



# **The Concept of Evolutionary Computing for Robust Surgical Endoscope Tracking and Navigation**



**Ying Wan**

School of Computing and Communications  
University of Technology, Sydney

This dissertation is submitted for the degree of  
*Doctor of Philosophy in Computer Science*

Faculty of Engineering and  
Information Technology

April 2017



I would like to dedicate this thesis to my remarkable husband, my adorable son, my loving parents, and the memory of my grandmother who just passed away December, 2015.



## **Declaration**

I hereby declare that except specific references, the content of this dissertation is original and has not been submitted in whole or in part for consideration for any other degree or qualification from any other universities. This dissertation contains nothing which is the outcome of work that has been done in collaboration with others, except as specified in the text and acknowledgments. In general, this dissertation totally consists of about 65,000 words including appendices, bibliography, footnotes, tables, equations, and 150 figures.

Ying Wan  
April 2017



## **Acknowledgements**

I am using this opportunity to express my gratitude to thank Professors Xiangjian He and Xiongbiao Luo who are my principle and co-supervisors, respectively, for their many suggestions, constant encouragement, and support during my PhD candidature. Moreover, I am sincerely grateful to them for sharing their truthful and illuminating vision on my work and have the honor of studying and working with them in the past three years and six months.

I certainly appreciate the support of the International Research Scholarship (IRS) provided by the Faculty of Engineering and Information Technology (FEIT), University of Technology, Sydney (UTS). I absolutely appreciate the financial support for my living in the PhD study.

I wish to thank my fellow colleagues and the staffs of the faculty for providing various assistance for the completion of this research work. In particular, Qiang Wu, Wenjing Jia, Min Xu, Sheng Wang, Minqi Li, Tao Zeng, and David Du, for their invaluable supports.

Last but not least, I would like to thank my husband and my parents for their patient, understanding, and wide support. This PhD thesis could not have been available without their encouragements and financial assistance during the past three years and six months.





## Abstract

Navigated endoscopy is generally agreed to be the next generation of interventional or surgical endoscopy. It usually combines pre- and intra-operative imaging information to guide physicians during endoscopic procedures. However, endoscope three-dimensional motion tracking that spatially and temporally synchronizes various sensory information still remains challenging for developing different endoscopic navigation systems. To navigate or track the surgical endoscope, three modalities of sensory information are utilized in endoscopic procedures: (1) preoperative images, i.e., three-dimensional CT images, (2) two-dimensional video sequences from the endoscopic camera, and (3) location measurements, attaching an electromagnetic sensor at the endoscope distal tip for measuring the temporal endoscope movement. In this respect, endoscope tracking and navigation aims to fuse these various modalities information to accurately and robustly locate or fly through the endoscope at any interest of regions. Unfortunately, fusing the multimodal information is still an open issue due to the information incompleteness, e.g., image artifacts, tissue deformation, and sensor output inaccuracy in computer assisted endoscopic interventions.

This thesis work focuses on fusing the multimodal information for accurate and robust endoscope tracking and navigation. A novel framework of multimodal information fusion is proposed to use evolutionary computing for endoscopic navigation systems. Several main contributions of this dissertation are clarified as follows. First, the concept of evolutionary computing was initially introduced to assist minimally invasive endoscopic surgery. Next, this work modified two evolutionary algorithms of particle swarm optimizer and differential evolution and proposed an enhanced particle swarm optimizer (EPSO) and observation-driven adaptive differential evolution (OADE). EPSO can adaptively update evolutionary parameters in accordance with spatial constraints and the current observation. OADE performs a new mutation operation for DE methods by integrating the current observation of sensor measurements and camera images, which can control the perturbation velocity and the direction of each individual during evolution, to enhance the DE performance. Additionally, the improved evolutionary computing algorithms are applicable to computer vision tasks, e.g., object tracking, motion estimation, and stochastic optimization.

The experimental results demonstrate that the proposed evolutionarily computed endoscopic tracking and navigation approaches in this dissertation provide a more accurate and robust endoscopic guidance framework than state-of-the-art methods. Based static phantom data validation, the average guidance accuracy of the EPSO framework was about 3.0 mm, its average position smoothness was 1.0 mm, and its average visual quality was improved to 0.29. By evaluating on a dynamic phantom, the OADE approach reduces the tracking error from 3.96 to 2.89 mm, improves the tracking smoothness from 4.08 to 1.62 mm, and increases the visual quality from 0.707 to 0.741.

