UNIVERSITY OF TECHNOLOGY, SYDNEY

DOCTORAL THESIS

Bayesian Nonparametric Modeling and Its Applications

By

Minqi LI

A thesis submitted in fulfilment of the requirements

for the degree of Doctor of Philosophy

in the

School of Computing and Communications
Faculty of Engineering and Information Technology

October 2016

Declaration of 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.

Signed: ______________________________

Date: ______________________________

Abstract

Bayesian nonparametric methods (or nonparametric Bayesian methods) take the benefit of unlimited parameters and unbounded dimensions to reduce the constraints on the parameter assumption and avoid over-fitting implicitly. They have proven to be extremely useful due to their flexibility and applicability to a wide range of problems. In this thesis, we study the Bayesain nonparametric theory with Lévy process and completely random measures (CRM). Several Bayesian nonparametric techniques are presented for computer vision and pattern recognition problems. In particular, our research and contributions focus on the following problems.

Firstly, we propose a novel example-based face hallucination method, based on a nonparametric Bayesian model with the assumption that all human faces have similar local pixel structures. We use distance dependent Chinese restaurant process (ddCRP) to cluster the low-resolution (LR) face image patches and give a matrix-normal prior for learning the mapping dictionaries from LR to the corresponding high-resolution (HR) patches. The ddCRP is employed to assist in learning the clusters and mapping dictionaries without setting the number of clusters in advance, such that each dictionary can better reflect the details of the image patches. Experimental results show that our method is efficient and can achieve competitive performance for face hallucination problem.

Secondly, we address sparse nonnegative matrix factorization (NMF) problems by using a graph-regularized Beta process (BP) model. BP is a nonparametric method which lets itself naturally model sparse binary matrices with an infinite number of columns. In order to maintain the positivity of the factorized matrices, an exponential prior is proposed. The graph in our model regularizes the similar training samples having similar sparse coefficients. In this way, the structure of the data can be better represented. We demonstrate the effectiveness of our method on different databases.

Thirdly, we consider face recognition problem by a nonparametric Bayesian model combined with Sparse Coding Recognition (SCR) framework. In order to get an appropriate dictionary with sparse coefficients, we use a graph regularized Beta process prior for the dictionary learning. The graph in our model regularizes training samples in a same class to have similar sparse coefficients and share similar dictionary atoms. In this way, the proposed method is more robust to noise and occlusion of the testing images.

The models in this thesis can also find many other applications like super-resolution, image recognition, text analysis, image compressive sensing and so on.

Acknowledgements

Foremost, I wish to appreciate my principle supervisor Professor Xiangjian He for his patience, motivation, constant encouragement and help during my candidature. His guidance helped me in all the time of research. Without his support, I cannot finish my PhD study.

I would also like to express my gratitude to my Co-supervisor Dr. Richard Yida Xu for leading me into the wonderful world of Machine Learning, especially, Bayesian nonparametric models. Without his professional guidance and help, this thesis would not be completed.

I am grateful to Professor Kin-Man Lam in Hong Kong Polytechnic University for his insightful comments and valuable suggestions on my research. I am also grateful to Associate Professor Jian Zhang for his financial support and valuable advice on my PhD study in the first year. I wish to thank my colleagues and the staff in the Faculty of Engineering and Information Technology (FEIT): Qiang Wu, Wenjing Jia, Min Xu, Ruo Du, Sheng Wang, Chao Zen, Muhammad Abul Hasan, Mohammed Ambu Saidi, Khaled W. Aldebei, Ying Wan, Xuhui Fan, Ava Bargi, Shaukat Abidi, etc., for their invaluable help and support. I have the honour of studying and working with them in the past four years which is valuably stamped in my life.

Last but not least, I would like to thank my family for their unconditional support throughout my life.

To my parents for your love and support.

Contents

List of Figures

List of Tables

Abbreviations

GaP	**G**amma **P**rocess
DP	**D**irichlet **P**rocess
CRP	**C**hinese **R**estaurant **P**rocess
ddCRP	**d**istance **d**ependent **C**hinese **R**estaurant **P**rocess
BeP	**B**ernoulli **P**rocess
BP	**B**eta **P**rocess
IBP	**I**ndian **B**uffet **P**rocess
SCR	**S**parse **C**oding **R**ecognition

Symbols

$\mathbb{R}$	the set of reals
$\mathbb{N}$	the set of natural numbers
$\mathbb{E}[\cdot]$	expectation of a random variable
δ_θ	measure concentrated at θ
$\mathbf{c}_i^-$	the assignment set excluding c_i
$DP(\alpha, H)$	Dirichlet process with concentration parameter α and base measure H
$GaP(c, H)$	Gamma process with concentration parameter c and base measure H
$BP(\alpha, H)$	Beta process with positive scalar α and base measure H
$\mathcal{IW}(\cdot)$	normal-inverse-Wishart distribution
$\mathcal{MN}(\cdot)$	matrix-normal distribution
$\mathcal{MT}(\cdot)$	matrix-t distribution
$\|\cdot\|_{lp}$	lp norm
$G_{i,n}$	graph weight between data i and n