# Visual Object Tracking

Zijing Chen

Faculty of Engineering and Information Technology

University of Technology Sydney

A thesis submitted for the degree of

*Doctor of Philosophy*

2018

# Certificate of Original Authorship

I, Zijing Chen, declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This thesis is the result of a research candidature conducted with another University as part of a collaborative Doctoral degree.

This document has not been submitted for qualifications at any other academic institution.

Production Note:

Signature: Signature removed prior to publication.

Date: 29/10/2018

# Acknowledgements

I would like to express my gratitude to all those who helped me finish my doctoral studies.

First and foremost, I want to extend my heartfelt gratitude to my supervisor, Professor Dacheng Tao. His kind supervision, valuable suggestions and warm encouragement helped me to overcome difficulties and to successfully complete this thesis. His foresight and sagacity inspired me to find inspiration in cutting-edge research, which made the process of research a great joy. He deserves to be admired, not only for his great achievements in artificial intelligence, computer vision and machine learning, but for his diligence and courage which inspires others to expend a greater effort on their research.

Second, I would like to thank my co-supervisor, Dr Jun Li, who helped me pursue my passion for research. With incredible talent, he is full of inspiring and effective ideas which helped me survive the tedious and burdensome research tasks. He is extremely patient and is generous with his time, helping me to make gradual progress.

Third, I wish to express my sincere appreciation to Prof. Xinge You. He is a Dual-PhD supervisor at Huazhong University of Science and Technology. Although we are far away from each other, Prof. You gave me many suggestions and ideas via emails and numerous phone calls.

Fourth, I would like to pay tribute to Prof. Sean He, Dr Ling Chen, Dr Wei Liu, and Dr Guoqiang Zhang for their inspiring words, continued support, and elaborate instructions about my presentations and academic reports. They had a significantly positive influence on me during my Ph.D. career.

I would like to thank my excellent collaborators. Dr Ruxin Wang and Dr Qiang Li helped me both in my research work and life in Sydney. Dr Xiubao Jiang, Dr Weihua Ou, Dr Hasan Abdulwahid gave me many tips to help me find solutions, especially in solving complicated mathematical problems related to my research. Mr Huan Fu and Mr Baosheng Yu guided me to the right path on deep learning. I also worked closely with Mr Boxuan Zhong who is always optimistic and cheerful.

Special thanks to my husband Zhe Chen, who helped me to overcome depression, insomnia and pessimism. He was the one who encouraged me to grow as an individual and helped me to replace insecurity with confidence.

Finally yet importantly, I would like to express my special thanks to my parents. I really appreciate their support and the confidence they have placed in me.

# Abstract

Visual object tracking is a critical task in many computer-vision-related applications, such as surveillance and robotics. If the tracking target is provided in the first frame of a video, the tracker will predict the location and the shape of the target in the following frames. Despite the significant research effort that has been dedicated to this area for several years, this field remains challenging due to a number of issues, such as occlusion, shape variation and drifting, all of which adversely affect the performance of a tracking algorithm.

This research focuses on incorporating the spatial and temporal context to tackle the challenging issues related to developing robust trackers. The spatial context is what surrounds a given object and the temporal context is what has been observed in the recent past at the same location. In particular, by considering the relationship between the target and its surroundings, the spatial context information helps the tracker to better distinguish the target from the background, especially when it suffers from scale change, shape variation, occlusion, and background clutter. Meanwhile, the temporal contextual cues are beneficial for building a stable appearance representation for the target, which enables the tracker to be robust against occlusion and drifting.

In this regard, we attempt to develop effective methods that take advantage of the spatial and temporal context to improve the tracking algorithms. Our proposed methods can benefit three kinds of mainstream tracking frameworks, namely the template-based generative tracking framework, the pixel-wise tracking framework and the tracking-by-detection framework. For the template-based generative

tracking framework, a novel template based tracker is proposed that enhances the existing appearance model of the target by introducing mask templates. In particular, mask templates store the temporal context represented by the frame difference in various time scales, and other templates encode the spatial context. Then, using pixel-wise analytic tools which provide richer details, which naturally accommodates tracking tasks, a finer and more accurate tracker is proposed. It makes use of two convolutional neural networks to capture both the spatial and temporal context. Lastly, for a visual tracker with a tracking-by-detection strategy, we propose an effective and efficient module that can improve the quality of the candidate windows sampled to identify the target. By utilizing the context around the object, our proposed module is able to refine the location and dimension of each candidate window, thus helping the tracker better focus on the target object.

# Contents

# List of Figures

# List of Tables