Device-free Human Localization and Activity Recognition for Supporting the Independent Living of the Elderly

Wenjie Ruan

School of Computer Science

The University of Adelaide

This dissertation is submitted for the degree of

Doctor of Philosophy

Supervisors: Prof. Michael Sheng, A/Prof. Nickolas J.G. Falkner

Dr. Lina Yao and Prof. Xue Li

December 2017

To my mother and father,

my wife and my newborn baby,

my brother,

who made all of this possible,

for their endless encouragement and patience.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Wenjie Ruan

December 2017

Acknowledgements

I would like to thank all the people who have ever helped, supported and advised me in one way or another. Without them, this thesis would never be possible.

First of all, my biggest gratitude goes to my PhD supervisor Prof. Michael Sheng. Without his encouragement three years ago, I would not have had a chance to pursue my PhD at the University of Adelaide. Without his continuous and selfless guidance, I would never have accomplished my PhD research within three years. It is my great honor and luck to have Michael being my PhD supervisor. He has always been patient, passionate, encouraging throughout the whole journey of my PhD study. His many insightful suggestions and comments on my research have significantly improved the work in this thesis.

Secondly, I would like to greatly thank A/Prof. Nickolas J.G. Falkner, who has provided me countless valuable suggestions in academic writing. His many insights regarding how to deliver a native and coherent research idea have greatly increased the quality of this thesis. I also would like to sincerely thank Dr. Lina Yao. During the first year of my PhD journey, her enormous passion on research strongly inspired and encouraged me, leading me to the "door" of academic research. Her supervision with decisiveness and assertiveness has greatly sped up the progress of my PhD project. Moreover, I would like to give my enormous and sincere thanks to Prof. Xue Li, who has provided me many valuable suggestions and strong supports in both my PhD research and job hunting.

Thirdly, I would like to sincerely thank A/Prof. Tao Gu, who has provided many valuable suggestions and assistance on stepping in the research area of pervasive computing. I also want to express heartfelt thanks to Dr. Lei Yang, Dr. Longfei Shangguan and Ms. Peipei Xu for their generous help and valuable discussion in the research. I also want to express my thanks to Dr. Yongrui Qin, who has offered me so many assistance on living in Adelaide. Moreover, I would like to thank other members in our lab, including Dr. Xianzhi Wang, Dr. Ali Shemshadi, Abdullah Alfazi, Dr. Yihong Zhang, Dr. Wei Zhang, Xiu Fang, Tran Khoi Nguyen, and Zhigang Lu for the valuable discussions in group readings and their companionship throughout the PhD journey.

Fourthly, I would like to express my sincere appreciation to University of Adelaide and Australia Research Council for funding my PhD study. I also deeply thank Tsinghua National Lab for Information Science and Technology and Dr. Lei Yang for accepting me and supporting my six-month visiting there. Being a visiting PhD student there has significantly broaden my research horizon and contributed to this thesis.

Lastly, I would express my countless thanks to my parents and my little brother, for their love and support. I am also deeply grateful to my wife, who has accompanied me in University of Adelaide during last two years of my PhD journey. She not only looked after my living but also joined my research, providing me enormous assistance in experiments and data analysis. It is my fortune to have her in my life. Finally, this thesis is the best gift to my newborn baby, wishing he has a colorful and meaningful life.

Abstract

Given the continuous growth of the aging population, the cost of health care, and the preference that the elderly want to live independently and safely at their own homes, the demand on developing an innovative living-assistive system to facilitate the independent living for the elderly is becoming increasingly urgent. This novel system is envisioned to be *device-free*, *intelligent*, and *maintenance-free* as well as deployable in a residential environment. The key to realizing such envisioned system is to study low cost sensor technologies that are practical for device-free human indoor localization and activity recognition, particularly under a clustered residential home. By exploring the latest, low-cost and unobtrusive RFID sensor technology, this thesis intends to design a new device-free system for better supporting the independent living of the elderly. Arising from this live-assistive system, this thesis specifically targets the following *six* research problems.