In conclusion, the concept of evolutionary computation is a promising strategy to improve endoscopic tracking and navigation for minimally invasive surgery. The validation demonstrated its effectiveness to improve the guidance accuracy, visual quality, and tracking smoothness during endoscopic surgery. Future work includes surgical data validation, real-time processing, and translation to clinical applications.

**Keywords**

Minimally invasive surgery — Bronchoscope — Endoscopy — Endoscopic navigation — Endoscopic video processing — Computer-assisted intervention — 2D/3D registration — Pre- and intra-operative imaging — Electromagnetic tracking — Evolutionary computing — Particle swarm optimizer — Differential evolution

# Table of contents

<b>List of figures</b>	<b>xv</b>
<b>List of tables</b>	<b>xxi</b>
<b>1 Minimally invasive endoscopic surgery</b>	<b>1</b>
1.1 Minimally invasive surgery . . . . .	2
1.2 Surgical endoscopy . . . . .	3
1.2.1 Diagnosis and Staging . . . . .	3
1.2.2 Therapy and Treatment . . . . .	4
1.3 Navigated Bronchoscopy Overview . . . . .	5
1.4 Related research field . . . . .	6
1.5 Dissertation organization . . . . .	6
<b>2 Endoscope tracking and navigation</b>	<b>9</b>
2.1 State of the art . . . . .	9
2.2 Current problems and challenge . . . . .	10
2.3 New concept: Evolutionary computing . . . . .	11
2.4 Definition and formulation . . . . .	12
2.4.1 Endoscopic camera geometry . . . . .	13
2.4.2 Parameterization . . . . .	13
2.5 Multimodal information fusion . . . . .	14
<b>3 Calibration and registration</b>	<b>21</b>
3.1 Coordinate Definitions . . . . .	21
3.1.1 Bronchoscopic Camera Geometry . . . . .	22
3.2 Hand-eye calibration . . . . .	25
3.2.1 Tsai's method . . . . .	25
3.3 EM-CT registration . . . . .	26
3.3.1 Adaptive marker-free registration . . . . .	27

3.3.2	Experiments . . . . .	37
3.3.3	Result . . . . .	39
3.3.4	Discussion . . . . .	41
3.3.5	Conclusions . . . . .	43
<b>4</b>	<b>Enhanced particle swarm optimization</b>	<b>53</b>
4.1	Methods and Materials . . . . .	53
4.1.1	Multimodal information . . . . .	55
4.1.2	Particle Swarm Optimization . . . . .	55
4.1.3	Enhanced particle swarm optimization . . . . .	56
4.1.4	Application to endoscopic guidance . . . . .	59
4.1.5	Validation . . . . .	62
4.2	Results . . . . .	64
4.3	Discussion . . . . .	66
4.3.1	Effectiveness . . . . .	66
4.3.2	Potential Limitations . . . . .	66
4.4	Conclusions . . . . .	67
<b>5</b>	<b>Observation-driven differential evolution</b>	<b>75</b>
5.1	Purpose . . . . .	76
5.1.1	Camera 3-D Motion Representation . . . . .	77
5.1.2	3-D Virtual Bronchial Tree Model . . . . .	79
5.2	Differential Evolution . . . . .	80
5.2.1	Mutation . . . . .	80
5.2.2	Crossover . . . . .	81
5.2.3	Selection . . . . .	81
5.2.4	Remarks on DEs . . . . .	82
5.3	Observation-Driven Adaptive Differential Evolution . . . . .	83
5.3.1	Method Overview . . . . .	83
5.3.2	Preprocessing . . . . .	84
5.3.3	Initialization and Randomization . . . . .	85
5.3.4	New Mutation Operation . . . . .	85
5.3.5	Fitness Computation . . . . .	88
5.3.6	Camera Pose Determination . . . . .	89
5.4	Experiments . . . . .	90
5.5	Results . . . . .	91
5.6	Discussion . . . . .	93

---

5.6.1	Effectiveness . . . . .	93
5.6.2	Limitations . . . . .	94
5.7	Conclusion . . . . .	95
<b>6</b>	<b>Conclusions and open remarks</b>	<b>113</b>
6.1	Conclusions . . . . .	113
6.2	Benefits to other fields . . . . .	115
6.3	Potential limitations . . . . .	115
6.4	Promising directions . . . . .	116
6.5	Closing remarks . . . . .	118
	<b>References</b>	<b>121</b>
	<b>Appendix A List of publications</b>	<b>127</b>



