

# **Inferring Human Pose and Motion from Images**

**Yifan Lu**

Research School of Engineering  
The College of Computer Sciences and Engineering  
The Australian National University

August 2011



---

# Acknowledgements

---

I am very grateful to be a member of the computer vision group at Australian National University. I would like to thank my supervisors Dr. Lei Wang, Professor Richard Hartley and Dr. Hongdong Li for invaluable direction and assistance in the development of this work. Thanks especially to Leslay Goldburg, Elspeth Davies and Marie Katselas for helping with conference travel and other issues in PhD life. Thank you also to Professor Henry Gardener for the encouragement and help to international students. Thank you to fellow students, Novi Quadrianto, Yuhang Zhang, Peter Carr, Tamir Yedidya, Luping Zhou and Cong Phuoc Huynh for their accompany. Last but not least, thanks to my parents for their continuous support.



---

# Abstract

---

As optical gesture recognition technology advances, touchless human computer interfaces of the future will soon become a reality. One particular technology, markerless motion capture, has gained a large amount of attention, with widespread application in diverse disciplines, including medical science, sports analysis, advanced user interfaces, and virtual arts. However, the complexity of human anatomy makes markerless motion capture a non-trivial problem: I) parameterised pose configuration exhibits high dimensionality, and II) there is considerable ambiguity in surjective inverse mapping from observation to pose configuration spaces with a limited number of camera views. These factors together lead to multimodality in high dimensional space, making markerless motion capture an ill-posed problem. This study challenges these difficulties by introducing a new framework. It begins with automatically modelling specific subject template models and calibrating posture at the initial stage. Subsequent tracking is accomplished by embedding naturally-inspired global optimisation into the sequential Bayesian filtering framework. Tracking is enhanced by several robust evaluation improvements. Sparsity of images is managed by compressive evaluation, further accelerating computational efficiency in high dimensional space.



---

# Contents

---

<b>Acknowledgements</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Problem Formalisation . . . . .	2
1.2 Challenges . . . . .	3
1.2.1 Ambiguity and Multimodality . . . . .	3
1.2.2 High Dimensionality . . . . .	5
1.2.3 Subject Specific Modelling . . . . .	6
1.2.4 Computational Performance . . . . .	6
1.3 Thesis Outline . . . . .	6
<b>2 Literature review</b>	<b>9</b>
2.1 Global Optimisation and Template based Generative Approach . . . . .	9
2.1.1 Global Optimisation . . . . .	10
2.1.1.1 Annealed Particle Filter . . . . .	10
2.1.1.2 Covariance Scaled Sampling . . . . .	11
2.1.1.3 Interacting Simulated Annealing and Hybrid Approach	12
2.1.2 Template Model Generation . . . . .	15
2.1.2.1 Articulated ICP and Subject Specific Model Generation	15
2.1.2.2 Performance Capture . . . . .	17
2.1.3 Object Localisation and Pose Estimation using a Graphic Model .	21
2.2 Learning Based Approach . . . . .	22
2.2.1 Tracking by Regression . . . . .	22
2.2.2 Gaussian Process Dynamical Model . . . . .	24
2.2.3 Dimensionality Reduction and Manifold Learning on Visual Track- ing . . . . .	26
2.3 Graph Based Image Segmentation and Tracking . . . . .	29
2.3.1 Simultaneous Segmentation and Pose Estimation . . . . .	31
2.3.2 Transductive Image Segmentation . . . . .	32
2.3.3 Laplacian Matrix and Tracking . . . . .	36
<b>3 Architecture Overview and Sequential Tracking Pipeline</b>	<b>39</b>
3.1 Architecture of Human Motion Capture . . . . .	39
3.2 Sequential Bayesian Filtering Framework . . . . .	43
3.3 Particle Filter on Visual Tracking . . . . .	46

