IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 16, Issue 6 Ser. IV (Nov. – Dec. 2020), PP 01-05
www.iosrjournals.org

Assorting Faces by Singular Value Decomposition

M.D. Alsulami
Department of Mathematics, College of Sciences and Arts at Alkamil/ University of Jeddah, Saudi Arabia

Abstract: *Singular value decomposition was used to characterize a matrix containing a set of images. These images were of 100 individuals smiling or not smiling, and they were classified by this characteristic through three different methods: structural similarity, principle component analysis with hyperplane separation, and a neural networks. All three methods yielded results within a 7% error threshold.*

Key Word*: Singular value decomposition; Classifying faces; Data mining and matrices; Neural networks; Principal component Analysis.*

Date of Submission: 01-12-2020 Date of Acceptance: 15-12-2020

I. Introduction

We developed and compared several algorithms for image classification using known techniques, and tested with a data set of smiling and non-smiling faces. Singular Value Decomposition, along with other methods, was used to characterize sample images in order to predict the class for test images. Three different methods were used to determine whether test images were smiling or not:

Structural similarity (SSIM), Principle Component Analysis (PCA) with hyperplane separation, and the use of a neural network. The structural similarity technique compares regions of the images, based on the pixel mean and variance. It then classifies the image based on semblance to sample images of both classes. The PCA technique creates a hyperplane which divides eigenfaces between the classes, then classifies test images by plotting them and determining which side of the plane that they lie. Lastly, the neural network technique is a black box mechanism that finds a nonlinear division between the eigenfaces.

The data set used in this analysis is from the FEI Face data, which was created by the Artificial intelligence Laboratory in Sao Bernardo do Campo, Sao Paulo, Brazil [5]. The images show 100 males and 100 females smiling and not smiling (400 images total), taken between June of 2005 and March of 2006. This data set is standardized and cropped, which allows for cleaner analysis. There are two likely characteristics to analyze from this data set: male vs. female, and smiling vs. not smiling. The latter denominator was chosen to characterize the data set. For each method, 200 images were reserved for the training set (100 people; 100 smiling images, 100 non-smiling images), and the other 200 images constituted the test set.

Methods Used:

II. Singular Value Decomposition (SVD)

Sets of data can be seen as a matrix, with its columns listing the attributes of a datum and the rows listing the various values for each attribute for the set of data. The trends for the data can be determined by decomposing the matrix into a product of matrices. These decomposition are often into three matrices, singular value decomposition (SVD) is an effective form of this technique. The original matrix is often denoted as A which is decomposed into $U\Sigma V^T$ where U is holds the left singular vectors of A which are the eigenvectors of AA^T, Σ is a diagonal matrix containing the singular values corresponding to U, and V holds the right singular vectors of A which are the eigenvectors of A^TA. The singular values in Σ are the square root of correspond eigenvalues for U and V. The singular vectors in U describes the variance in the data are orthogonal to each other, they are ordered from the largest variance to the least. Thus, the Σ matrix is the weighting of each vector in U to recreate the original data [3].

The SVD technique can be used to classify data by comparing the singular values derived from the decomposition between data. Therefore SVD can be used to classify images based upon a particular characteristic. The images are converted into column vectors, so all of the pixel values of an image are in one vector. The vectors are then concatenated together into a matrix which is then decomposed into the singular value form. The singular vectors in U are the "Eigen faces" for the overall image and the Σ contains the singular values for the corresponding vectors thus providing the weighting of the eigenfaces. In theory, a new face (image) could be approximated from the eigenfaces, thus giving a method of facial analysis.

Singular value decomposition was used in all three of our classification methods (SSIM, PCA, and Neural Networks). SVD is an invaluable tool for extracting important information from images.

III. Structural Similarity (SSIM)

Structural Similarity (SSIM) compares two images by comparing pixel data in various windows of the images. This technique was originally designed for image quality assessment [7], where a

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

Compressed or distorted image is compared to an original image, and an index is assigned to indicate how similar the two images are. This technique has also had success in more perception-based applications, such as object movement tracking in videos [2]. An SSIM score is a value between zero and one, where a score of one indicates that the two images are identical. The SSIM of a single window is given by:

where x and y are windows of the two images, μ is the mean pixel value, σ2 is the variance of pixel values, σxy is the covariance, and c1, c2 are constants that stabilize the numerical calculation. The mean of several window SSIM scores determines the overall image SSIM score [7].

