
# OBJECT TRACKING IN VIDEO AND LOCALIZATION

By

Somaia Mohamed Mahmoud Mohamed Mohamed

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
2016

# OBJECT TRACKING AND NAVIGATION BASED ON COMPUTER VISION

By
Somaia Mohamed Mahmoud Mohamed Mohamed

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. M.Fekri

……..

Professor of Electronics and Communications
Department
Faculty of Engineering, Cairo University

Prof. Dr. M.Hesham

……..

Professor of Engineering Math and Physics
Department
Faculty of Engineering, Cairo University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2016

# OBJECT TRACKING AND NAVIGATION BASED ON COMPUTER VISION

By
Somaia Mohamed Mahmoud Mohamed Mohamed

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. M.Fekri, Thesis Main Advisor

_____
Prof. Dr. Elsayed Eissa Hemayed, Internal Examiner

_____
Prof. Dr. Mohamed Ebrahim El-Adawy, External Examiner
Faculty of Engineering, Helwan University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2016

**Engineer's Name:** Somaia Mohamed Mahmoud
Mohamed Mahmoud
**Date of Birth:** 03./02/1980.
**Nationality:** Egyptian
**E-mail:** Smohamed32@gmail.com
**Phone:** 01003914061
**Address:** 521, Mogwra 3, Aked awal,
5th district, 6th October city, Giza.

Insert photo here

**Registration Date:** 01/ 03/2010
**Awarding Date:** / / 2016
**Degree:** Doctor of Philosophy
**Department:** Electronics and Communications Engineering
**Supervisors:**
Prof. Magdi Fekri
Prof. Mohamed Hesham

**Examiners:**
Prof. Mohamed Ebrahim EL-Adawy (External examiner) Faculty of Engineering, Helwan University.
Prof. ELsayed Eissa Hemayed (Internal examiner)
Porf. Magdi Fekri Mohamed (Thesis main advisor)

**Title of Thesis:**

Object Tracking and Navigation Based on Computer Vision

**Key Words:**
Computer Vision, Object tracking, SIFT, Localization, Visual Odometry

**Summary:**

This work proposes novel modify on algorithm for two different applications in computer vision namely localization based on vision and object tracking in video. The proposed work in this thesis is intended to improve the conventional system for object tracking through using wavelet transform and morphological operations. The new algorithm succeeds in removing the noise and smooth the image, as well as scale invariant feature transform (SIFT) to increase the number of features which, in turn, improves the matching stage. The proposed procedure has been applied on many consequent frames, under different circumstances. The implementation results show high accuracy and performance, compared with the other conventional procedures such as Wavelet-based object tracking algorithms, and SIFT-based object tracking algorithms.

On the other hand, the proposed navigation algorithm determines the location of the moving objects by using a computer vision system. A single camera system is used to select suitable features and a stereo camera system to obtain the locations of the object. This significantly improves the overall accuracy, compared with the conventional systems that mainly depend on maps to determine the location of the object by comparing the features of the image with the features of the map. The implementation results of the proposed object localization show high performance that achieved 10 meters' accuracy on the trajectory with a length of 165 meters.

# Acknowledgments

I wish to express my gratitude to my supervisor, Professor Magdi Fekti Mohamed, who was abundantly helpful and offered invaluable assistance, support and guidance. I was very pleased to work under his supervision.

Deepest gratitude is also due to my co-supervisor Professor Mohamed Hesham for his vision, proposed ideas, valuable discussion and constructive suggestions during my studies.

Special thanks also to my co-supervisor Professor Aboelmagd Noureldin for hosting and supporting me during my visit to Queen's University.

Special thanks to my colleagues in my research group at the Navigation Lab in RMC and Queen's University, for their wonderful friendship, support and valuable discussion.

I wish to express my love and gratitude to my beloved parents, brother and sister for their support and encouragement.

# Dedication

*To My Beloved Family*

*My Parents,*

*My Brother,*

*and My sister*

*(Thank You So Much)*

# Table of Contents

**Acknowledgments** ..... **i**
**Dedication** ..... **ii**
**Table of Contents** ..... **iii**
**List of Tables** ..... **v**
**List of Figures** ..... **vi**
**List of Equations** ..... **viii**
**Nomenclature** ..... **ix**
**List of symbols** ..... **xi**
**Abstract** ..... **xii**

**CHAPTER 1 : INTRODUCTION** ..... **1**

1.1 PROBLEM STATEMENT ..... 1
1.1.1 Object Tracking ..... 1
1.1.2 Navigation ..... 1
1.2 THESIS SCOPE ..... 2
1.3 THESIS ORGANIZATION ..... 3

