A STUDY OF OCCLUSION, MULTI-SCALE METHODS AND ILLUMINATION EFFECTS IN PCA-BASED VISION SYSTEMS

A Thesis Submitted to the Committee

in Partial Fulfilment of the Requirements

For the Degree of Master of Science

in the Department of Computer Science

University of Saskatchewan

by

Li Zhao

May 20, 1997

Permission To Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Computer Science
University of Saskatchewan
Saskatoon, Saskatchewan, Canada
S7N 0W0

Abstract

Principal component analysis (PCA) is a traditional pattern recognition technique. Since Turk and Pentland [27] first applied it to human face recognition, PCA-based methods have become an important approach in object recognition in computer vision [28]. Much research has been done in this area thereafter. Nonetheless most has been done in high-level vision. Some inherent problems such as occlusion, segmentation and illumination effects have not been accounted for. Indeed, these problems are general challenges to every machine vision system.

This thesis presents a comprehensive study of these problems. New algorithms are given to deal with these fundamental problems.

To account for occlusion, Mosaic Image Method is presented. In this method, the object is represented as a collection of features and their relative positions (topology). This is a local and global method. Although this method is created to deal with the occlusion problem, to the author's surprise, it turns out that this is a better representation than the optimum one in traditional PCA-based vision systems. Thorough experiments are conducted to test Mosaic Image Method. More than 110,000 test images are used in these experiments. The proposed method can accommodate up to 53% occluded parts with a more than 95% correct recognition rate. To the author's knowledge, this is the best result in the presence of occlusion in PCA-based vision systems.

In traditional PCA-based vision systems, it is assumed that the object can be easily segmented from the environment. This is only true in simple scenes. One method to get around the segmentation problem is to apply multi-scale methods such as the pyramid method or the scale space method. In multi-scale methods, the computation of eigenimages in different scales is computationally intensive. Hence it poses a main problem concerning its usage. In this thesis, an efficient method is presented. This method is only exactly true when the similarity condition holds. In general, it trades accuracy for efficiency. A theoretical error analysis is given for the general situation. Thorough experiments are conducted to test the proposed method. It is found that this algorithm indeed gives good representations in different scales.

Different illumination can drastically change the appearance of an object. To address illumination effects, a theoretical method based on the physics of object reflection and global illumination analysis is presented. In the first part of this method, an analytical closed-form formula is given to compute the covariance matrix. In the second part, a framework is developed to deal with the changes of light source type, number, intensity, and other illumination effects such as interreflection, shadow and specular highlight. An equation called the illumination equation is derived which accounts for all these illumination effects. Two thorough experiments are conducted to test the proposed method. Experimental results confirm that the proposed method is able to capture all these illumination effects satisfyingly.

Combining these three parts, a general PCA-based vision system can be formed. It is robust to both occlusion and lighting changes and is able to process images in multi-scales by an efficient algorithm. The resulting study forms a new perspective on PCA-based vision systems.

Acknowledgements

First, I thank Professor Y.H. Yang for accepting me as his research assistant. I thank him for his valuable advice and discussions with me on academic issues. I thank him for encouraging me to improve my speaking and writing English. In these two years, I feel that my understanding of Computer Vision, Image Processing, Computer Graphics, and Computer Science in general has progressed quite a lot. For this, I thank him and his Computer Vision and Graphics Lab where most of my time in Saskatoon is spent.

I thank Professor J. Greer and Professor C. McCrosky for being my thesis committee members. I thank Professor K. Takaya for being my external examiner. I thank Professor J. Greer for helping me to improve my technical writing. Partially because of him, now, I am confident in technical writing style. I thank Professor C. McCrosky for helping me to improve my technical writing. From him, I really learned something which I already applied to my thesis writing. I thank Professor K. Takaya's comments on my thesis which help me to improve my thesis writing.

I thank Professor A.D. Jepson of University of Toronto for his giving me the improved version of SVD code and for the discussions with me. I thank Dr. M.J. Black of Xerox Palo Alto Research Center for the discussions among him, Professor A.D. Jepson, and me on the paper *EigenTracking* by which I know what the eigenimages and the reconstructed images in that paper actually are.

I thank Dr. Larry Gritz of the George Washington University for his wonderful rendering tool *Blue Moon Rendering Tool*. All the rendered images in this thesis are rendered by this tool.

Table of Contents

Permission To Use **i**

Abstract **ii**

Acknowledgements **iii**

Table of Contents **I**

List of Tables **IV**

List of Figures **V**

1 Introduction **1**

1.1 Introduction to PCA-Based Vision Systems 1

1.2 Organization of This Document 6

2 Principal Component Analysis (PCA) **7**

2.1 Introduction 7

2.2 A Review of Principal Component Analysis 7

3 Mosaic Image Method: A Local and Global Method **12**

3.1 Introduction 12

3.2 Why Occlusion Is a Problem in Traditional PCA-Based Vision Systems . . . 13

3.3 Mosaic Image Method (I): Representation 17

3.4 Experiment to Test the New Representation: A Better Representation than the Optimum Representation 20
3.4.1 Experimental Setup 20
3.4.2 Using the New Representation Purely Globally 21
3.4.3 Using the New Representation Locally and Globally by Mosaic Image Method 24
3.4.4 Using the New Representation Purely Locally: the New Representation Is A Global Representation 24
3.5 Mosaic Image Method (II): Simple Recognition and Reconstruction Algorithms 26
3.6 Experimental Results 30
3.6.1 Generating Test Image with Occluded Parts 30
3.6.2 All Test Images Without Occlusion are Correctly Recognized 31
3.6.3 Experiment 1: Occlusion by Solid Squares 32
3.6.4 Experiment 2: Occlusion by Real Objects 33
3.7 Summary 34