# List of figures

1.1	Examples of diagnostic and therapeutic ancillary tools usually used during interventional bronchoscopy (Images courtesy of Olympus, Europe). . . . .	4
2.1	Composite view of an endoscope tracking and navigation system that contains four windows providing with useful information: Endoscopic video sequences (upper left), 2D virtual rendering images (lower left) generated from previously acquired CT data and corresponding to video images, a bird's view (upper right) displaying the endoscope trajectory, and a slice view (lower right) showing an axial slice, in which a cross cursor marks the current endoscope location. All windows were synchronized to provide critical visual and structural information during endoscopic intervention. . . . .	15
2.2	Camera geometry with three coordinate systems . . . . .	16
2.3	Various spatial transformations among different coordinate systems involved in endoscopic navigation systems . . . . .	17
2.4	General flowchart of evolutionary computing . . . . .	18
2.5	Multimodal information including the CT images used to generate 2D virtual rendering images, endoscopic video images, and positional sensor measurements involved in an electromagnetically guided endoscopic procedure . . . . .	19
2.6	Relationships of different coordinate systems in electromagnetically navigated endoscopic procedures . . . . .	19
3.1	Several coordinate systems are defined when using EMT-based methods during bronchoscopic navigation. . . . .	22
3.2	Pinhole camera model in the camera coordinate system. $C$ denotes the camera center and is the coordinate origin. The principal point $(p_x, p_y)$ is the intersection of the principal axis and the image plane that is located in front of the focal length. . . . .	23
3.3	Relationship between the world and camera coordinate systems. . . . .	24

3.4	Hand-eye calibration setup from camera pose estimation and EM sensor outputs. . . . .	26
3.5	CT images for the airway tree volume segmentation . . . . .	28
3.6	Automatically segmented airway volume from CT images . . . . .	29
3.7	Extracted airway centerlines from the segmented airway . . . . .	30
3.8	The model drawing for calibrating the spatial relationship between the ET sensor and the endoscope tip center . . . . .	31
3.9	The calibration model was produced by using a 3-D printer . . . . .	32
3.10	Insert the endoscope with an integrated ET sensor into the calibration model	33
3.11	Attached an ET sensor at the endoscope distal tip and calibrate the endoscope tip center . . . . .	34
3.12	Generated multiple points on basis of sensor measurement $\mathbf{p}_i$ and endoscope tip center $\hat{\mathbf{p}}_i$ , and assign centerline $\mathbf{c}_k$ to transformed point ${}^{ct}\mathcal{T}_{et}\mathbf{p}_i$ . . . . .	35
3.13	Obtained projected point $\mathbf{q}_i$ in terms of distance $d({}^{ct}\mathcal{T}_{et}\mathbf{p}_i, \mathbf{c}_k)$ . . . . .	36
3.14	Static phantom (a) and endobronchial ultrasound endoscope with an ET (AURORA) sensor fixed at its distal tip (b). . . . .	44
3.15	Dynamic phantom (a) and endoscope with an ET (Ascension) sensor fixed at its distal tip (b) . . . . .	45
3.16	Plotted fiducial registration error $f_e$ in static phantom validation (Experiment 6)	46
3.17	Plotted distance to centerlines in static phantom validation (Experiment 4) .	46
3.18	Plotted target tracking error $t_e$ in static phantom validation (Experiment 3) .	47
3.19	Transformed points were plotted along assigned centerlines in static phantom validation (Experiment 14). Points ( <i>green</i> ) transformed by our method were closer to assigned centerlines ( <i>blue</i> ) than points ( <i>red</i> ) transformed by Deguchi et al. [1]. . . . .	47
3.20	Comparison of positions estimated by Deguchi et al. [1] and our method in static phantom validation (Experiment 2). Positions at <i>blue</i> points were obtained using marker-based ET-CT transformation ${}^{ct}\mathcal{T}_{et}^*$ , whose fiducial error was about 0.81 mm. Positions at <i>green</i> points estimated by our method were almost overlapping <i>blue</i> points which were far from positions at <i>red</i> points tracked by Deguchi et al. [1]. The tracking accuracy of our method is significantly better than Deguchi et al. [1]. . . . .	48
3.21	Plotted fiducial registration error $f_e$ in dynamic phantom validation (Experiment 7) . . . . .	48
3.22	Plotted distance to centerlines in dynamic phantom validation (Experiment 15)	49
3.23	Plotted target tracking error $t_e$ (Experiment 12) in dynamic phantom validation	49