---

3.3.1	Recursive Bayesian Filtering . . . . .	46
<b>4</b>	<b>Subject Specific Body Shape Modelling and Automatic Initialisation</b>	<b>51</b>
4.1	Related Works . . . . .	52
4.2	Generic Human Body Skeleton . . . . .	55
4.3	Human body shape . . . . .	58
4.3.1	Needle based Body Shape Parameterisation . . . . .	59
4.3.1.1	Contour from Needle Projection . . . . .	66
4.3.2	Data-Driven Body Shape Parameterisation . . . . .	69
4.3.2.1	Dynamic Bone Length and Collision Bounding Box Adjustment . . . . .	71
4.4	Automatic Initialisation . . . . .	74
4.4.1	Using Needle based Body Parameterisation . . . . .	75
4.4.2	Using Data-Driven Shape Parameterisation . . . . .	77
<b>5</b>	<b>Nature Inspired Global Optimisation</b>	<b>81</b>
5.1	Simulated Annealing . . . . .	82
5.1.1	Simulated Annealing Particle Filter . . . . .	85
5.2	Particle Swarm Optimisation . . . . .	88
5.2.1	Algorithm Description . . . . .	91
5.3	Covariance Matrix Adaptation Evolution Strategy . . . . .	92
5.3.1	Evolution Strategy . . . . .	93
5.3.2	Covariance Matrix Adaptation . . . . .	95
5.3.2.1	Selection and Recombination . . . . .	95
5.3.2.2	Adapting the Covariance Matrix . . . . .	96
5.3.2.3	Step-Size Control . . . . .	102
5.4	Covariance Matrix Adaptation Annealing . . . . .	104
5.4.1	Problems in Dynamic Settings . . . . .	104
5.4.2	Covariance Matrix Adaptation Annealing Algorithm . . . . .	105
5.4.2.1	Perturbation Matrix and Particle Velocity Update . . . . .	107
5.4.3	Experiments with Benchmark Optimisation Problems . . . . .	110
5.4.3.1	Ackley Problem . . . . .	110
5.4.3.2	Rastrigin Problem . . . . .	111
5.4.3.3	Griewank Problem . . . . .	114
5.4.3.4	Rosenbrock Problem . . . . .	116
<b>6</b>	<b>Robust Evaluation Model</b>	<b>119</b>
6.1	Incremental Relaxation by Fast March Method . . . . .	120
6.1.1	Fast March Method . . . . .	123
6.1.2	Experiments . . . . .	126
6.2	Colour and Texture Incorporation . . . . .	128
6.2.1	Illumination Invariant Colour Difference . . . . .	133
6.2.2	Experiments . . . . .	136
6.3	Maximisation of Mutual Information . . . . .	140

6.4	Gradual Sampling for Annealed Particle Filter . . . . .	143
6.4.1	Connection between Gradual Sampling and Annealing Variable .	147
6.4.2	Experiments and Discussion . . . . .	149
<b>7</b>	<b>Compressive Evaluation</b>	<b>155</b>
7.1	Compressive Sensing . . . . .	156
7.1.1	Signal Sparse Representation . . . . .	157
7.1.2	L1 Minimisation and Reconstruction . . . . .	158
7.1.3	Incoherence Sampling . . . . .	159
7.1.4	Restricted Isometry Property . . . . .	161
7.1.5	RIP Random Sensing . . . . .	162
7.2	Discrete Wavelet Transform . . . . .	163
7.3	Compressive Annealed Particle Filter . . . . .	166
7.3.1	Restricted Isometry Property and Pairwise Distance Preservation	168
7.3.2	Multilevel Wavelet Likelihood Evaluation on Compressive Measurements . . . . .	170
7.3.2.1	Construct Increasing Wavelet Coefficient Image . . . . .	171
7.4	Experiments . . . . .	172
<b>8</b>	<b>Conclusion</b>	<b>177</b>
<b>A</b>	<b>Appendix</b>	<b>181</b>
A.1	Perspective Projection . . . . .	181
A.2	Importance Resampling . . . . .	183
A.3	Human Body Segments and Joint Angle Ranges . . . . .	185
A.4	Parameterisations of Three DOF Rotations . . . . .	185
A.4.1	Rotation Matrices . . . . .	185
A.4.2	Euler Angles . . . . .	191
A.4.3	Axis Angle . . . . .	192
A.4.4	Optimisation on Axis Angle . . . . .	193
<b>Bibliography</b>		<b>197</b>

