Markov Random Fields with Unknown Heterogeneous Graphs

Zhenhua Wang
School of Computer Science
The University of Adelaide

This dissertation is submitted for the degree of
Doctor of Philosophy

Dec 6 2014

For my parents and my girlfriend.

Acknowledgements

It would not have been possible to write this doctoral thesis with the help and support of the people around me, to only some of whom it is possible to give particular mention here.

I would like to gratefully and sincerely thank my principle supervisor Professor Anton van den Hengel for his guidance, understanding, patience during my graduate studies at the University of Adelaide. As the founder and director of the Australian Computer Vision Centre (ACVT), he has provided me an excellent research platform and many chances to communicate and cooperate with outstanding researchers and experts. I appreciate his support, encouragement and suggestions on improving my English skills.

I also would like to gratefully and sincerely thank my co-supervisors: Dr. Qinfeng Shi and Dr. Anthony Dick. Dr. Shi taught me Support Vector Machines and Probabilistic Graphical Models at the beginning of my graduate research. He encouraged me to not only grow as a programmer and an experimentalist but also as an instructor and an independent thinker. I am not sure many graduate students are given the opportunity to develop their own individuality and self-sufficiency by being allowed to work with such independence. Dr. Dick give me a lot of help on my study. I especially thank for his recommendation on reading English novels, and his help in joining the Summer Retreat in Victor Harbour arranged by the School of Computer Science.

I would like to give my sincere thanks to Professor Chunhua Shen for his guidance on learning Convex Optimisation. It has always been inspiring and pleasant to discuss with him.

My research has been benefited from discussions with Dr. John Bastian, Dr. Peng Wang, Dr. Cesar Cadena, Dr. Pawan Kumar, Guosheng Lin, Zhen Zhang, for which I thank them.

I thank Mr. Peter McMillan for his advice on English study.

My special thanks also go to Chinese Scholarship Council (CSC) and the University of Adelaide. The former funds me the stipends and the latter waives me the tuition fee for my PhD study.

Also I would like to thank my parents, my brother and my sister for their encouragement and support.

In the end, I would like to gratefully thank my girlfriend. She was always there cheering me up and stood by me through the good times and bad.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma 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 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 Universitys digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed: ______________________

Date: ______________________

Zhenhua Wang
Dec 6 2014

Publication

This thesis is based on the content of the following conference and journal papers:

- Zhenhua Wang, Qinfeng Shi, Chunhua Shen, Anton van den Hengel. Bilinear Programming for Human Activity Recognition with Unknown MRF Graphs. In *Proceedings of International Conference on Computer Vision and Pattern Recognition (CVPR)*, Portland, USA, June 2013.

- Zhenhua Wang, Qinfeng Shi, Chunhua Shen, Pawan Kumar, Anton van den Hengel. Linear and Quadratic Relaxations for Inferring MRF Labels and Graphs Simultaneously. Submitted to some conference. Under review.

- Zhenhua Wang, Zhiyi Zhang, Nan Geng, A Message Passing Algorithm for MRF Inference with Unknown Graphs and Its Applications. *Asian Conference on Computer Vision (ACCV)*, accepted as oral presentation, Singapore, 2014.

In addition, I have co-authored the paper:

- Qinfeng Shi, Mark Reid, Tiberio Caetano, Anton van den Hengel, and Zhenhua Wang. A Hybrid Loss for Multiclass and Structured Prediction. In *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, to appear, 2014.

Abstract

Markov Random Fields have been widely used in computer vision problems, for example image denoising, segmentation and human action recognition. The structure of the graph can be determined using human heuristics or domain knowledge, or can be learned from data when assuming graphs are homogeneous in topology. However, there are many applications for heterogeneous graphs. This research concentrates on estimating heterogeneous graphs and labels simultaneously from the observation. The joint estimation of graphs and labels is formulated into maximising a joint likelihood inference. Unfortunately, these inference problems are generally NP-complete, and we thus develop novel algorithms for effectively and efficiently finding approximate solutions. We also demonstrate how to learn Markov random field parameters from data with our inference techniques.