**CHAPTER 2 : LITERATURE REVIEW** ..... **4**

2.1 OBJECT TRACKING ..... 4
2.1.1 Preprocessing and Postprocessing Stage ..... 7
2.1.1.1. Morphological Operations ..... 8
2.1.1.1.1. Dilation operation: ..... 8
2.1.1.1.2. Erosion Operation: ..... 8
2.1.1.2. Wavelet ..... 9
2.1.2 Segmentation Stage ..... 13
2.1.3 Background Subtraction ..... 13
2.1.4 Foreground Subtraction ..... 14
2.1.5 Object Representation ..... 15
2.1.6 Features ..... 18
2.1.6.1. SIFT Features ..... 18
2.1.6.2. SURF Features ..... 22
2.1.6.3. Harris Corner Detector ..... 24
2.2 NAVIGATION ..... 25
2.2.1 Visual Simultaneous Localization and Mapping (V-SLAM) ..... 26
2.2.2 Visual Odometry (VO) ..... 27
2.2.3 Motion Estimation ..... 28
2.2.4 Camera Calibration ..... 28
2.2.4.1. Calibration of Stereo Camera ..... 29
2.2.4.2. Pinhole Camera Model ..... 32
2.2.4.3. Single Camera System ..... 34
2.2.4.4. Stereo Camera System ..... 34
2.3 THE PERFORMANCE EVALUATION ..... 35

**CHAPTER 3 : PROPOSED ALGORITHMS** ..... **39**

3.1 OBJECT TRACKING ..... 39
3.1.1 The algorithm of Object Tracking ..... 39
3.2 LOCALIZATION ..... 42
3.2.1 The algorithm of Visual Odometry ..... 42

**CHAPTER 4 : EXPERIMENTAL SETUP AND RESULTS ..... 46**

4.1 TRACKING ALGORITHM EXPERIMENTS ..... 46
4.1.1 Wavelet Family Selection ..... 46
4.1.2 Object Tracking Experiments ..... 56
4.1.2.1. SIFT Features and Background Subtraction Experiments ..... 57
4.1.2.2. SIFT features and Morphological Operation Experiments ..... 59
4.1.2.3. SURF features and Morphological Operation Experiments ..... 60
4.2 LOCALIZATION EXPERIMENTS ..... 64
4.2.1 Calibration Step ..... 64
4.2.2 Localization experiments using single and stereo camera system ..... 65
4.2.2.1. Experiment 1 Drawing Trajectory in Image Plane using Single Camera ..... 67
4.2.2.2. Experiment 2 Drawing Trajectory in Image Plane using Single Camera ..... 67
4.2.2.3. Experiment 3 Short Trajectory using Stereo Camera System and SIFT Features ..... 68
4.2.2.4. Experiment 4 Short Trajectory using Stereo Camera System and SURF Features ..... 69
4.2.2.5. Experiment 5 Long Trajectory using Stereo Camera System and SIFT Features ..... 69

**CHAPTER 5 : DISCUSSION AND CONCLUSIONS ..... 71**

**References ..... 72**
**Appendix A: Object Tracking Code ..... 76**
**Appendix B: Navigation Code ..... 83**

# List of Tables

Table 2.1 The extracted features SIFT and SURF after salt and pepper noise in the images. ..... 23
Table 3.1: The procedure of motion detection ..... 40
Table 3.2: Proposed algorithm ..... 43
Table 4.1 Matching percent of DB2 wavelet on the dataset ..... 47
Table 4.2 Matching percent of DB3 wavelet on the dataset ..... 48
Table 4.3 Matching percent of DB4 wavelet on the dataset ..... 49
Table 4.4 Matching percent of DB5 wavelet on the dataset ..... 50
Table 4.5: Matching percent Sym2 wavelet with different levels ..... 51
Table 4.6: Matching percent of Sym4 wavelet on the dataset ..... 52
Table 4.7: Matching percent of Coif1 wavelet on the dataset ..... 53
Table 4.8 Matching percent of Coif2 wavelet on the dataset ..... 54
Table 4.9 Comparison of the best basis of families' wavelet ..... 54
Table 4.10: output performance of object tracking using vision-traffic data ..... 57
Table 4.11: output performance of tracking object A in vision-traffic data ..... 59
Table 4.12: output performance of tracking object A in vision-traffic data using SURF ..... 61

