University of Technology, Sydney

Faculty of Engineering and Information Technology
SCHOOL OF COMPUTING AND COMMUNICATIONS

PhD Thesis

Tracking and Fine-Grained Activity Recognition in Depth Videos

Prepared By: Sari Awwad

Principal Supervisor: Prof. Massimo Piccardi

Co-Supervisor: Dr. Richard Xu

November, 2016

Certificate of Authorship and Originality

Title: **Tracking and Fine-Grained Activity Recognition in Depth Videos**

Author: **Sari Awwad**

Date: **November , 2016**

Degree: **PhD**

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of author

Acknowledgements

Foremost, I would like to express my sincere gratitude to my principal supervisor, Professor Massimo Piccardi for the continuous support in my PhD study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better supervisor and mentor for my PhD study.

My sincere thanks also go to my home country “Jordan”, my sponsor “The Hashemite University” and Dr. Ahmad Otoom for giving me the greatest opportunity to complete my PhD with a greatest superviso, and my thanks also goes to Dr. Bashar and Dawacom company.

I would like to thank my mother for her continuous prayer for me, and thanks for my brothers: Osama, Samer, Saed, Sameh, and my lovely sister Joman for their continuous supports.

I thank my colleagues: Dr. Shaukat Abidi, Mrs. Fairouz Hussein, Dr. Ava Bargi and Dr. Ehsan Zare Borzeshi, they are a great group, thank you all.

To a wonderful companion, great gratitude to my wife, Lina. Her support, encouragement, quiet patience and unwavering love were undeniably the bedrock upon which the past fourteen years of my life have been built. Her tolerance of my occasional vulgar moods is a testament in itself of her unyielding devotion and love.

Also, I thank Lina’s parents, Mr. Asad and Mrs. Somayya, for their unending encouragement and support.

To the lovely kids, my daughters Farah and Aseel, and my son Osama who have decorated my life and made it full of happiness and joy.

Finally, this thesis is dedicated to the soul of my father, may Allah forgive him and grant him his highest paradise (Ameen).

Sari Awwad.

November, 2016, Sydney

Abstract

Tracking and activity recognition in video are arguably two of the most active topics within the field of computer vision and pattern recognition. Historically, tracking and activity recognition have been performed over conventional video such as color or grey-level frames, either of which contains significant clues for the identification of targets. While this is often a desirable feature within the context of video surveillance, the use of video for activity recognition or for tracking in privacy-sensitive environments such as hospitals and care facilities is often perceived as intrusive.

For this reason, this PhD research has focused on providing tracking and activity recognition solely from *depth* videos which offer a naturally privacy-preserving visual representation of the scene at hand. Depth videos can nowadays be acquired with inexpensive and highly-available commercial sensors such as Microsoft Kinect and Asus Xtion. The two main contributions of this research have been the design of a specialised tracking algorithm for tracking in depth data, and a fine-grained activity recognition approach for recognising activities in depth video. The proposed tracker is an extension of the popular Struck algorithm, an approach that leverages a structural support vector machine (SVM) for tracking. The main contributions of the proposed tracker include a dedicated depth feature based on local depth patterns, a heuristic for handling view occlusions in depth frames, and a technique for keeping the number of support vectors within a given budget, so as to limit computational costs. Conversely, the proposed fine-grained activity recognition approach leverages multi-scale depth measurements and a Fisher-consistent multi-class SVM. In addition to the novel approaches for tracking and activity recognition, in this thesis we have canvassed and developed a practical computer vision application for the detection of hand hygiene at a hospital. This application was developed in collaboration with clinical researchers from the Intensive Care Unit of Sydney's Royal Prince Alfred Hospital. Experiments presented through the thesis confirm that the proposed approaches are effective, and either outperform the state of the art or significantly

reduce the need for sensor instrumentation. The outcomes of the hand-hygiene detection were also positively received and assessed by the clinical research unit.

Contents

List of Figures

List of Tables

Abbreviations

DP	**D**epth **P**atterns
EM	**E**xpectation **M**aximization
FV	**F**eature **V**ector
HAI	**H**ealthcare **A**ssociated **I**nfections
KPF	**K**alman **P**article **F**ilter
LDP	**L**ocal **D**epth **P**atterns
LDPT	**L**ocal **D**epth **P**atterns for **T**racking
MEI	**M**otion **E**nergy **I**mages
MHT	**M**ultilple **H**ypothesis **T**racking
PCA	**P**rincipal **C**omponent **A**nalysis
PTB	**P**rinceton **T**racking **B**enchmark
RFID	**R**adio **F**requency **I**dentification **D**evice
ROI	**R**egion **O**f **I**nterest
RPAH	**R**oyal **P**rince **A**lfred **H**ospital
SMO	**S**equential **M**inimal **O**ptimization
SVM	**S**upport **V**ector **M**achine
WHO	**W**orld **H**ealth **O**rganisation