Robust Online Visual Tracking

Zhibin Hong

Faculty of Engineering and Information Technology

University of Technology, Sydney

A thesis submitted for the degree of

Doctor of Philosophy

2015

Certificate of Original Authorship

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.

Student: Zhibin Hong
Date: 22/09/2015

I would like to dedicate this thesis to my loving parents
Liduan Lin and *Cheng Hong*

Acknowledgements

I would like to take this good opportunity to appreciate my advisors, several professors, my colleagues, my friends and my family for their significant help during my doctoral study in University of Technology (UTS), Sydney, Australia.

First of all, I would like to express my sincere appreciation and deep gratitude to my advisor supervisor **Prof. Dacheng Tao**. He is the one who led me to the field of academic research. He always gives me plenty of freedom to explore and timely constructive suggestions to help me out of difficulties. I can always benefit and learn a lot from various detailed discussions with him. It is hard to imagine I could have finished this thesis without his high scientific standards, unlimited patience, generous support, constant encouragement and guidance. I also wish to express my sincere appreciation to **Prof. Lianwen Jin** who was my advisor when I was in South China University of Technology for master study. I am so grateful for his generosity, guidance and support. He is the one so important to my career and life track. I would not have come to the field of computer vision and would not have come to UTS and met Prof. Dacheng Tao without him. I am also deeply indebted to **Dr. Xue Mei** who led me to the field of visual tracking. His expertise, guidance and encouragement significantly helped me for the completion of this thesis and kept me away from many detours during the journey of exploration. I sincerely appreciate him for his valuable time for beneficial discussions, constructive suggestions and timely support. I would also like to express my appreciation to **Dr. Chaohui Wang** as a close mentor, collaborator and elder brother, for his kindly support and timely help. In addition, I appreciate Prof. Dacheng Tao, Prof. Lianwen Jin,

Dr. Xue Mei and Dr. Chaohui Wang for their guidance, suggestions and help to my future career in the final stage of PhD study.

I have been fortunate to work in UTS and Centre for Quantum Computation and Intelligent Systems (QCIS). I wish to express my appreciation to Prof. Chengqi Zhang, for his full support for me to attend those top conferences including ECCV, ICCV and CVPR, and for his amazing work of gathering so many brilliant researchers in QCIS. I am grateful to Prof. Xingquan Zhu, Prof. Massimo Piccardi, Prof. Maolin Huang, Dr. Danil Prokhorov, Dr. Lin Chen, Dr. Jun Li for their help during my doctoral study. Moreover, I also want to give special thanks to my excellent collaborators: Jia Wu, Zhe Xu, Zhe Chen, for their brilliant work and timely support, and to my dear colleagues and friends I met in QCIS: Prof. Weifeng Liu, Dr. Wei Bian, Dr. Tianyi Zhou, Dr. Naiyang Guan, Dr. Nannan Wang, Dr. Meng Fang, Dr. Mingsong Mao, Dr. Hongshu Chen, Dr. Shirui Pan, Mingming Gong, Tongliang Liu, Changxing Ding, Maoying Qiao, Dr. Guodong Long, Dr. Jing Jiang, Barbara Munday, Chunyang Liu, Dr Shengzheng Wang, Dr. Yong Luo, Lianyang Ma, Xiaoyan Li, Fei Gao, Jie Gui. Bozhong Liu, Prof. Bo Du, Dianshuang Wu, Dr. Ting Guo, Dr. Lianhua Chi, Kailing Guo, Ruxin Wang, Qiang Li, Weilong Hou, Shaoli Huang, Chang Xu, Chen Gong, Sujuan Hou, Haishuang Wang, Qin Zhang, Yali Du, Zhongwen Xu, Dr. Yi Yang, Hao Xiong, Prof. Shigang Liu, Prof. Xianhua Ben, Prof. Wankou Yang, Xiyu Yu, Jiang Bian, Guoliang Kang, Liu Liu, for the inspiring discussions, kind support and companionship.

I am also grateful to all the other friends: Zhida Guo, Qihao Chen, Xueshen Xian, Shixin Lin, Hepeng Ling, Yingying Li, Ming Chen, Sheng Dai, YingLi Li, Ruifang Chen, Shi Chen, Jiawei Xuan, Yanhua Zhang, Shanlin Ye, Xueying Chen, Tao Wang, Weiqi Tang, Chunyu Wang, Ang Li, Qihan Zhao, Danyan Lei, for their support and company during both joyful and stressful times.

Finally, I would like to express my deeply felt gratitude to my fam-

ily: my parents, my grandparents, my lovely brother, my uncles and aunties, my cousins, for their endless love, encouragement and full support throughout my study and life.

Abstract

Visual tracking plays a key role in many computer vision systems. In this thesis, we study online visual object tracking and try to tackle challenges that present in practical tracking scenarios. Motivated by different challenges, several robust online visual trackers have been developed by taking advantage of advanced techniques from machine learning and computer vision.

In particular, we propose a robust distracter-resistant tracking approach by learning a discriminative metric to handle distracter problem. The proposed metric is elaborately designed for the tracking problem by forming a margin objective function which systematically includes distance margin maximization, reconstruction error constraint, and similarity propagation techniques. The distance metric obtained helps to preserve the most discriminative information to separate the target from distracters while ensuring the stability of the optimal metric.

To handle background clutter problem and achieve better tracking performance, we develop a tracker using an approximate Least Absolute Deviation (LAD)-based multi-task multi-view sparse learning method to enjoy robustness of LAD and take advantage of multiple types of visual features. The proposed method is integrated in a particle filter framework where learning the sparse representation for each view of a single particle is regarded as an individual task. The underlying relationship between tasks across different views and different particles is jointly exploited in a unified robust multi-task formulation based on LAD. In addition, to capture the frequently emerging outlier tasks, we decompose the representation matrix to two collaborative components which enable a more robust and accurate approximation.

In addition, a hierarchical appearance representation model is proposed for non-rigid object tracking, based on a graphical model that exploits shared information across multiple quantization levels. The tracker aims to find the most possible position of the target by jointly classifying the pixels and superpixels and obtaining the best configuration across all levels. The motion of the bounding box is taken into consideration, while Online Random Forests are used to provide pixel- and superpixel-level quantizations and progressively updated on-the-fly.

Finally, inspired by the well-known Atkinson-Shiffrin Memory Model, we propose MUlti-Store Tracker, a dual-component approach consisting of short- and long-term memory stores to process target appearance memories. A powerful and efficient Integrated Correlation Filter is employed in the short-term store for short-term tracking. The integrated long-term component, which is based on keypoint matching-tracking and RANSAC estimation, can interact with the long-term memory and provide additional information for output control.

Contents

List of Figures

List of Tables