Firstly, to deal with severe missing readings of passive RFID tags, this thesis proposes a novel tensor-based low-rank sensor reading recovery method, in which we formulate RFID sensor data as a high-dimensional tensor that can naturally preserve sensors' spatial and temporal information. Secondly, by purely using passive RFID hardware, we build a novel *data-driven* device-free localization and tracking system. We formulate human localization problem as finding a location with the maximum posterior probability given the observed RSSIs (Received Signal Strength Indicator) from passive RFID tags. For tracking a moving target, we mathematically model the task as searching a location sequence with the most likelihood under a Hidden Markov Model (HMM) framework. Thirdly, to tackle

the challenge that the tracking accuracy decreases in a cluttered residential environment, we propose to leverage the Human-Object Interaction (HOI) events to enhance the performance of the proposed RFID-based system. This idea is motivated by an intuition that HOI events, detected by pervasive sensors, can potentially reveal people's interleaved locations during daily living activities such as watching TV or opening the fridge door.

Furthermore, to recognize the resident's daily activities, we propose a device-free human activity recognition (HAR) system by deploying the passive RFID tags as an array attached on the wall. This HAR system operates by learning how RSSIs are distributed when a resident performs different activities. Moreover, considering that falls are among the leading causes of hospitalization for the elderly, we develop a fine-grained fall detection system that is capable of not only recognizing regular actions and fall events simultaneously, but also sensing the fine-grained fall orientations. Lastly, to remotely control the smart electronic appliances equipped in an intelligent environment, we design a device-free multi-modal hand gesture recognition (HGR) system that can accurately sense the hand's in-air speed, waving direction, moving range and duration around a mobile device. Our system transforms an electronic device into an active sonar system that transmits an inaudible audio signal via the speaker and decodes the echoes of the hand at its microphone.

To test the proposed systems and approaches, we conduct an intensive series of experiments in several real-world scenarios by multiple users. The experiments demonstrate that our RFID-based system can localize a resident with average 95% accuracy and recognize 12 activities with nearly 99% accuracy. The proposed fall detection approach can detect 90.8% falling events. The designed HGR system can recognize six hand gestures with an accuracy up to 96% and provide more fine-grained control commands by incorporating hand motion attributes.

