

Almost autonomous training of mixtures of principal component analyzers

Mohamed E.M. Musa ^{a,*}, Dick de Ridder ^b, Robert P.W. Duin ^b, Volkan Atalay ^c

^aDepartment of Computer Engineering, Cankaya University, €O_gretmenler Caddesi No. 14, Balgat, Ankara, Turkey

^bFaculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands

^cDepartment of Computer Engineering, Middle East Technical University, TR-06531 Ankara, Turkey

Received 18 March 2003; received in revised form 21 January 2004

Abstract

In recent years, a number of mixtures of local PCA models have been proposed. Most of these models require the user to set the number of submodels (local models) in the mixture and the dimensionality of the submodels (i.e., number of PC's) as well. To make the model free of these parameters, we propose a greedy expectation maximization algorithm to find a suboptimal number of submodels. For a given retained variance ratio, the proposed algorithm estimates for each submodel the dimensionality that retains this given variability ratio. We test the proposed method on two different classification problems: handwritten digit recognition and 2-class ionosphere data classification. The results show that the proposed method has a good performance. _ 2004
Published by Elsevier B.V.

Keywords: PCA mixture model; EM algorithm; Regularization.