We used SSIM to compare two alternate reconstructed faces to an original image. The reconstructed faces were formed by using eigenfaces (from SVD) on a training set. For our purposes, SSIM is more appropriate than other methods of index comparison such as peak signal-to-noise ratio (PSNR) and mean squared error (MSE), which compare images pixel-by-pixel. The code used to calculate SSIM scores was taken from Zhou Wang at the University of Waterloo (ssim_index.m)[6].

IV. Principle Component Analysis (PCA) with Hyperplane Separation

Principal component Analysis is a method of comparing and classifying orthogonal sets of data. The orthogonal data is generally generated using SVD, and the data can be separated with a hyperplane in n-dimensional space, where n is the number of components in each data point. The goal of this method is to find important aspects of complex data, and to separate the data in a simple manner [1].

We used eigenfaces as the orthogonal data, and separated the corresponding eigenface weights of the smiling/non-smiling faces with an n-dimensional hyperplane.

V. Neural Networks

Neural networks were designed to imitate biological nervous systems; they consist of several "neurons," or interconnected nodes that learn by example. Neural network applications include machine learning problems, such as pattern recognition and data classification [4]. Classifying data with neural networks involves constructing a nonlinear regression of a separation curve with a training set, then testing this rule with a test set. As more nodes are used in this algorithm, the regression becomes more detailed, and a neural network scheme that uses too many nodes constitutes overfitting.

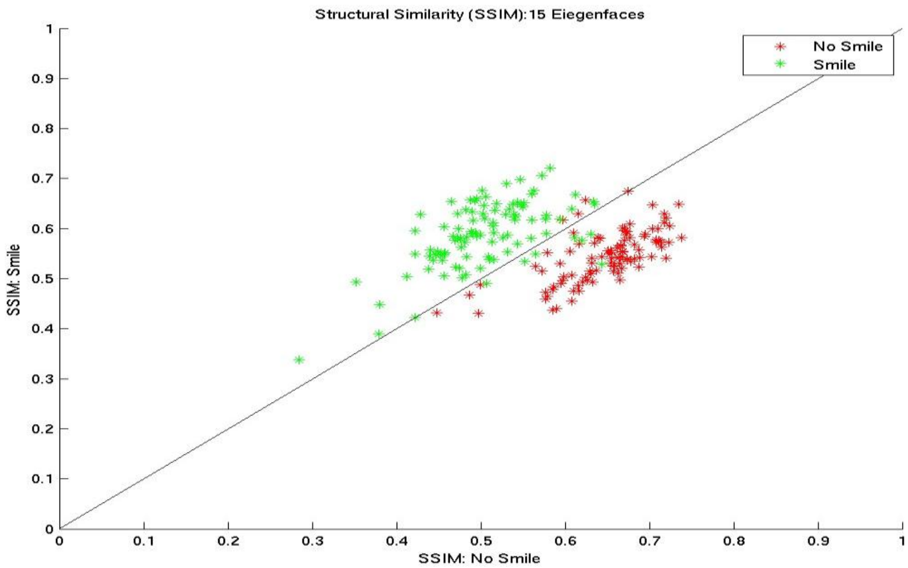
When classifying faces, neural networks worked by the input of the SVD into the pre-included nodal network in MATLAB. The neural network created a nonlinear division between the eigenfaces of the training set, thus creating an algorithm to classify the test set faces.