3.24	Examples of comparison of tracking results of using different methods on Experiments 5 and 8 during dynamic phantom validation: Plotted endoscope positions estimated by Deguchi et al. [1] ( <i>red</i> points) and our method ( <i>green</i> points) against marker-based estimations ( <i>blue</i> points). . . . .	50
3.25	Examples of comparison of tracking results of using different methods on Experiments 5 during dynamic phantom validation. First column show frame numbers selected uniformly every 100 frames, and second column displays corresponding videos images, and other columns show virtual images generated from tracking results using different methods of Klein et al. [2], Deguchi et al. [1], and ours. The proposed method performs much better than other two approaches. . . . .	51
3.26	Examples of comparison of tracking results of using different methods on Experiments 5 during dynamic phantom validation. TThe first column show frame numbers selected uniformly every 100 frames, and the second column displays corresponding videos images, and other rows show virtual images generated from tracking results using different methods of Klein et al. [2], Deguchi et al. [1], and ours. The proposed method performs much better than other two approaches. . . . .	52
4.1	Multimodal information including the CT images used to generate 2-D virtual rendering images, endoscopic video images, and positional sensor measurements involved in an electromagnetically guided endoscopic procedure.	54
4.2	Population propagation to obtain new particle states. The <i>red</i> , <i>green</i> , and <i>blue</i> points (or solid lines) denote the current particle $\mathbf{x}_{i,j}$ , the local individual best particle $\mathbf{p}_{i,j}$ , and the global all best $\mathbf{g}_{i,j}$ at iteration $j$ , respectively. The <i>red</i> , <i>green</i> , and <i>blue</i> dash lines display vectors $\omega\mathbf{v}_{i,j}$ , $\lambda_1\eta_1(\mathbf{p}_{i,j} - \mathbf{x}_{i,j})$ , and $\lambda_2\eta_2(\mathbf{g}_{i,j} - \mathbf{x}_{i,j})$ (Eq. 4.2) to determine vector $\mathbf{v}_{i,j+1}$ (the short <i>white</i> dash line) combined with $\mathbf{x}_{i,j}$ to compute new particle $\mathbf{x}_{i,j+1}$ (the long <i>white</i> dash line) in Eq. 4.3. Vector $\omega\mathbf{v}_{i,j}$ depends on the current observation (Eq. 4.4). .	57
4.3	Our endoscopic guidance framework using EPSO for the multimodal information integration. . . . .	60
4.4	Detect image patches with specific structures from camera image $\mathbf{I}_k$ for the similarity or the fitness computation. A <i>yellow</i> square denotes a selected patch with its center ( <i>green</i> dot). . . . .	68
4.5	Experimentally determined particle and iteration numbers: $M = 20$ and $N = 8$ .	69
4.6	Plotted the guidance accuracy of the five methods on Data 3. . . . .	70
4.7	Position and orientation smoothness of the five methods on ten experiments.	71

