Ranking Principal Components in Face Spaces Through AdaBoost.M2 Linear Ensemble

Tiene A. Filisbino

Gilson Antonio Giraldi

Carlos Eduardo Thomaz

Laboratório Nacional de Computação Científica-LNCC/MCTI

Centro Universitário da FEI

November 6th, 2016

Neutral Happy Sad Fear Angry

Radboud Database

JAFFE Database

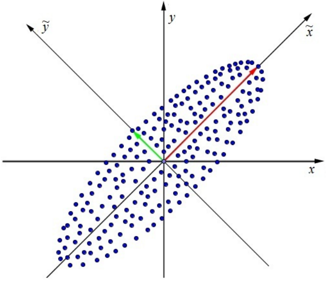
$$\boxed{m - dimensional}$$

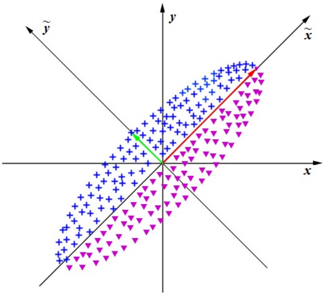
$$\Big\updownarrow [\,T\,]$$

$$\boxed{m' - dimensional}$$

$$m^{'} << m$$

- Linear Techniques: Principal Components Analysis (PCA)





(a) Scatter plot and PCA directions. (b) The same population but distinguishing patterns plus (+) and triangle (▼).

The problem of ranking components obtained by subspace learning techniques is addressed by the discriminant principal components analysis (DPCA) that aims to identify the most discriminant subspaces using weights given by separating hyperplanes [Thomaz and Giraldi, 2010].[1] $\Rightarrow$ Support Vector Machine (SVM)

[1] Thomaz, C. E. and Giraldi, G. A. A new ranking method for principal components analysis and its application to face image analysis.Image Vision Comput,2010.

Procedure for DPCA: for Two classes

1: Determine P_{pca}.

2: Form the matrix $\overline{\Theta}$ given by $\overline{\Theta} = (P_{pca})^T \widetilde{\Theta}$, where $\widetilde{\Theta}$ is the centered data matrix $\widetilde{\Theta} = [\widetilde{\mathbf{x}}_1, \widetilde{\mathbf{x}}_2, \cdots, \widetilde{\mathbf{x}}_M]$;.

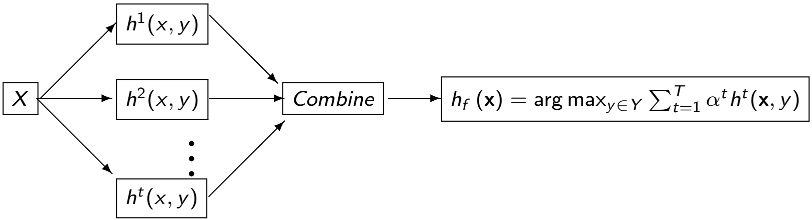
3: Compute ϕ_{svm}.

4: Sort the discriminant weights such that $|\phi_1| \geq |\phi_2| \geq ... \geq |\phi_{m'}|$.

5: Select the principal components according to the obtained $|\phi_i|$ sequence.

Ensemble: AdaBoost.M2

The AdaBoost.M2 algorithm belongs to the class of boosting procedures [Zhou, 2012][2] .

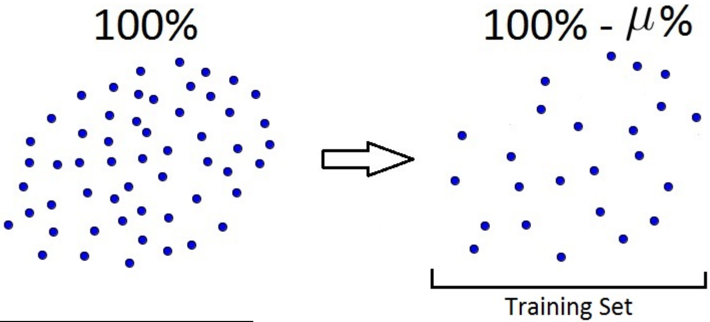


SVM ensemble: $y \in \{1, 2, ..., N\}$

[2] Zhi-Hua Zhou.Ensemble Methods: Foundations and Algorithms, 1st edition, 2012.

The Multi-Class.M2 DPCA pipeline consists of the following steps:

1. Apply PCA technique for dimensionality reduction in order to eliminate redundancy.
2. Compute a linear weak SVM ensemble ϕ_y^t, based on the "one-against-all" SVM multi-class approach, where y is a class.
   - We build a weak SVM , which discards a percentage μ of the samples in the original data set to generate the training set [Garcia and Lozano, 2007].[3]



[3] Elkin Garcia and Fernando Lozano. Boosting Support Vector Machines.In Proceedings of International Conference of Machine Learning and Data Mining, Germany,2007.