Implementation and Result:

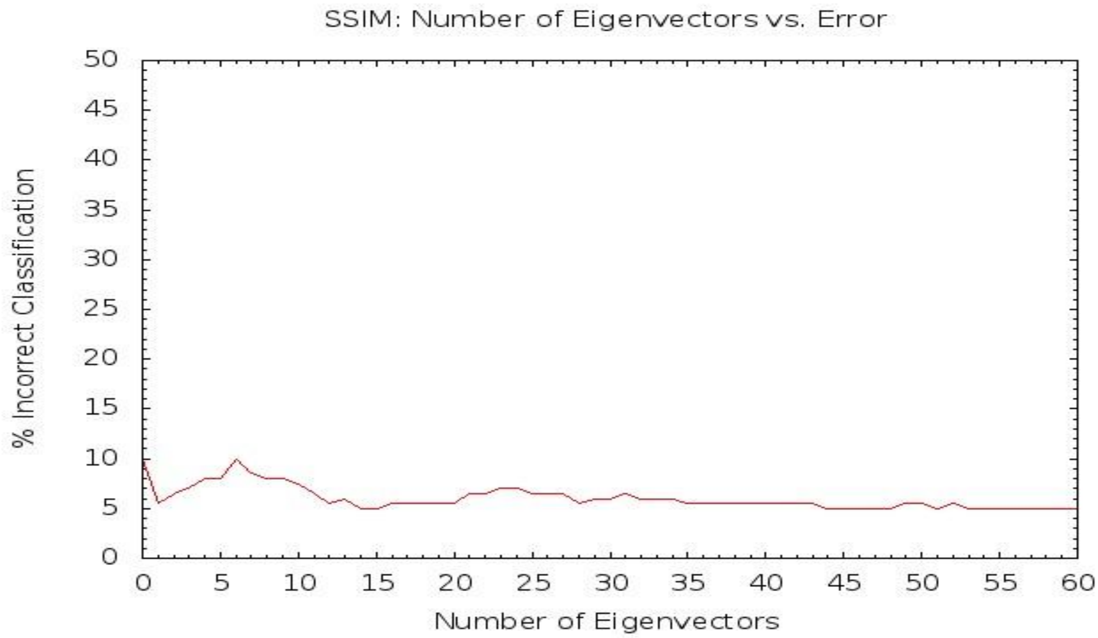
VI. Structural Similarity (SSIM)

Of the 200-image training set, the 100 training images of each group (smiling/non-smiling) were used to form two sets of eigenfaces via SVD. Each image in the test set was reconstructed using the first few smiling eigenfaces, then with the first few non-smiling eigenfaces. Each reconstruction was compared to the original test image, and an SSIM score was assigned to each comparison. The higher SSIM score determined the classification of the image. The reconstruction was done using various numbers of eigenfaces, and the errors were compared.

Here is a plot of the resulting SSIM scores of the test set, where 15 eigenfaces were used to reconstruct each image:



Here is the classification error with a variety of reconstructions:

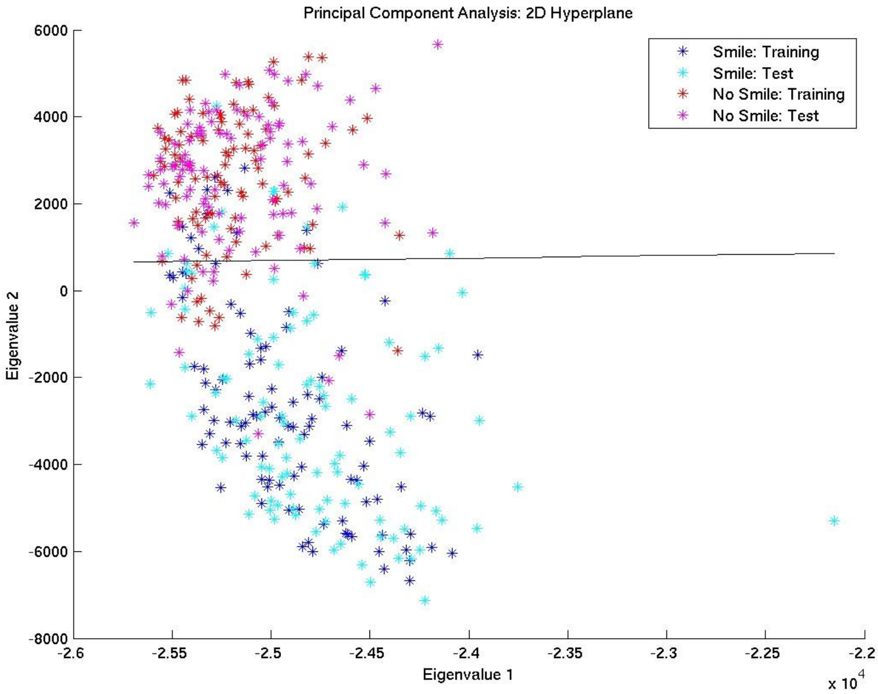


This technique separates the two classes well, and the simple comparison of SSIM scores seems to be an optimal line of separation. The minimum classification error is 5%, and the error starts to level off around 15 eigenfaces. Most of the misclassified images have similar SSIM scores for both the smiling and non-smiling classes. The error converges slowly, and there does not seem to be a definite optimal number of eigenfaces to use.