Table of Contents

List of Figures **xix**

List of Tables **xxvii**

1 Introduction **1**

1.1 System Overview 3

1.2 Challenges 4

1.3 Summaries of Key Chapters 5

1.3.1 Recovering Missing Readings for Corrupted Sensor Data via Low-Rank Tensor Completion 5

1.3.2 Device-free Human Localization and Tracking Using Passive RFID Tags 6

1.3.3 Enhanced Device-free RFID-based Indoor Localization and Tracking through Human-Object Interactions 8

1.3.4 Device-free Human Activity Recognition based on Passive RFID Tag-Array 9

1.3.5 Enabling the Fine-grained Device-free Fall Detection 10

1.3.6 Realizing Human-Machine Interactions Using Touch-free Hand Gestures 11

1.4 Summary 12

2 Literature Review **15**

2.1 Missing Sensor Reading Recovery 15

2.1.1 Matrix Completion Techniques 16

2.1.2 Tensor Completion Techniques 17

2.2 Device-free Human Localization and Tracking 18

2.2.1 Wearable Devices based Techniques 18

2.2.2 Device-free Techniques 19

2.3 Human Activity Recognition 22

2.4 Fall Detection 24

2.5 Hand Gesture Recognition 27

2.5.1 Wearable Devices based Gesture Recognition 28

2.5.2 Device-free Gesture Recognition 29

2.6 Summary 31

3 Recovering Missing Sensor Readings via Low-Rank Tensor Completion **33**

3.1 Introduction 34

3.2 Problem Formulation 38

3.3 Robust Low-Rank Spatio-Temporal Tensor Recovery 40

3.4 Experiments 45

3.4.1 Comparison Methods 45

3.4.2 Evaluations on Synthetic Data 46

3.4.3 Evaluations on RFID Sensory Data 49

3.5 Conclusion 51

4 Device-free Human Localization and Tracking Using Passive RFID Tags **53**

4.1 Introduction 54

4.2 Preliminary 59

4.2.1 Backscatter Radio Communication 59
4.2.2 Received Signal Strength Indicator (RSSI) 60
4.2.3 Intuitions Verification 62
4.3 Problem Formulation 65
4.4 Localizing Stationary Subject 66
4.4.1 Gaussian Mixture Model based Localization 67
4.4.2 *k* Nearest Neighbor based Localization 69
4.4.3 Kernel-based Localization 70
4.4.4 Discussion 71
4.5 Tracking a Moving Subject 71
4.5.1 Transition Matrix 74
4.5.2 Emission Matrix 75
4.5.3 Viterbi Searching 77
4.5.4 Latency Reduction 78
4.6 Evaluation 79
4.6.1 Hardware Deployment 79
4.6.2 Evaluation Metrics 80
4.6.3 Micro Experiments 80
4.6.4 Field Experiments 88
4.6.5 Parameters Selection 92
4.7 Conclusion 96

5 Enhancing RFID-based Device-free Indoor Localization and Tracking through Human-Object Interactions 97

5.1 Introduction 98
5.2 Preliminary 103
5.2.1 Received Signal Strength Indicator (RSSI) 103

5.2.2 Human-Object Interactions (HOI) 103
5.3 HOI-Loc Overview 104
5.3.1 Problem Definition 105
5.3.2 Solution 105
5.4 Localization 106
5.4.1 RSSI Probability 108
5.4.2 HOI Probability 111
5.5 Tracking 113
5.5.1 Transition Strategy 115
5.5.2 Viterbi Searching 116
5.5.3 Forward Calibration 117
5.6 Implementation and Evaluation 118
5.6.1 Evaluation Metrics 120
5.6.2 Localization 121
5.6.3 Tracking 124
5.6.4 Beyond the Limits 127
5.7 Conclusion 129

6 Device-free RFID-based Human Activity Recognition 131

6.1 Introduction 132
6.2 Background 136
6.2.1 Application Scenarios 136
6.2.2 Observations and Problem Formulation 137
6.3 The Proposed Approach 139
6.3.1 Tag Deployment 139
6.3.2 Steady Activity Recognition 145
6.3.3 Activity Sequence Recognition 146

6.4 Experiments 149
6.4.1 Experimental Settings 151
6.4.2 Results 152
6.5 Conclusion 161

7 Fine-grained Device-free Fall Detection based on Passive RFID Tag Array 163

7.1 Introduction 164
7.2 Hardware and Intuitions 167
7.3 System Architecture 171
7.3.1 Activity Sensing Phase 171
7.3.2 Profile Construction Phase 171
7.3.3 Fall Detection Phase 171
7.3.4 Falling Direction Sensing Phase 172
7.3.5 Altering and Update Phase 172
7.4 Device-free Fine-grained Fall Detection 172
7.4.1 Fall Detection 174
7.4.2 Falling Direction Sensing 177
7.5 Evaluation 180
7.5.1 Evaluation Metrics 180
7.5.2 Sensing Normal Activities and Falls 180
7.6 Discussion 189
7.6.1 Computation Cost 189
7.6.2 Hardware 189
7.6.3 Detection Methods 190
7.6.4 Limitations 190
7.7 Conclusion 191

8 Realizing Human-Machine Interactions Using Touch-free Hand Gestures 193
8.1 Introduction 194
8.2 Preliminaries 197
8.2.1 Doppler Effect 198
8.2.2 COTS Speakers & Microphones 199
8.3 Empirical Studies and Challenges 200
8.3.1 Weak Echo Signal 200
8.3.2 Audio Signal Drift 202
8.4 System Conceptual Overview 203
8.5 Realizing the *AudioGest* System 206
8.5.1 FFT Normalization 207
8.5.2 Audio Signal Segmentation 209
8.5.3 Doppler Effect Interpretation 211
8.5.4 Transforming Frequency Shift Area into Hand Velocity 213
8.5.5 Gesture Recognition 215
8.6 Evaluation 220
8.6.1 Hardware 220
8.6.2 Testing Participants 221
8.6.3 Collection of Ground Truth 221
8.6.4 Evaluation Metrics 222
8.6.5 Micro-Test Benchmark 223
8.6.6 In-suit Experiments 231
8.6.7 Comparing with the State-of-the-Art 233
8.7 Discussion 239
8.7.1 Separation of the Speaker and Microphone 239
8.7.2 Gesture Trajectory 239