# List of Figures

Figure 2.1: Block diagram of DWT Color SIFT method [11] .....6
Figure 2.2: Sub-band decomposition of Image [11].....6
Figure 2.3: Generic background subtraction techniques and Post processing block diagram [15] .....7
Figure 2.4: Impact of dilation using 3 x3 structure elements [18] .....8
Figure 2.5: Impact of erosion using 3 x3 structure elements [18].....9
Figure 2.6 Procedure for the application of the wavelet decomposition [21] .....9
Figure 2.7 Haar wavelet filter [22].....10
Figure 2.8: Daubechies wavelet filter (DB2) [22].....10
Figure 2.9: Daubechies wavelet filter (DB3) [22].....11
Figure 2.10: Daubechies wavelet filter (DB4) [22].....11
Figure 2.11: Daubechies wavelet filter( DB5)[22].....11
Figure 2.12: Symlets wavelet filter (Sym2)[22].....12
Figure 2.13: Symlets wavelet filter (Sym4) [22].....12
Figure 2.14: Coiflets wavelet filter (Coif1) [22] .....12
Figure 2.15: Coiflets wavelet filter (Coif2) [22] .....13
Figure 2.16: Classification of segmentation algorithm .....14
Figure 2.17: Foreground mask for an outdoor scene.....15
Figure 2.18: Object representation methods [24].....16
Figure 2.19: Object representations. (a) Centroid point, (b) Multiple points, (c) Rectangular patch, (d) Elliptical patch, (e) Object articulated, (f) Object skeleton, (g) Object contour, (h) Object contour, (i) Object silhouette [24]. .....17
Figure 2.20: Difference of Gaussian in different scal [28].....19
Figure 2.21:Maxima and minima of the difference-of-Gaussian images are detected by comparing a pixel X to its 26 neighbors in 3x3 regions at the current and adjacent scales (marked with circles) [28].....20
Figure 2.22: Image gradient in (a) and Keypoint descriptors in (b) [28] .....21
Figure 2.23: Samples for the SIFT features matching two different views.....22
Figure 2.24: The difference between SFT and SURF [32] .....24
Figure 2.25: Classification of image points using eigenvalues .....25
Figure 2.26: Navigation methods .....26
Figure 2.27: General block diagram of Visual Odometry [4] .....28
Figure 2.28: Vehicle with stereo camera system.....30
Figure 2.29: Checkerboard calibration pattern.....30
Figure 2.30: Pairs of images are used for the calibration of a stereovision system from 1 to 6.....31
Figure 2.31: Pairs of images are used for the calibration of a stereovision system from 1 to 12.....31
Figure 2.32: Pinhole model and reference frames.....32
Figure 2.33: Camera systems (a) single camera (b) stereo system.....34
Figure 2.34: Stereo Camera System .....34
Figure 2.35: Spatial overlap between $GT_i$ and $ST_j$ in frame k .....36
Figure 3.1: Block diagram of the procedure.....41
Figure 3.2: Block diagram of the proposed algorithm .....44
Figure 3.3: Flowchart of the proposed algorithm.....45
Figure 4.1 Matching percent of DB2 with different levels .....47

Figure 4.2: Matching percent of DB3 with different levels .....48
Figure 4.3: Matching percent of DB4 with different levels .....49
Figure 4.4: Matching percent of DB5 with different levels .....50
Figure 4.5: Matching percent of Sym2 with different levels.....51
Figure 4.6: Matching percent of Sym4 with different levels.....52
Figure 4.7: Matching percent of Coif1 with different levels.....53
Figure 4.8 : Matching percent of Coif2 with different levels.....54
Figure 4.9: The best one here is DB2,2 because DB5,2 transition is greater than DB2,2 .....55
Figure 4.10: Sym2 is suitable.....56
Figure 4.11: Coif 1and 2.....56
Figure 4.12: DB2,2 and sym2,2 are more suitable, and the transition is less than the coif1,2.....56
Figure 4.13: object tracking outputs of vision-traffic data at t=0.5.....58
Figure 4.14: object tracking outputs of vision-traffic data at t = 0.9.....58
Figure 4.15: tracking object A on 100 frames using SIFT and morphological operation .....60
Figure 4.16: Tracking object A on 100 frames using SURF and morphological operation without t.....62
Figure 4.17: Tracking object A on 100 frames using SURF and morphological operation with t=0.9 .....63
Figure 4.18: Samples from the frames that used in navigation .....66
Figure 4.19: samples from frames used in the navigation.....66
Figure 4.20: Path of trajectory in image plane using single camera (a) and ground truth table in (b).....67
Figure 4.21: Path of trajectory in image plane using single camera (a) and from ground truth table in (b).....68
Figure 4.22: output of stereo camera system and SIFT compared with output of ground truth table.....68
Figure 4.23: Output of stereo camera system SURF compared with output of ground truth table.....69
Figure 4.24: output of stereo camera system comparable with output of ground truth table for 66 sec trajectory .....70