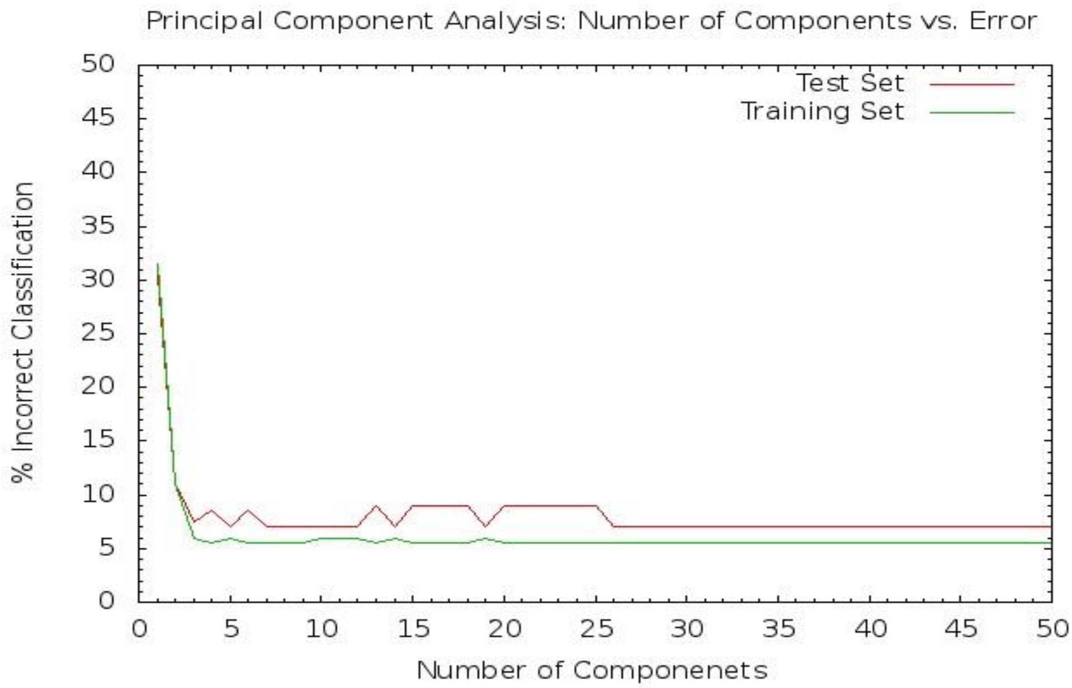
VII. Principle Component Analysis (PCA) with Hyperplane Separation

The training set of 200 pictures was used to form one set of eigenfaces via SVD. The weights of the first few eigenfaces were taken as the principal components, and a hyperplane that best separated the two groups (smiling/non-smiling) was found. This was done by finding a hyperplane orthogonal to the vector connecting each group's mean eigenvalues. The optimal offset of the hyperplane was found via linear search. Each of the 200 images in the test set was reconstructed, and we determined which side of the hyperplane the components of the image belonged to. This determined the classification of the test image.

