Facial Image Restoration and Retrieval Through Orthogonality

Dayong Tian

Faculty of Engineering and Information Technology

University of Technology Sydney

A thesis submitted for the degree of

Doctor of Philosophy

2017

To my loving parents
Jubao Tian and *Shuxia Song*
and my wife
Yiwen Wei

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.

Dayong Tian

Acknowledgements

I would like to express my special appreciation and thanks to my supervisor Professor Dacheng Tao. He accepted my application three years ago and gave me the chance to change my life. In last three years, he consistently taught me how to research. From reading literatures to writing my own papers, from implementing others' methods to devising my own methods, I learned a lot under his instructions.
I also wish to give special thanks to Chaoyue Wang, Guoliang Kang and Jiang Bian. Without their helps, my life in Sydney would not be so easy. I am grateful to Baosheng Yu and Zhe Chen. Our work stations are neighboring. Their opinions deeply impressed me. I would like to give my gratitude to Changxing Ding, Kede Ma, Maoying Qiao and Tongliang Liu for their helps on my research.
I would like to give thanks to my friends I met in Conversation@UTS, especially to Sakada Ko. Ko is an exchange student from Kyoto University of Foreign Studies. He shared lots of his experiences on learning English to me.
It is my fortune to met my friends, Archer, Berry, Beryl, Billy, Carrie, Celine, Chain, CK, Danny, David, Emma, Peyton, Rhonda, Vivi, Vivian, William, Wilton and Zoe. I cannot remember how may times Berry and Chain mixed drinks for us, how many times I was drunk with Rhonda and how many times Zoe drove me home. It is them who make my life diverse and give me completely different experiences. I will miss every minute spent with them.
Finally, I would like to express my gratitude to my family, my parents and my wife, for their encouragement and support.

Abstract

Orthogonality has different definitions in geometry, statistics and calculus. This thesis studies how to incorporate orthogonality to facial image restoration and retrieval tasks. A facial image restoration method and three retrieval methods were proposed.

Blur in facial images significantly impedes the efficiency of recognition approaches. However, most existing blind deconvolution methods cannot generate satisfactory results, due to their dependence on strong edges which are sufficient in natural images but not in facial images. A novel method is proposed in this report. Point spread functions (PSF) are represented by the linear combination of a set of pre-defined orthogonal PSFs and similarly, an estimated intrinsic sharp face image (EI) is represented by the linear combination of a set of pre-defined orthogonal face images. In doing so, PSF and EI estimation is simplified to discovering two sets of linear combination coefficients which are simultaneously found by the proposed coupled learning algorithm. To make the method robust to different kinds of blurry face images, several candidate PSFs and EIs are generated for a test image, and then a non-blind deconvolution method is adopted to generate more EIs by those candidate PSFs. Finally, a blind image quality assessment metric is deployed to automatically select the optimal EI.

On the other hand, the orthogonality is incorporated into the proposed Unimodal image retrieval method. Hashing methods have been widely investigated for fast approximate nearest neighbor searching in large datasets. Most existing methods use binary vectors in lower dimensional spaces to represent data points that are usually real vectors of higher dimensionality. The proposed method divides the hashing

process into two steps. Data points are first embedded in a low-dimensional space, and the Global Positioning System (GPS) method is subsequently introduced but modified for binary embedding. Data-independent and data-dependent methods are devised to distribute the satellites at appropriate locations. The proposed methods are based on finding the tradeoff between the information losses in these two steps. Experiments show that the data-dependent method outperforms other methods in different-sized datasets from 100K to 10M. By incorporating the orthogonality of the code matrix, both data-independent and data-dependent methods are particularly impressive in experiments on longer bits.

In social networks, heterogeneous multimedia data correlates to each other, such as videos and their corresponding tags in YouTube and image-text pairs in Facebook. Nearest neighbor retrieval across multiple modalities on large data sets becomes a hot yet challenging problem. Hashing is expected to be an efficient solution, since it represents data as binary codes. As the bit-wise XOR operations can be fast handled, the retrieval time is greatly reduced. Few existing multi-modal hashing methods consider the correlation among hashing bits. The correlation has negative impact on hashing codes. When the hashing code length becomes longer, the retrieval performance improvement becomes slower. The proposed method incorporates a so-called minimum correlation constraint which can be treated as a generalization of orthogonality constraint. Experiments show the superiority of the proposed method becomes greater as the code length increases.

Deep neural network is expected to be an efficient way for multi-modal hashing. We propose a hybrid neural network which consists of a convolutional neural network for facial images and a full-connected neural network for tags or labels. The minimum correlation regularization is imposed on the parameters of output layers. Experiments validates the superiority of the proposed hybrid neural network.

Contents

List of Figures