8.7.3 Noise Disturbance to Human 240
8.8 Conclusion 240

9 Conclusion and Future Work 243
9.1 Conclusions 243
9.2 Open Issues for Future Work 246
9.2.1 Sensor Data Recovery 247
9.2.2 Device-free Indoor Localization and Tracking 248
9.2.3 Device-free Human Activity Recognition 249
9.2.4 Device-free Fall Detection 250
9.2.5 Device-free Hand Gesture Recognition 251

References 253

APPENDIX A Convergence Proof 269

APPENDIX B Examples of Denoising and Segmentation in AudioGest 271

APPENDIX C Multi-modal Hand Detection Examples 273

List of Figures

1.1 The overall conceptual framework of the proposed system 4

2.1 Design Space: comparing to related fall detection systems 25

3.1 Matrix formulation *vs* Tensor formulation 35

3.2 Relative errors for different known elements ($\rho_n = 0.1$, $a = 1$) 47

3.3 Relative errors for different known elements ($\rho_n = 0.25$, $a = 1$) 47

3.4 Relative errors for different corruption percentages ($\rho_o = 1$, $a = 1$) 47

3.5 Iteration numbers for different known elements ($\rho_n = 0.15$, $a = 1$) 47

3.6 Iteration numbers for different known elements ($\rho_n = 0.3$, $a = 1$) 47

3.7 Iteration numbers for different corruption percentages ($\rho_o = 1$, $a = 1$) . . . 47

3.8 Left: The phenomena of RSSI readings loss in passive RFID tags; Right: The missing rates of RSSI readings from a practical Human Activity Recognition system built upon a passive RFID tag-array 49

3.9 (a) Experimental testbed of RFID sensor array; (b) Relative errors for different tag-array size with 20% missing values 50

4.1 The general idea of the proposed DfP localization and tracking system . . . 56

4.2 Backscatter communication mechanism 59

4.3 Path loss illustration 60

4.4 RSSI variation with distance 61

4.5 The RSSI readings cluster in differentiable spaces when a person appears in different locations . . . 63
4.6 The system architecture . . . 64
4.7 RSSI distribution pattern and fitted by GMM . . . 68
4.8 Localization results of different methods . . . 72
4.9 Localization accuracy comparision with k changes . . . 76
4.10 HMM based methods . . . 78
4.11 Hardware deployment . . . 79
4.12 Multiple RSS fields and testing paths . . . 81
4.13 Tracking errors on three paths (*CT: Constraint Transition; CLT: Constraint-Less Transition*) . . . 84
4.14 Average tracking errors . . . 86
4.15 Tracking error CDF . . . 86
4.16 House layout and tracking paths . . . 87
4.17 Localization accuracy in Senario 1 . . . 89
4.18 Localization accuracy in Senario 2 . . . 89
4.19 Localization accuracy in Senario 3 . . . 90
4.20 Tracking errors on three paths . . . 90
4.21 Tracking error CDF . . . 91
4.22 Tracking errors with tag numbers . . . 92
4.23 k value and GMM component number . . . 93
4.24 Window size in forward calibration . . . 94
4.25 Stationary data *vs* dynamic data . . . 95

5.1 Intuition of *HOI-Loc* . . . 99
5.2 *HOI-Loc* system overview . . . 103
5.3 RSSIs clustering in different HD spaces for subject in different locations . . . 108

