Paper Review 1

Two-Dimensional PCA : A New Approach to Appearance-Based Face Representation and Recognition
Jian Yang, David Zhang, *Senior Member IEEE*, Alejandro F. Frangi, and Jing-yu Yang

Summary

Representation of digital images and extraction of features from the representation is the first step towards most image processing/computer vision applications. In this research work, the authors describe 2D-Principal Component Analysis (2DPCA) as an alternative image representation technique, and record the performance of this feature representation on a variety of face image datasets for the purpose of facial recognition. One of the major dimensionality reduction and feature extraction techniques used prior to the introduction of this work were Principal Component Analysis (PCA), and it's variants such as kernel-PCA and Independent Component Analysis (ICA). The authors compare the performance of these image representation techniques in terms of the time taken for extracting the features, and the performance of these features for facial recognition on three datasets.

While using PCA for facial recognition, each image is converted into a 1D vector, and from this high-dimensional vector space, the covariance matrix is computed. This makes it a computationally expensive technique. The main idea behind the authors' work is to save computation time by avoiding such a high dimensional vector space. The 2DPCA technique achieves this by constructing an image covariance matrix directly from the original image matrices in the dataset, by estimating a projection matrix using which each image matrix is projected onto a feature vector, under the constraint that the trace of the covariance matrix of the projected feature vectors is maximum. Owing to the maximal variance of the projected feature vectors, this covariance matrix is referred to as *image scatter matrix*, and the constraint as *generalized total scatter criterion.* Infact, the size of the covariance matrix computed by 2DPCA is much smaller than the one computed by PCA and it's variants.

In this research study, to compare the performance of 2DPCA with PCA, the experiments were conducted on three face image databases - the ORL database, the AR database and Yale face database; under a variety of conditions. The ORL database contains images from 40 individuals, each providing 10 images, some of which taken at different times with varying facial expressions. In the initial set of experiments, five image samples per individual were used for training and five for testing. Furthermore, for 2DPCA, the 10 largest eigenvalues were used as the projection axes. In Fig.4 in the paper, the authors compare the reconstructed images of the dataset by adding the first d (eigenvalues) subimages obtained by both 2DPCA and PCA. It can be observed that the image reconstructed using the first 10 eigenvalues obtained using 2DPCA was able to form an image closer to the original image, as compared to the first 40 eigenvalues obtained using PCA. Furthermore, in Table 1 and

Table 2, the authors compare the top recognition accuracy and the top computation time (over a variety of train-test split ratio), for both 2DPCA and PCA. It was observed that 2DPCA required more coefficients to represent the images, but the recognition accuracy was significantly better than that of PCA (some cases with a statistical significance level of 0.05), and a much better computation time. In Table 3, 2DPCA was compared with other techniques such as Fisherfaces, ICA and Kernel-PCA using both train-test split and leave-one-out strategy, and 2DPCA outperformed each of the techniques.

The AR database contains over 4000 color face images of 126 people (70 men, 56 women). Experiments were conducted on this database under conditions under conditions where there was a variation over time, facial expressions and lighting conditions. In this research study, face images of 120 (65 men, 55 women) people (those who participated in two sessions of the data collection, separated by two weeks) were selected by cropping and normalizing each of the image to 50*40 pixels, such that each contained only the face portion of the image. Fig.6 in the paper shows a comparative sample of one of the subjects displaying the same variety of expressions in both the sessions. The performance of 2DPCA with respect to PCA is recorded over the three varieties of conditions, and as can be seen in Table 2 in the paper, it outperformed PCA with a better recognition accuracy in each of the conditions, with a significantly lower feature extraction time.

The Yale database consists of 165 images of 15 individuals, and it was tested under conditions where both facial expressions and illumination varied. In this series of experiments, leave-one-out strategy was adopted and 2DPCA was compared with PCA, ICA and Kernel-PCA. Even in this case, as can be seen in Table 5 in the paper, 2DPCA outperformed each of the mentioned techniques in terms of recognition accuracy. To further evaluate the statistical significance of the results obtained through the series of experiments, the authors employed null hypothesis statistical test on the performance of 2DPCA and PCA. Based on this test, if the p-value turns out to be less than 0.05, then the performance difference between the algorithms is considered to be statistically significant. Based on this, while 2DPCA performed better than PCA in few of the trials for the ORL database; it performed better than PCA in all the trials for the AR database and Yale database.

2DPCA is an unsupervised technique, which is simpler to use for image feature extraction since it deals with image matrices directly, and computes the covariance matrix more accurately as compared to PCA. Secondly, 2DPCA performs better than PCA in all the trials, though it is statistically insignificant on few of the trials in the ORL dataset. Also, 2DPCA is computationally much faster than PCA in all the experiments. These are the major strengths of this technique as compared to PCA. The authors pointed out that one of the major disadvantages of this technique was that more coefficients are needed to represent the image, thus having a more memory requirement for the implementation.

As a follow up of this work, the authors point out that some aspects of 2DPCA are yet to be explored, such as the mean square error between the approximation of the image reconstructed using small number of principal components and the original image, which is minimal in case of PCA. Also, the authors point out that a particular research direction could be to explore techniques which help to reduce the number of coefficients required to represent the image. Zhang et. al in [1], propose a technique as an extension of 2DPCA, in which they consider the row and column directions simultaneously, as opposed to only the row direction in 2DPCA. It performs the same or higher recognition accuracy than 2DPCA, requiring much fewer coefficients set for image representation. A variety of local feature-based techniques have been proposed lately, especially those based on deep learning which have been shown to perform better for the purpose of image representation. Hinton et.al [2] discuss the use of neural networks for transforming the high dimensional data into low-dimensional code. They discuss a pre-training step consisting of learning a stack of Restricted Boltzmann Machines (RBM) for learning feature representations. However, as a further study, a comparitive study analyzing the performance of this feature learning technique with 2DPCA and $2D^2$PCA, using the same classifier for facial recognition can be performed to study the recognition accuracy. However, with deep learning techniques for image representation, the feature extraction time cannot be analyzed and compared.

References

[1] D. Zhang, Z. Zhou. "Two-Directional Two-Dimensional PCA for Efficient Face Representation and Recognition", Neurocomputing, vol. 69, no. 1, pp. 224-231, 2005.

[2] Hinton, Geoffrey E. and Ruslan Salakhutdinov. "Reducing the dimensionality of data with neural networks." Science 313 5786 (2006): 504-7 .