4.8	Comparison of visual quality and computational time of the five methods. . . . .	72
4.9	Visualized video and virtual images to investigate the visual quality of different methods. First column shows frame numbers selected uniformly at every 100 frames, and second column shows their corresponding real images. Other columns display virtual images generated from tracking results using the methods discussed above. Our method displays much better quality. . . . .	73
5.1	CT image segmentation for obtaining a 3-D bronchial tree model where a virtual camera flies through and generates 2-D virtual images by different camera poses. . . . .	78
5.2	Illustration of crossover step in DE. . . . .	81
5.3	Processing flowchart of our OADE tracking method. . . . .	84
5.4	Illustration of our mutation strategy to generate mutant vector $\mathbf{V}_{i,j}^k$ in accordance with three perturbations including observation $\Omega_i(\mathbf{E}^k - \mathbf{E}^{k-1})$ and two difference vectors of $F_i^b(\mathbf{X}_{best,j}^k - \mathbf{X}_{i,j}^k)$ and $F_i^r(\mathbf{X}_{r_1^i,j}^k - \mathbf{X}_{r_2^i,j}^k)$ . . . . .	86
5.5	An example of one bronchoscopic image with structural information . . . . .	87
5.6	Extract structural regions in one bronchoscopic image to calculate similarity during fitness computation step. One <i>yellow</i> square denotes one patch and a <i>green</i> point is its center. . . . .	97
5.7	Experimentally determined population size: $P = 25$ . . . . .	99
5.8	Experimentally determined generation number $G = 2$ or $3$ . . . . .	100
5.9	Plotted tracking error of six methods evaluated in Experiment 20. . . . .	102
5.10	Plotted tracking smoothness of six methods evaluated in Experiment 6. . . . .	103
5.11	Plotted visual quality of six methods evaluated in Experiments 8 and 14. . . . .	104
5.12	Comparison of different similarity measures and optimization frameworks for 3-D bronchoscope motion estimation in Experiment 9. Tracking error of the MoMSE+SMC method [3] was 3.95 mm and $9.99^\circ$ while the MoMSE+OADE approach was 3.50 mm and $9.58^\circ$ . Our method shows best performance with 2.69 mm and $8.09^\circ$ . . . . .	105
5.13	Comparison of tracking (position and orientation) smoothnesses of six methods during 21 experiments. . . . .	106
5.14	Visual quality of six methods during 21 experiments and computational time of five methods. . . . .	107

- 5.15 Plotted tracking results (Experiments 5 & 21) as camera motion paths on pre-built 3-D bronchial tree model using the OADE method (*blue dots*) and the method of Luo et al. [3] (*green dots*). Dots demonstrate that our method overlaps more ground truth dots or follow longer camera movement paths than Luo et al. [3] (*cyan dots* show ground truth). The errors of Luo et al. [3] were about 3.92 mm (Experiment 5) and 4.23 mm (Experiment 21). However, our method provided the tracking errors about 2.78 mm (Experiment 5) and 3.32 mm (Experiment 21). Dots in the *circle* marks further prove that our method outperforms Luo et al. [3], since many *green* dots were located outside the bronchial tree. . . . . 108
- 5.16 Examples of visual comparison of tracking results of Experiment 18. First column shows frame numbers selected uniformly every 200 frames, and second column shows corresponding real images. Other columns (from third to sixth) display virtual images generated from tracking results using methods of [4], [5], [6], [3], [7], and ours, respectively. Our proposed framework shows the best performance. . . . . 109
- 5.17 Continuous bronchoscopic video frames 420~489 (*left* → *right*, *top* → *bottom*) of Experiment 7 for illustrating that our method can tackle respiratory motion and image artifacts. Image artifacts, such as illumination changes (e.g., Frames 420, 421, and 422), collisions with bronchial walls (e.g., Frames 423~427), and jumps (e.g., Frames 428 and 49), occurred frequently. . . . 110
- 5.18 Virtual bronchoscopic images that were estimated from the method of Luo et al. [3] correspond to bronchoscopic video images in Fig. 5.17. . . . . 111
- 5.19 Virtual bronchoscopic images that were estimated from our proposed OADE method correspond to bronchoscopic video images in Fig. 5.17. Our method still tracks bronchoscope motions successfully and greatly outperforms the method of Luo et al. [3] since these virtual images resemble video images much better than in Fig. 5.18. . . . . 112



# List of tables

3.1	Quantitative comparison of fiducial registration error $f_e$ , target tracking error $t_e$ , and distance between transformed points and assigned centerlines on 15 experiments and their $p$ -values during static phantom validation (Eqs. 3.23, 3.24, and 3.13) (unit: mm). . . . .	38
3.2	Quantitative comparison of fiducial registration error $f_e$ , target tracking error $t_e$ , and distance between transformed points and assigned centerlines on 15 experiments and their $p$ -values during dynamic phantom validation (Eqs. 3.23, 3.24, and 3.13) (unit: mm). . . . .	40
4.1	Comparison of guidance accuracy (position, orientation) in terms of the ground truth datasets and the estimated results of using the five different methods of M1 [4], M2 [5], M3 [8], M4 [9], and M5 (our EPSO-based method). 65	
5.1	Comparatively quantified average accuracy, smoothness, and visual quality of tracking results of six methods evaluated in six ground truth datasets. They were calculated in terms of Eqs. 5.24~5.27. Our OADE outperformed other methods. . . . .	101