5.4 RSSIs from different locations are bounded by isolated HD polyhedrons . . 109
5.5 Localization accuracy for proposed PPI and traditional *k*NN 111
5.6 Localization result based on RSSI signal (*k*=2) 111
5.7 Localization result of fusing HOI events with RSSI signal (*k*=2) 113
5.8 HMM tracking mechanism by fusing RSSI signal and HOI events 117
5.9 Experiment settings and paths 118
5.10 Sensors and RFID hardare deployment *Testing Area: master bedroom:* $3.6m \times 4.8m$, *bedroom:* $3m \times 3.2m$, *kitchen:* $3.6m \times 4.6m$ 119
5.11 Localization result for Stationary Scenario 120
5.12 Localization result for Dynamic Scenario 121
5.13 Localization result for Mixed Scenario 122
5.14 Compare tracking accuracy of *HOI-Loc* with other state-of-the-art systems . 124
5.15 Trackng error CDF (cumulative distribution function) for different device-free methods 125
5.16 Mean tracking errors using different tag numbers 126
5.17 Tracking errors for mutiple residents 127
5.18 Confusion matrix of detecting four basic postures 128

6.1 Proposed lightweight setup: a person performs different activities between the wall deployed with an RFID array and an RFID antenna. The activities can be recognized by analyzing the corresponding sensing data collected by the RFID reader. 133
6.2 (a) Histogram of RSSI from activity *sit leaning left*; (b) Histogram of RSSI from activity *sit leaning right* 137
6.3 RSSIs from 9-tag array for a fall with different orientations 138
6.4 Illustration of RSSI fluctuations of falling right and falling left: RSSIs of tag 1, tag 2 and tag 3 (top) and RSSIs of tag 7, tag 8 and tag 9. 140

6.5 Illustrative examples of tag correlations 142
6.6 RFID tags/reader/antenna (left); Lab setting (middle) and Bedroom setting (right) 149
6.7 Predefined orientation-sensitive activities 150
6.8 An example of activity changes 152
6.9 Activity classification comparison with Top N tag selection in (a) lab and (b) bedroom environments 153
6.10 Selected tags 155
6.11 Accuracy comparison with tag selection and without tag selection using different training sizes: (a) lab and (b) bedroom 156
6.12 Performance comparison on different window sizes using 30s and 60s strategies without tag selection and with tag selection (a) lab and (b) bedroom . . 158
6.13 Recognition latency: blue dot vertical line indicates the ground-truth time point of activity change, pink dot vertical line indicates the recognition time point detected by our proposed approach. 160

7.1 RSSIs variation patterns when falls occur 165
7.2 Hardware Deployment 167
7.3 RSSIs variation patterns when a subject falls from different status 168
7.4 RSSIs variation patterns when a subject falls to different directions from standing 169
7.5 System Architecture 170
7.6 Intuition of angle-based outlier detection 173
7.7 Intuition of pABOD 176
7.8 Outline of DTW based kNN 179
7.9 Room layout and three representative action paths 181
7.10 Types of normal activities 181

7.11 Different falls in the experiments 182

7.12 Regular activity categories and boundaries 183

7.13 Confusion Matrix and Detection Performance 184

7.14 Detection rate and false detection rate varies with the boundaries size (X-axis only shows the lower boundary, so upper boundary should be $100\% - LowerBoundary$, the boundary range should be $UpperBoundary - LowerBoundary$) 185

7.15 Confusion Matrix of DTW based kNN ($k = 3$) 186

7.16 Accuracy of classifying falling direction varies with parameter k 186

7.17 Detect fall events in action paths 187

8.1 Illustration of Doppler Frequency Shift 199

8.2 Speakers and microphones in COTS mobile devices 200

8.3 The Doppler frequency shifts caused by different hand gestures and waving speeds 201

8.4 The sound signal drifts for different mobile devices at different time slots . 202

8.5 Overview of the system for hand gesture detection 205

8.6 Left Figure: raw audio spectrogram; Right Figure: audio spectrogram after FFT normalization 207

8.7 All spectrums of audio signal frames: each line represents a spectrum of each frame 208

8.8 Left Figure: the spectrogram after continuous frame subtraction; Right Figure: the spectrogram after the square calculation 209

