image descriptors . . . 12
    - 2.1.1.6 Global binary image descriptors . . . 13
  - 2.1.2 CNN image features . . . 14
- 2.2 Vehicle visual localization . . . 14
  - 2.2.1 Linear based methods . . . 15
    - 2.2.1.1 SeqSLAM . . . 15
    - 2.2.1.2 SMART . . . 16
    - 2.2.1.3 ABLE . . . 16
  - 2.2.2 Statistics based methods . . . 17
    - 2.2.2.1 Particle filter . . . 17
    - 2.2.2.2 Markov localization . . . 18

3 Condition Invariant Image Matching 19
- 3.1 Introduction . . . 19
- 3.2 Image pre-processing . . . 19
  - 3.2.1 Image down-sampling . . . 20

3.2.2 Image illumination normalization . . . 20
3.2.3 Sky blacking . . . 21
3.3 Original LDB . . . 21
3.4 LDB drawbacks . . . 23
3.5 Extended local difference binary . . . 25
3.5.1 Exponential grid-side size growth . . . 25
3.5.2 Modified representative functions . . . 27
3.6 Image matching . . . 27
3.7 Panoramic images . . . 28

4 Multi Hypothesis Markov Localization **31**
4.1 Markov localization . . . 31
4.2 Modified Markov localization filter . . . 32
4.3 Generating multiple hypotheses at street intersections . . . 33
4.4 Localization algorithm . . . 33

5 Experimental Work **38**
5.1 Experiment setup . . . 38
5.2 Datasets used in experiments . . . 39
5.2.1 Alderley dataset . . . 39
5.2.2 Surfers Paradise dataset . . . 39
5.2.3 Highway dataset . . . 39
5.2.4 CBD dataset . . . 40
5.3 Measuring image matching accuracy . . . 41
5.3.1 Alderley results . . . 42
5.3.2 Surfers Paradise results . . . 45
5.3.3 Highway results . . . 48
5.3.4 CBD results . . . 51
5.3.5 Summary of image matching results and comparison with other algorithms . . . 54
5.4 Accuracy under viewpoint change . . . 57
5.5 Descriptors processing time . . . 57
5.6 Measuring localization algorithm accuracy . . . 59
5.6.1 Highway results after using MHML . . . 60
5.6.2 CBD results after using MHML . . . 61
5.7 Discussion . . . 62

6 Conclusion **65**
6.1 Thesis summary . . . 65
6.2 Achieved results . . . 67
6.3 Future research directions . . . 68
6.3.1 Use ELDB as a local image features descriptors . . . 68
6.3.2 Use Binary-CNN based image features . . . 69
6.3.3 Use CNN semantic segmentation to discard image area covered by moving obstacles . . . 70

References **71**

Appendix A Random-LDB **77**

Appendix B Local ELDB Descriptor **81**

Appendix C AdaBoost Cell-Pairs Selection **83**

List of Tables

5.1 Percent of correctly matched images in different datasets using linear and exponential ELDB1 and ELDB2, LDB, BoW/SURF, and NetVLAD. . . . 54
5.2 Percent of correctly matched images as a function in lane difference. Computed using the Highway dataset. . . . 57
5.3 Results of different algorithms on Highway dataset, where η is the percentage of the correctly estimated locations, D is the maximum continuous uncertainty distance in meters, and F_{Search} and F_{Track} are the cycle rates at search and track phases, respectively. Algorithms without F_{Search} value in table are the algorithms that assume the initial vehicle location is known. . . . 61
5.4 Results of different algorithms on CBD dataset. . . . 63

B.1 Local ELDB matching results using Bikes image sequence. . . . 82
B.2 Local ELDB matching results using UBC image sequence. . . . 82
B.3 Local ELDB matching results using Trees image sequence. . . . 82

List of Figures

2.1 Matching two images using local extracted features [42]. 8
2.2 SIFT image descriptor. Left: grid is centered over the key point location with same orientation and scale of the key-point. Center: computing image local gradient over grid area. Right: computing cells gradient-histograms. 9
2.3 BRISK pattern [39]. 11
2.4 FREAK pattern [2]. 12
2.5 HOG. 13
2.6 Difference matrix used two match pair of image sequences. 16

3.1 Sample of day and night images for the same place, and their illumination normalized versions. 22
3.2 Histogram for the different C values of the sky pixels and the ground pixels. Each of the two curves are normalized by the total number of sky and ground pixels respectively. The overlap between the two classes is 8.6% [71]. 23
3.3 Illumination normalization process for a day-time image with sky covers large percent of its area, as in (a), generates noisy values in the sky area as shown in (b), which reduces matching accuracy. Sky blacking algorithm solves this problem by detecting sky area in the original image, and mask this area in the illumination normalized image, to generate the final image (C) with all sky area pixels equal zero. 24
3.4 The difference between linear and exponential growing grids. Randomly selected cell-pairs are shown in each grid of the two figures, where each cell-pair is marked by a line connecting the centers of the two cells. 26
3.5 Image matching cycle. 28
3.6 Projecting cylindrical co-ordinates to rectangular co-ordinates. 29