Our contributions are as follows. First, we show that estimating graphs and performing maximum a posteriori inference can be achieved simultaneously by solving a bilinear programming problem, for which we provide a branch and bound solver. Second, we relax the inference problem into a tighter non-convex quadratic programming problem, and propose a convex-concave procedure algorithm to solve the non-convex quadratic programming. Third, we derive the partial-dual of a mixed integer programming relaxation of the inference problem, which admits a scalable message passing-style algorithm. Lastly, we show how to learn the parameters of Markov random fields with the structured max-margin training and the proposed inference algorithms.

We evaluate the proposed algorithms on both synthetic data and real applications including human action/activity recognition and semantic image segmentation. Our inference algorithms usually outperform the state-of-the-art. Within our proposed algorithms, the quadratic programming algorithm often performs better than the bilinear programming and the message passing algorithms, while the message passing algorithm is the most efficient.

Contents

List of Figures

List of Tables

Nomenclature

Mathematical Notations

a activity variable

$\mathcal{E}$ the set of all possible edges

$\boldsymbol{\theta}$ vector representation of all unary and binary potentials

$\mathcal{O}$ the set of constraints of the non-convex QP formulation

χ the variable of the QP formulation

$\mathbf{Q}$ the matrix of the quadratic term in convex QP formulation

$\deg(i)$ the degree of node i

$\mathbf{diag}(\mathbf{d})$ a diagonal matrix formed by the vector $\mathbf{d}$

$\mathrm{dis}(y_i, y_j)$ the distance function of y_i, y_j used to generate random potentials

$F(\mathbf{x}, \mathbf{y}, G; \mathbf{w})$ discriminant function

E Edge set of G

$\emptyset$ empty set

G Graph

$\mathbb{1}(\cdot)$ indicator vector

$\mathbf{e}_c(y_i)$ the indicator vectors of label

$\mathbf{e}_o(o_i)$ the indicator vectors of body pose

$\mathbf{e}_p(r_{ij})$ the indicator vectors of relative body position

Φ_2 the binary joint feature map

Φ_1 the unary joint feature map

$\otimes$ Kronecker tensor

$\Delta(\mathbf{y}, \hat{\mathbf{y}})$ the cost of predicting $\mathbf{y}$ as $\hat{\mathbf{y}}$

$\delta(y, \hat{y})$ the cost of predicting y as $\hat{y}$

$\mathcal{L}(\Omega)$ the length of Ω

Φ_{lb} the lower bound function

$\mu_{ij}(y_i, y_j)$ the joint distribution of random variables y_i, y_j

$\mathcal{M}$ set contains all constraints of the mixed integer bilinear programming problem

$\mu_i(y_i)$ the distribution of random variable y_i

$Z(\boldsymbol{\theta})$ partition function

$\mathbf{w}$ the parameter vector of potential function

$\mathbf{w}_1$ the parameter vector of unary potential

$\mathbf{w}_2$ the parameter vector of pairwise potential

o_i the body pose of person i

Ω rectangle domain

Ω_{init} the initial rectangular domain

r_{ij} the relative body position of person i to person j

$\preceq$ negative semi-definite matrix notation

h graph sparsity parameter

$\theta_{st}(y_s, y_t)$ pairwise (binary) potential function of y_s, y_t

$\theta_i(y_i)$ unary potential function of y_i

$\hat{\Theta}$ the matrix of the quadratic term in the non-convex QP formulation

Θ block diagonal matrix defined in Section 3.1

$\mathcal{U}$ uniform distribution

Φ_{ub} the upper bound function

V Node set of G

$\mathbf{x}$ input (observation)

$\mathcal{X}$ input space

y a random variable

$\mathbf{y}$ vector contains all y variables

$\mathcal{Y}$ output (labelling) space

z_{ij} the variable that represents the existence of edge (i, j)

$\mathbf{z}$ vector contains all z_{ij} variables

Z-mst MST

Abbreviations

BLP bilinear programming

BP belief propagation

CCCP convex-concave procedure

CRF conditional random field

HAR human action recognition

HoG/HoF histogram of gradients/optical flow

LP linear programming

MAP maximum a posteriori

MCSVM multi-class support vector machine

MRF Markov Random Field

PCA principle component analysis

PDMP partial-dual based message passing

QP quadratic programming

SSVM structured support vector machine

SVM support vector machine