8.9 Left Figure: the spectrogram after Gaussian Smooth Filter; Right Figure: the segmented area where Doppler Frequency shift happens 210

8.10 The hand moving path with its generated audio spectrogram. Left Figure: hand moving from Right to Left; Right Figure: hand moving along Clockwise Circle 212

8.11 The illustration of transforming frequency shifts into hand velocity, in-air duration and waving range 214
8.12 Six hand waving scenarios: (a) Up-to-down hand waving; (b) Down-to-up hand waving; (c) Right-to-left hand waving; (d) Left-to-right hand waving; (e) Anticlockwise hand circling; (f) Clockwise hand circling 218
8.13 The three mobile devices used for testing 220
8.14 The illustration of handsize measurement and participant information . . . 221
8.15 The 3-axis accelerometer in smartwatch 221
8.16 The average gesture classification accuracy for different mobile devices and users 223
8.17 The Confusion Matrix for the gesture classification 224
8.18 The hand in-air duration estimation error for different mobile devices and users 225
8.19 The average speed-ratio estimation error of hand moving for mobile devices and users 225
8.20 The average range-ratio estimation error of hand moving for different users 225
8.21 The gesture detection accuracy with parameter *H-size* 226
8.22 The gesture detection accuracy with parameter σ 226
8.23 The gesture detection accuracy with gesture signal threshold 226
8.24 The device orientation angle with its detection accuracy 229
8.25 The device-hand distance with its detection accuracy 229
8.26 The average detection accuracy for different scenarios 229
8.27 The detection accracy with and without denoising 232
8.28 The average gesture classification accuracy for in-suit test 232
8.29 The average estimation error of hand in-air duration for in-suit test 232
8.30 The average speed-ratio estimation error of hand movement for in-suit test . 232

8.31 SoundWave detects the frequency shift based on a percentage-threshold method. *For one peak case, it detects the bandwidth of the amplitude drops below 10% of the tone peak. For a large frequency shift casing two peaks, it performs a second scan (if the second peak $\geq$ 30%) and repeats the first scan to find the bandwidth drops from the second peak.* 233

8.32 Experimental Case 1: a slow-speed clockwise hand circling 236

8.33 Experimental Case 2: a fast-speed clockwise hand circling 237

B.1 Denoised spectrograms of different hand gestures with various speeds and their segmentation results: waving hand (a) from Right to Left; (b) from Up to Down; (c) Anticlockwise Circle; (d) Clockwise Circle; (e) Clockwise Circle with fast speed; (f) Clockwise Circle with slow speed 272

C.1 The echo spectrograms and the detected hand motion attributes: (a) Up-Down; (b) Down-Up. *We can distinguish different hand gestures via the waving directions, being similar to current hand-gesture recognition systems.* 274

C.2 The echo spectrograms and the detected hand motion attributes: (a) Right-Left; (b) Anticlockwise Circle. *We can distinguish different hand gestures via the waving directions, being similar to current hand-gesture recognition systems.* 275

C.3 The echo spectrograms and the detected hand motion attributes for a same hand waving: (a) Fast-Speed Clockwise Circling; (b) Slow-Speed Clockwise Circling. *We can distinguish hand gestures (a) and (b) by the speed-ratios even though their waving trajectories are same.* 276

C.4 The echo spectrograms and the detected hand motion attributes for a same hand waving: (a) Small-Range Clockwise Circling; (b) Large-Range Clockwise Circling. *We can recognize hand gesture (a) and (b) by their range-ratios even though their waving trajectories are same, which enables our multi-modal hand motion detection and to advance current related systems.* 277

List of Tables

2.1 Comparison of typical device-free localization systems 20

4.1 Localization accuracies of different methods by using different ratios of training data 82

5.1 The percentage improvements for the accuracy of our method over the other approaches 123

6.1 Confusion matrix with tag selection in lab 157

6.2 Confusion matrix with tag selection in bedroom 157

8.1 Calculation time and resolution vs. frame sizes 227

8.2 Comparison of typical device-free HGR systems 235