IDEA GROUP PUBLISHING
ITB12301
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

This paper appears in the publication, *Biometric Image Discrimination Technologies*
**by** David Zhang, Xiaoyuan Jing, and Jian Yang © 2006, Idea Group Inc.

## Chapter VI

# Solutions of LDA for Small Sample Size Problems


## ABSTRACT

*This chapter shows the solutions of LDA for small sample-size (SSS) problems. We first give an overview on the existing LDA regularization techniques. Then, a unified framework for LDA and a combined LDA algorithm for SSS problem are described. Finally, we provide the experimental results and some conclusions.*


## INTRODUCTION

It is well known that Fisher LDA has been successfully applied in many practical problems in the area of pattern recognition. However, when LDA is used for solving SSS problems, like face identification, the difficulty that we always encounter is that the within-class scatter matrix is singular. This is due to the high-dimensional characteristic of a face image. For example, face images with a resolution of 100×100 will result in a 10,000-dimensional image vector space, within which the size of the within-class scatter matrix is as high as 10,000×10,000. In real-world problems, it is difficult or impractical to obtain enough samples to make the within-class scatter matrix nonsingular. In this singular case, the classical LDA algorithm becomes infeasible.

So, it is necessary to develop a feasible algorithm for LDA for the high-dimensional and SSS case. Generally, there are two popular strategies for LDA in such cases. One strategy is transform-based; that is, before LDA is used for feature extraction, another procedure is first applied to reduce the dimension of the original feature space. The other strategy is algorithm-based; that is, to find an algorithm for LDA that can deal with the singular case directly.

The typical transform-based methods include fisherfaces (Belhumeur, Hespanha, & Kriegman, 1997), EFM (Liu & Wechsler, 2000, 2001), uncorrelated LDA (Jin, Yang, Hu, et al., 2001; Yang, Yang, & Jin, 2001) and so on. These methods can also be subdivided into two categories. In the first category, such as fisherfaces, EFM and the discriminant eigenfeatures technique (Swets & Weng, 1996), PCA is first used for dimensional reduction. Then, LDA is performed in the PCA-transformed space. Since the dimension of the PCA-transformed space is usually much lower than the original feature space, the within-class scatter matrix is certain to be nonsingular. So, the classical LDA algorithm becomes applicable. This type of approach is generally known as PCA plus LDA. In the second category of approaches, like uncorrelated LDA and the method adopted in Yang, Yang, and Jin (2001), another K-L transform technique is used instead of PCA for dimensional reduction. Although the methods mentioned above can avoid the difficulty of singularity successfully, they are approximate because some potential discriminatory information contained in some small principal components is lost in the PCA or K-L transform step. In addition, the theoretical foundation of the above methods is not clear yet, by far. For instance, why select PCA (or K-L transform) for dimensional reduction beforehand? Is any important discriminatory information lost in the PCA process because the criterion of PCA is not identical to that of LDA? These essential problems remain unsolved.

Some typical algorithm-based methods were developed by Hong and Yang (1991), Liu and Yang (1992), Guo, Huang, and Yang (1999), Guo, Shu, and Yang (2001), and Chen, Liao, and Ko (2000). Hong and Yang's method (1991) of avoiding singularity is to perturb the singular within-class scatter matrix into a nonsingular one. The methods of Liu and Wechsler (2001), Guo, Huang, and Yang (1999), and Guo, Shu, and Yang (2001), are based on a mapping technique that transforms the singular problem into a nonsingular one. Their idea is good and the developed theory provides a solid foundation for solving this difficult problem. It can be considered that Chen, Liao, and Ko's method is a special case of Guo, Huang and Yang's approach (Chen, Liao, & Ko, 2000; Guo, Huang, & Yang, 1999). Chen, Liao, and Ko (2000) merely emphasize the discriminatory information within the null space of the within-class scatter matrix and overlook the discriminatory information outside of it. Instead, Guo, Huang, and Yang (1999) and Guo, Shu, and Yang (2001) take those two aspects of discriminatory information into account at the same time. However, the methods mentioned above have a common disadvantage; that is, the algorithms have to run in a high-dimensional, original feature space. So, these methods are all very computationally expensive in the high-dimensional case. Differing from the above LDA methods, a novel direct LDA (DLDA) approach was proposed recently by Yu and Yang (2001). Although DLDA was claimed to be an exact algorithm of LDA in the singular case, in fact, a part of the important discriminatory information is still lost by this method, as demonstrated by the experiments in this chapter.

In this chapter, our focus is on an LDA algorithm for the high-dimensional and SSS case. We attempt to give a theoretically optimal, exact and more efficient LDA algorithm that can overcome the weaknesses of the previous methods. Towards achieving this