4 Eigenpyramid and Eigenscale Space **37**
4.1 Segmentation vs. Multi-Scale Schemes 37
4.2 Similarity Condition and a Special Case 38
4.2.1 What Does Theorem 1 Mean? 42
4.2.2 Edge Map and Relationships among Three Eigenimages 43
4.3 General Cases and Theoretical Error Analysis 44
4.4 Experimental Results 47
4.5 Summary 49

5 On Illumination Effects in PCA-Based Vision Systems **60**
5.1 Introduction 60
5.2 Comparison with Former Research 61
5.3 A Theoretical Method to Compute Covariance Matrix of All Images in a Light Pyramid Under Single Directional Light Source 64
5.3.1 Preparations 64

5.3.2 A New Method to Compute Eigenimages in PCA-Based Vision Systems 68
5.4 General Illumination Effects 73
5.4.1 Light Sources 73
5.4.2 Interreflection 75
5.4.3 Light Intensity Changes 77
5.4.4 Illumination Equation 79
5.4.5 Implications of Illumination Equation 80
5.5 Combination with Mosaic Image Method 81
5.5.1 Better Approximation 81
5.5.2 Tractable Computation 82
5.5.3 Accommodating Occlusion, Shadow and Specular Highlight 82
5.6 Experiments 84
5.6.1 Experiment (I): Computer Graphics Simulations 84
5.6.2 Experiment (II): Real Objects 88
5.6.3 Using More than Three Eigenimages in the Proposed Method 89
5.7 Summary 89
5.8 Direction for Further Research 91

6 Conclusion **104**
6.1 Summary 104
6.2 Contributions 105

A Glossary **107**

Bibliography **110**

List of Tables

3.1 The results of the experiment with rectangular solid occluded parts 32
3.2 The results of the experiment with the occluder squares in random distribution 33
3.3 The results of the experiments with the rectangular occluded part of a candy box 34
3.4 The results of the experiment with distributed occluders of a candy box . . . 34

5.1 Same light sources with different light intensities 86
5.2 Specification for the test images rendered under illumination of multiple directional light sources 87
5.3 Specification of the lighting conditions for the real image experiment 88

List of Figures

1.1 A PCA-based vision system. 4

3.1 Mosaic Image Method, a local and global method 18

3.2 To recognize an object, at least some key features should be observed 18

3.3 Mask of an object 20

3.4 A box with complex texture 21

3.5 All sample mosaic images in a special location of the mosaic mask 22

3.6 Optimum representation and new representation 23

3.7 Using new representation purely globally 23

3.8 Using new representation locally and globally by Mosaic Image Method . . . 25

3.9 Using new representation purely locally 27

3.10 Test images with solid square occluders 30

3.11 A candy box with complex texture. 31

3.12 Test images with parts of a candy box as occluders 31

3.13 Some experimental results with solid square occluders 33

3.14 Some experimental results with occluded parts of a candy box 35

4.1 Operation of smooth-and-down-sampling 39

4.2 An image and its smooth-and-down-sampling version 39

4.3 A diagram for Eq. (4.7). 40

4.4 The relationships among three groups of eigenimages 44

4.5 Images of a Coke can 45

4.6 Some model images of a Coke can and a 7up can 47

4.7 The first ten eigenimages in three different scales 50
4.8 The first ten eigenimages generated by the SVD method and the SD method 51
4.9 The ratio of pixels in the first eigenimage by the SD method to pixels in that by the SVD method 52
4.10 The ratio diagrams of the first five eigenimages by the SD method and the SVD method 53
4.11 Eigenimages from 10th to 14th by the SVD method 54
4.12 Eigenimages from 10th to 14th by the SVD method and the SD method . . . 54
4.13 The ratio of pixels in 14th eigenimage by the SD method to pixels in that by the SVD method 55
4.14 The difference between the 14th eigenimage generated by the SVD method and that generated by the SD method 55
4.15 Reconstruction by eigenimages generated by the SVD method and the SD method 56
4.16 The ratio of the reconstructed images by the SVD and the SD representations for the first image in Fig. 4.15 57
4.17 The ratio of the reconstructed images by the SVD and the SD representations for the second image in Fig. 4.15 58
4.18 The ratio of the reconstructed images by the SVD and the SD representations for the third image in Fig. 4.15 59

5.1 A Lambertian surface 65
5.2 Approximation 67
5.3 The light pyramid 68
5.4 A light source 74
5.5 Mosaic Image Method, a local and global method. 81
5.6 A polygonal mesh human face 84
5.7 The three synthesized primary images 85
5.8 The mean image and the first three eigenimages 85
5.9 Same light sources with different intensities 93

5.10 The ratio α estimated for every mosaic image in every image 94
5.11 Multiple light sources test. 95
5.12 An image of the experiment setup 96
5.13 Test results for area light sources without accounting for interreflection . . . 97
5.14 Test results for area light sources accounting for interreflection 98
5.15 The three primary images of a celebrity's face 99
5.16 The eigenimages and mean images for the real image test 99
5.17 Test for real images 100
5.18 Comparison of reconstructed images with different sizes of mosaic images . . 101
5.19 The original images and the smoothed reconstructed images by three eigenimages 102
5.20 Using different numbers of eigenimages in the proposed methods 103