5.1 Satellite images for Highway dataset vehicle trajectory [51]. In this dataset the vehicle took the same trajectory in both database image acquisition trip and real-time tripe. The trajectory is marked by a blue line in the image. . 40
5.2 Satellite images for CBD dataset vehicle trajectories [51]. Database trajectory is shown in the left image marked by blue line, and real-time trajectory is shown in the right image marked by red line. 41
5.3 Accuracies of different descriptors as a function of descriptor length, for Alderley dataset. 42
5.4 Precision-recall curves of different descriptors, for Alderley dataset. . . . 43
5.5 Correct matches curves of Alderley dataset. Percent of correct matches of ELDB1$_{Exponential}$, ELDB2$_{Exponential}$, and LDB are 26.56%, 22.58%, and 20.77%, respectively. 43
5.6 Difference matrices of Alderley dataset using LDB and ELDB1$_{Exponential}$. The dark line represents correct matches appears more clearly in ELDB1$_{Exponential}$ figure, which means that ELDB1$_{Exponential}$ has better discrimination power. 44

5.7 Accuracies of different descriptors as a function of descriptor length, for Surfers Paradise dataset. . . . 45
5.8 Precision-recall curves of different descriptors, for Surfers Paradise dataset. 46
5.9 Correct matches curves of Surfers Paradise dataset. Percent of correct matches of $ELDB1_{Exponential}$, $ELDB2_{Exponential}$, and LDB are 78.67%, 76.28%, and 58.8%, respectively. . . . 46
5.10 Difference matrices of Surfers Paradise dataset using LDB and $ELDB_{Exponential}$. 47
5.11 Accuracies of different descriptors as a function of descriptor length, for Highway dataset. . . . 48
5.12 Precision-recall curves of different descriptors, for Highway dataset. . . . 49
5.13 Correct matches curves of Highway dataset. Percent of correct matches of $ELDB1_{Exponential}$, $ELDB2_{Exponential}$, and LDB are 51.65%, 45.56%, and 27.31%, respectively. . . . 49
5.14 Difference matrices of Highway dataset using LDB and $ELDB_{Exponential}$. 50
5.15 Accuracies of different descriptors as a function of descriptor length, for CBD dataset. . . . 51
5.16 Precision-recall curves of different descriptors, for CBD dataset. . . . 52
5.17 Correct matches curves of CBD dataset. Percent of correct matches of $ELDB1_{Exponential}$, $ELDB2_{Exponential}$, and LDB are 70.56%, 66.84%, and 42.66%, respectively. . . . 52
5.18 Difference matrices of CBD dataset using LDB and $ELDB_{Exponential}$. . . . 53
5.19 Sample of day/night frames that ELDB1 succeeded to match while LDB failed. The complexity of matching these images is mainly attributed to: bad visibility conditions due to a raining storm in the two image pairs in 5.19a, crowded traffic that causes variation between the images in 5.19b, and gradient inversion due to the presence of many artificial lighting sources in 5.19c, in addition to gradient inversion at sky-area boarders in all images. 55
5.20 Sample of panoramic day/night frames that ELDB1 succeeded to match while LDB failed. Each sub-figure shows query image in the top, and its database match in the bottom. The major difficulties in matching these images are: gradient inversion problem at sky area boarders and at locations of artificial lighting sources, and surrounding vehicles appear in the images that cause variation between the two images, like in 5.20b. Figures 5.20a and 5.20b are from Highway dataset, while 5.20c and 5.20d are from CBD dataset. . . . 56
5.21 Comparison between processing time of different descriptors at two different maximum grid resolutions. . . . 58
5.22 Precision-recall curves of Highway dataset after using MHML. . . . 60
5.23 Correct matches curves of Highway dataset after using MHML. . . . 61
5.24 Precision-recall curves of CBD dataset after using MHML. . . . 62
5.25 Correct matches curves of CBD dataset after using MHML. . . . 62

A.1 Image cells of random size and random locations. . . . 78
A.2 Accuracy as a function of descriptor length in bits, for Highway dataset. . . . 79
A.3 Accuracy as a function of descriptor length in bits, for CBD dataset. . . . 80

A.4 Descriptor processing time as a function of descriptor length in bits. RLDB has higher processing time, while ELDB$_{Linear}$ and LDB almost have identical curves. 80

Nomenclature

ABLE	Able for Binary-appearance Loop-closure Evaluation
BRIEF	Binary Robust Independent Elementary Feature
BRISK	Binary Robust Invariant Scalable Key-points
CCD	Charge Coupled Device
CNN	Convolutional Neural Networks
EKF	Extended Kalman Filter
ELDB	Extended Local Difference Binary
FAB-MAP	Fast Appearance-Based Mapping
FREAK	Fast Retina Key-point
GP-GPU	General-Purpose Graphics Processing Units
GPS	Global Positioning System
HOG	Histogram of oriented gradient
IMU	Inertial Measuring Unit
LDB	Local Difference Binary
LiDAR	Light Detection And Ranging
MHKF	Multi-Hypothesis Kalman Filter
MHML	Multi-Hypothesis Markov Localization
ORB	Oriented FAST and Rotated BRIEF
PF	Particle Filter
PROSAC	Progressive Sample Consensus
RANSAC	Random Sample Consensus
SAD	Sum of Absolute Difference
SeqSLAM	Sequence SLAM
SIFT	Scale Invariant Feature Transform
SMART	Sequence Matching Across Route Traversals
SURF	Seeded Up Robust Features
VLAD	Vector of Locally Aggregated Descriptor
YOLO	You Only Look Once