3. Combine the discriminant weights computed through the separating weak SVM hyperplanes using AdaBoost.M2 as follows [Filisbino et al., 2015][4].

$$|\Phi_{i,y}| = \left| \sum_{t=1}^{T} \alpha^t \phi_{i,y}^t \right|, y \in Y \quad , i = 1, 2, \ldots, m' \tag{1}$$

4. Compute $v(i) = \max_{y \in Y} \{|\Phi_{i,y}|\}$,
5. Sort discriminant weights: $v(1) \geq v(2) \geq \ldots v(m')$
6. Select the principal components according to the obtained $v(i)$ sequence.

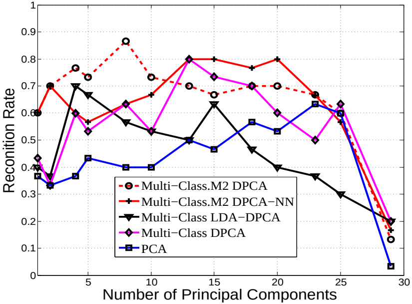
[4] Filisbino, T., Leite, D., Giraldi, G., and Thomaz, C. (2015). Multi-class discriminant analysis based on svm ensembles for ranking principal components. In 36th Ibero-Latin Am. Cong. on Comp. Meth. in Eng. (CILAMCE).

Neutral Happy Sad Fear Angry

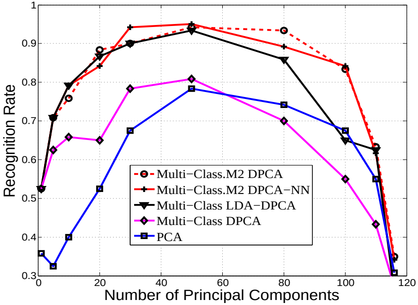
Radboud Database

JAFFE Database

$$d_i(\mathbf{x}_r) = \sum_{j=1}^{k} \frac{1}{\lambda_j}(x_{rj} - \widehat{x}_{ij})^2$$



3 classes(JAFFE)



3 classes(Radboud)

$$I = \widehat{\mathbf{x}} + \delta \cdot \mathbf{q}_i, \tag{2}$$

where $\widehat{\mathbf{x}}$ is the global mean, $\delta \in \{\pm j \cdot \bar{\lambda}^{0.5}, \quad j = 0, \pm 3\}$, and $\bar{\lambda}$ is the average eigenvalue of the total covariance matrix of PCA.

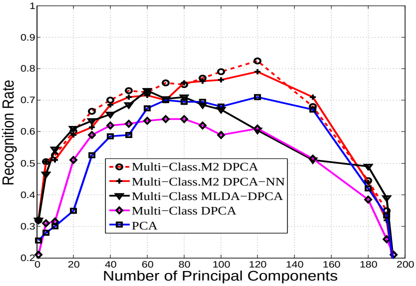
(a) (b) c

(d) (e) (f)

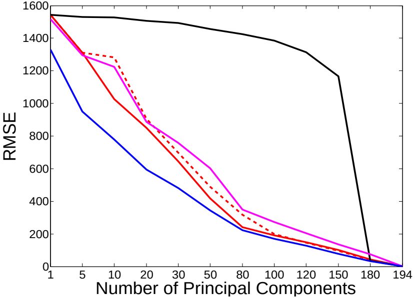
Visualization of the changes described by a principal direction, using the Radboud face database for three-class experiments, selected by:
(a)-(c) Multi-Class.M2 DPCA and (d)-(f) PCA

$$RMSE^{l,i}(k) = \sqrt{\frac{\sum_{j=1}^{M} ||P.I_k^l.P^T\mathbf{x}_j - \mathbf{x}_j||^2}{M}}, \quad (3)$$

where the index l is a methodology, i is a class, I_k^l is a truncated identity matrix that keeps the selected subspace with dimension k, $P = P_{pca}$, and $\|\cdot\|$ is the usual 2-norm.



5 classes(Radboud)



5 classes(Radboud)

- This paper proposes an extension of the DPCA technique for multi-class problems, named Multi-Class.M2 DPCA algorithm.
- In general, Multi-Class.M2 DPCA allow higher recognition rates using less linear features than standard PCA.
- Further work: use Bagging instead of AdaBoost.M2.

Filisbino, T., Leite, D., Giraldi, G., and Thomaz, C. (2015).
Multi-class discriminant analysis based on svm ensembles for ranking principal components.
In *36th Ibero-Latin Am. Cong. on Comp. Meth. in Eng. (CILAMCE)*.

Garcia, E. and Lozano, F. (2007).
Boosting Support Vector Machines.
In *Proceedings of International Conference of Machine Learning and Data Mining (MLDM'2007)*, pages 153–167, Leipzig, Germany. IBal publishing.

Thomaz, C. E. and Giraldi, G. A. (2010).
A new ranking method for principal components analysis and its application to face image analysis.
Image Vision Comput., 28(6):902–913.

Zhou, Z.-H. (2012).
Ensemble Methods: Foundations and Algorithms.
Chapman & Hall/CRC, 1st edition.

Thanks!

Questions?