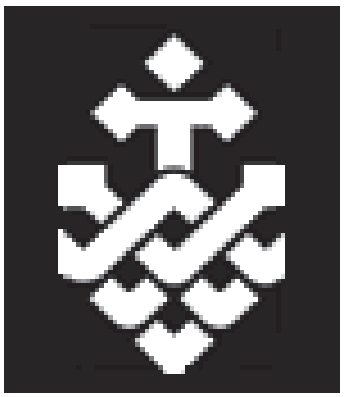
Semi-supervised and Unsupervised Extensions to Maximum-Margin Structured Prediction



Shaukat Abidi

Faculty of Engineering and Information Technology

University of Technology Sydney

A thesis submitted for the degree of

Doctor of Philosophy

2016

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's Name : Shaukat Abidi
Signature of Student :
Date : 18/July/2016

Abstract

Structured prediction is the backbone of various computer vision and machine learning applications. Inspired by the success of maximum-margin classifiers in the recent years; in this thesis, we will present novel semi-supervised and unsupervised extensions to structured prediction via maximum-margin classifiers.

For semi-supervised structured prediction, we have tackled the problem of recognizing actions from single images. Action recognition from a single image is an important task for applications such as image annotation, robotic navigation, video surveillance and several others. We propose approaching action recognition by first partitioning the entire image into "superpixels", and then using their latent classes as attributes of the action. The action class is predicted based on a graphical model composed of measurements from each superpixel and a fully-connected graph of superpixel classes. The model is learned using a latent structural SVM approach, and an efficient, greedy algorithm is proposed to provide inference over the graph. Differently from most existing methods, the proposed approach does not require annotation of the actor (usually provided as a bounding box).

For the unsupervised extension of structured prediction, we considered the case of labeling binary sequences. This case is important in a detection scenario, where one is interested in detecting an action or an event. In particular, we address the unsupervised SVM relaxation recently proposed in (Li et al. 2013) and extend it for structured prediction by merging it with structural SVM. The main contribution of the proposed extension (named Well-SSVM) is a re-organization of the feature map and loss function of structural SVM that permits finding the violating labelings required by the relaxation. Experiments on synthetic and real datasets in a

fully unsupervised setting reveal a competitive performance as opposed to other unsupervised algorithms such as k-means and latent structural SVM.

Finally, we approached the problem of unsupervised structured prediction by M^3 Networks. M^3 Networks are an alternative formulation of maximum-margin structured prediction that can satisfy the complete set of constraints for decomposable feature and loss functions; hence, the entire set of constraints is considered during the search for the optimal margin as opposed to Structural SVM. In the thesis, we present the interpretation of M^3 Networks in Well-SSVM, thus allowing us to use in a semi-supervised and unsupervised scenario.

Acknowledgements

I would like to take this opportunity to acknowledge enormous support from my supervisors, very useful suggestions from my group and a friendly environment of our lab especially during coffees, lunches and dinners.

First of all, I would like to thank my principal supervisor Professor Mary-Anne Williams who provided me the opportunity to come to Australia for PhD. Without her continuous support and supervision, I would have never gotten a chance to explore beautiful practical applications of robotics and computer vision.

I would like to convey big thanks to my co-supervisor Professor Massimo Piccardi, who nurtured my technical knowledge for doctoral degree. His stream of ideas kept me occupied due to which I was able to explore amazing research path. His continuous technical support especially regular meetings, sometimes at coffee shops, has played a central role in the formulation of my technical abilities.

I am thankful to all of my friends, colleagues and group members of Magic Lab and Surveillance Lab: Dr. Benjamin Johnston, Dr. Xun Wang, Dr. Saleha Raza, Dr. Rony Novianto, Wei Wang, Pramod Parajulli, Jonathan Vitale, Mahya Mirzae, Ali Raza, Sylvan Rudduck, Nima Ramezani, Robert Lange, Sari Awwad, Fairouz Hussein, and Ava Bargi. I am thankful to Professor Ivor Tsang, who provided valuable insights and critical reviews for my work. I am grateful to the visiting professors of Magic lab who shared their research experience with me: Professor Pavlos Peppas, Professor Peter Gärdenfors and Associate Professor Sajjad Haider. I would like to thanks Benjamin Johnston and Xun Wang again, with whom I developed demos for robots.

In the end, I would like to thanks my parents without whom I would never be a person I am at the moment and will be in the future.

Contents

List of Figures

List of Tables

Dedicated to my Parents