Here is a plot of the two principal eigenvalues, with the separation of a 2-dimensional hyperplane (2 components):



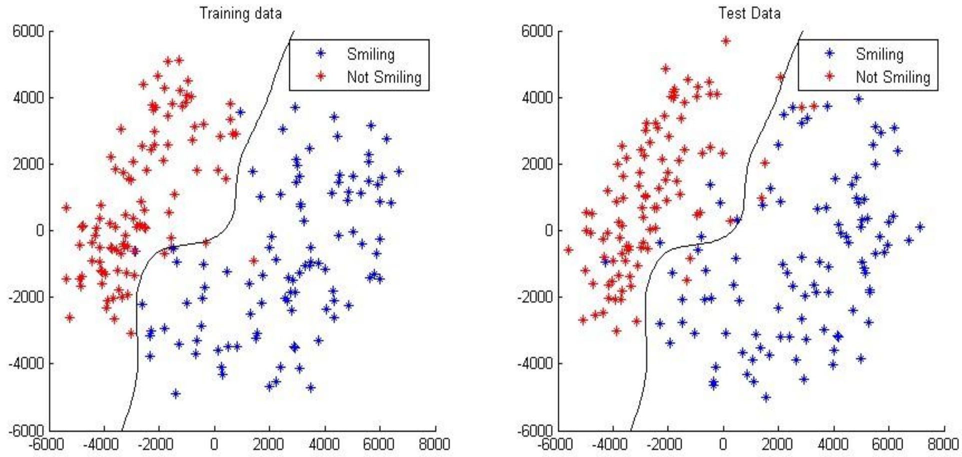
Here is a plot of the error from various n-dimensional hyperplanes:



The 2-dimensional hyperplane seems to separate the two groups well, and there does not seem to be much difference between the training set performance and the test set performance. The training set and test set errors seem to level off at around 3-5 components, the minimum test set error is 7%, and the minimum training set error is 5.5%.

VIII. Neural Networks

Using the neural network tools provided by MATLAB, a neural net was constructed using the weights of our principal components (eigenfaces) as inputs. The training algorithm used was scaled conjugate gradient backpropagation. The training data is comprised from half of our facial images along with their classification (smiling or not smiling). Results of training a network with two inputs (the weights of our two primary eigenfaces) and 20 hidden nodes are shown.



The test error in this case is 7.0%. Increasing the number of inputs or increasing the number of nodes used in the network did not reduce this error.

IX. Conclusion

All of the methods have comparable accuracy in determining the classification of an image.

SSIM has a slightly lower minimum error rate at 5%, compared to 7% of the other two techniques. This may be because the SSIM technique explicitly separates the two classes when calculating eigenfaces, or because no curve-fitting of the training set is required, thus eliminating the danger of overfitting. The SSIM technique, however, requires more eigenfaces to accurately classify the test images than the other two methods, which slows down computation time; SSIM also requires two reconstructions of the test image, while PCA and neural networks require only one.

For two-component separation, neural networks are more accurate than PCA, indicating that a nonlinear regression offers an advantage in this scenario. The PCA and neural network techniques have similar results overall, as the separation of the principal components is close to linear, especially when more eigenfaces are included in the classification decision.

All methods required a low number of eigenfaces to achieve optimal results; SSIM required around 15, PCA required around 5, and neural networks required 2. This is promising, as this allows for efficient classification algorithms in production code.

References

[1]. Abdi, Hervé, Lynne J. Williams. "Principal component analysis." Wiley Interdisciplinary Reviews: Computational Statistics. 2.4 (2010).

[2]. Loza, Artur, Lyudmila Mihaylova, NishanCanagarajah, and David Bull. "Structural Similarity-BasedObject Tracking in Video Sequences." Proc. of the 9th International Conf. on Information Fusion. (2006).

[3]. Rainer Gemulla, Pauli Miettinen. "Data Mining and Matrices 03 – Singular Value Decomposition." Max-Planck-InstitutInformatik, April 25, 2013.

[4]. Stergiou, Christos, DimitriosSiganos. "Neural Networks." .Imperial College London. Web. 9 Dec 2013.

[5]. Thomaz, Carlos Eduardo. "FEI Face Database." Centro Universitário da FEI, 14 Mar 2012.

[6]. Wang, Zhou. "The SSIM Index for Image Quality Assessment." University of Waterloo, 20 Jan2012.

[7]. Wang, Zhou, Alan C. Bovik, Hamid R. Sheikh, Eero P. Simoncelli. "Image Quality Assessment: From Error Visibility to Structural Similarity." IEEE Transactions on Image Processing. 13.4 (2004).

M.D. Alsulami. "Assorting Faces by Singular Value Decomposition." *IOSR Journal of Mathematics (IOSR-JM)*, 16(6), (2020): pp. 01-05.