# List of Equations

(2.1).....15
(2.2).....15
(2.3) .....19
( 2.4) .....19
(2.5) .....19
(2.6).....20
(2.7).....20
(2.8) .....22
(2.9) .....24
(2.10).....32
(2.11).....32
(2.12).....32
(2.13).....32
(2.14).....33
(2.15).....33
(2.16).....33
(2.17).....33
(2.18).....33
(2.19).....35
(2.20).....35
(2.21).....35
(2.22).....36
(2.23) .....36
(2.24) .....36
(2.25) .....36
(2.26) .....36
(2.27) .....37
(2.28) .....37
(2.29) .....37
(2.30) .....37
(2.31) .....37
(2.32) .....37
(2.33) .....37
(2.34) .....38
(2.35) .....38
(2.36) .....38
(3.1) .....40
( 3.2) .....40
(4.1) .....57

# Nomenclature

| | |
|---|---|
| **2-D** | Two-Dimensions |
| **3-D** | Three-Dimensions |
| **CDT** | Correct Detected Track |
| **Coif1** | Coiflets wavelet base 1 |
| **Coif2** | Coiflets wavelet base 2 |
| **CUPT** | Coordinate Update |
| **DB2** | Daubechies wavelet base 2 |
| **DB3** | Daubechies wavelet base 3 |
| **DB4** | Daubechies wavelet base 4 |
| **DB5** | Daubechies wavelet base 5 |
| **DR** | Detection Rate |
| **DWT** | Discrete Wavelet Transform |
| **FAR** | False Alarm Rate |
| **FAT** | False Alarm Track |
| **FN** | False Negative |
| **FP** | False Positive |
| **GPS** | Global Positioning System |
| **GT** | Ground Truth |
| **HH** | High High subbands of Wavelet transform |
| **HL** | High Low subbands wavelet transform |
| **HOG** | Histogram of Oriented Gradients |
| **HSV** | (**H**ue, **S**aturation, **V**alue) |
| **ICA** | Independent Component Analysis |
| **IDWT** | Inverse Discrete Wavelet Transform |
| **IMU** | Inertial Measurement Unit |
| **INS** | Inertial Navigation Sensor |
| **JPEG** | Joint Photographic Experts Group |
| **LBP** | Local Binary Patterns |
| **LH** | Low High sub band of Wavelet transform |
| **LIDAR** | Light Detection And Ranging |
| **LL** | Low Low subbands of Wavelet transform |
| **PCA** | Principle Component Analysis |
| **RANSAC** | Random sample consensus |
| **RDWT** | Redundant Discrete Wavelet transform |
| **SFM** | Structure From Motion |
| **SIFT** | Scale Invariant Features Transform |
| **SLAM** | Simultaneous Localization and Mapping |
| **ST** | System Truth |
| **SURF** | Speed Up Robust Features |
| **SVD** | Singular Value Decomposition |
| **Sym2** | Symlets wavelet base 2 |
| **Sym4** | Symlets wavelet base 4 |

| | |
|---|---|
| **TDF** | Track Detection Failure |
| **TF** | total number of frames in the video sequence |
| **TG** | total number of frames for the ground truth objects |
| **TN** | True Negative |
| **TP** | True Positive |
| **TRDR** | Tracker Detection Rate |
| **VBN** | Vision Based Navigation |
| **VO** | Visual Odometry |
| **V-SLAM** | Visual Simultaneous Localization and Mapping |
| **WT** | Wavelet Transform |

# List of symbols

| | |
|---|---|
| Alph_c | Skew of camera |
| cc | principle point of camera |
| Fc | Focal Length of camera |
| Kc | distortion of camera |
| $I_t(x,y)$ | Image in Spatial domain |
| $B_t$ | Background model |
| T | Experimental Threshold |
| $\mu_d$ | Mean |
| $\lambda_d$: | Standard deviation |
| $T_s$ | Foreground experimental threshold |
| L | scale space of an image |
| G | scale Gaussian |
| D | Difference of Gaussian |
| m | gradient magnitude |
| Ɵ | Gradient orientation |
| H | Hessian matrix |
| $L_{xx}$ | convolution of the second order derivative of Gaussian with an image in x direction |
| $L_{xx}$ | convolution of the second order derivative of Gaussian with an image in y direction |
| λ | Eigen values |
| R | The measure of corner response |
| tr | Trace of matrix |
| Det | determinant of matrix |
| A | spatial overlap between ground truth and system truth |
| O | Binary variable between ground truth and system truth |
| $T_{ov}$ | Threshold |
| $TR_{ov}$ | arbitrary threshold |
| d | Distance |