# SVD and PCA Techniques for Image Classification

Jon Armer
Kent, Wa.
USA

David Kaplan
Puyallup, Wa.
USA

***Abstract*—Singular Value Decomposition and Principal Component Analysis were investigated as image classifiers. Residuals were studied as classifiers in SVD, and the quality of compression and decompression was studied as a classifier studied in PCA.**

***Keywords—SVD; PCA; "eigenfaces"***

## I. Introduction

This project tackled two similar image classification tasks, using two closely related techniques, SVD and PCA.

One of the problems looked at in this paper was how to classify an expression in an image. While it is simple for a human to determine the expression on a person, it is difficult for a computer. A solution to this problem would be useful to A.I and other computer science areas. For an example, determining if a person is smiling could give the A.I. a better context of what is being said.

A second problem was that of algorithmically classifying the gender of a face in an image. The reasons for determining the gender of an individual from an image are many, such as grouping images into male and female before passing to another classifier, determining the gender ratio of customer walking through a store, or determining the gender of a social network profile without access to login info or APIs.

## II. Related Works

The SVD algorithm from [2] is used to try and classify the expression of a person in a picture. The concept is to convert each nxm training picture to a 1x(n*m) vector. Each of these vectors become the column vectors in a new matrix A. Using SVD the U matrix is determined for each expression. Using the U value for each expression the equation $\|(I-UU^T)z\|$, where z is the image you are trying to classify. This equation measures is what is called the residual.

The idea is that the image z should be better approximated in the subspace of the expression that is in z. The expression subspace with the minimal residual is how the algorithm classifies an image.

The 'eigenfaces' technique, using PCA, as described by Turk and Pentland [3] is the primary technique experimented with for the gender classification tasks. It was adopted by one of the authors of this paper, Kaplan [4], in 2005, for undergraduate classwork involving digit recognition. Similar work of Kekre [5], also showed the technique of projecting data into two or more different "eigenspaces", and checking the "quality" of the re-projection into Euclidean space, as basis for a classifier.

## III. Methodology

For SVD, It is possible to improve this algorithm by filtering the image beforehand. To demonstrate this idea the sobel filter was applied to each image. Sobel is an image filter that highlights edges in an image. This filter can help emphasize the parts of the image that are used to classify the image. It can also de-emphasize the parts of the image that are not needed, such as skin tone. The algorithm was applied to grayscale, grayscale polarized, and sober filtered images to compare the accuracy. Fig. 1 shows an example of the images used.

Fig. 1. Left image: Grayscale. Middle image: Grayscale polarize. Right image: Sober filtered.

As Turk and Pentland discuss, PCA can produce "eigenfaces", which are the eigenvectors of the covariance of the set of training face images. When projected onto the span of a set these eigenvectors, an image consisting of n*m pixels values (coordinates), is transformed into a set of weights representing coordinates in the new basis of these eigenvectors.

By creating a separate set of male and female eigenvectors, we create two "face spaces" to project each classification example into. Then, we reverse the projection, projecting the data point back into Euclidean space. The distance between the original and the re-projection is measured, and the class whose "face space" minimizes this re-projection is selected by the classifier.

The data set was taken from the same Essex [6] collection as Kekre's work, although a different subset of images was used. Many folders had missing images upon unpacking parts of the data set, but corrupted files were removed and a data set made up. A total of 38 male and 38 female individuals comprised the photos used. A set of 20 different images for each individual was used, with the exception of one female individual where only 19 images where available.

Each set of 20 images were from the “essex 94” or “easy” set, where the subjects were all facing forwards, and lighting and background was controlled. Of each set of 20 (or 19) images, 2 were held out for testing data. The rest was used for training data. Overall, 341 female and 342 images were used for training, and 38 male and 38 female images were used for testing. The faces were all centered reasonably well, and were 200 rows by 180 columns (portrait view).

Two image preprocessing techniques were tried for the RGB images. The first was to “flatten” the RGB images to grayscale, using the Matlab rgb2gray() function. The second was to separate the red, green, and blue channels and concatenate them as 3 horizontal images, similar to what was done by Kekre in separating the red, green and blue parts of the image. We will refer to this method as the ‘RGB side-by-side’ method for the rest of this paper. The following images show an original 200*180 image, the grayscale conversion, and the red, green, and blue channels stacked side-by-side.


Fig. 2. Top Left: original image: Top Right: grayscale. Bottom image: Red, Green, and Blue Channels.

Once the images were processed in both of these fashions, the average male and female image for each processed set was calculated. To calculate the male eigenfaces, the male average face was subtracted out of the data set (mean adjusted), and to generate the female eigenfaces, the female average face was subtracted out. One problem resolved by Turk and Pentland is that even with a modest 200*180 image size, our images lie in a 36000 dimensional space. However, if our M mean adjusted images are column vectors in a matrix A, then the covariance $C=A^TA$ is 36,000 rows by 36,000 columns, and we do not wish to calculate 36,000 eigenvectors and eigenvalues of this this covariance. Using the “trick” from the eigenfaces paper, we instead calculated the M eigenvectors of $AA^T$. As Turk and Pentland demonstrate “Consider the eigenvectors $\mathbf{v_i}$ of $A^TA$ such that

$$A^TA\mathbf{v_i} = \mathbf{u_i}\mathbf{v_i}$$

Premultiplying both sides by A we have

$$AA^T A\mathbf{v_i} = \mathbf{u_i}A\mathbf{v_i}$$

from which we see that $A\mathbf{v_i}$ are the eigenvectors of C.”

Turk and Pentland state that when M, the number of data examples, is much smaller than the dimension of the space, then there will only be M - 1 meaningful eigenvectors, the rest being essentially zero vectors.

In the experiment we attempt how few eigenvectors we can use to represent the variation in a data set. With our training set, M = 683.

The eigenvectors are sorted according to the magnitude of their eigenvalues. The eigenvector associated with the largest eigenvalue, second largest eigenvalue, …. nth largest value, will be called the 1st top eigenvector, 2nd top eigenvector, …. nth top eigenvector for the rest of this paper. Since an eigenvector has the same number of components and dimension as an image vector, it is possible to display eigenvectors.

Figs. 3 and 4 display some sample “eigenfaces” from our data set.

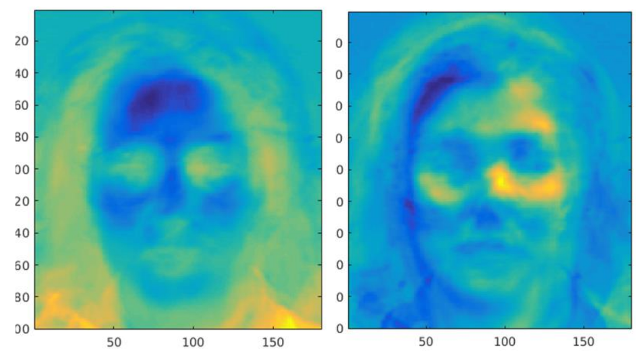

Fig. 3. Left: female eigenface, created from the grayscale faces, corresponding to largest eigenvalue. Right: Female eigenface corresponding to 5th largest eigenvalue.

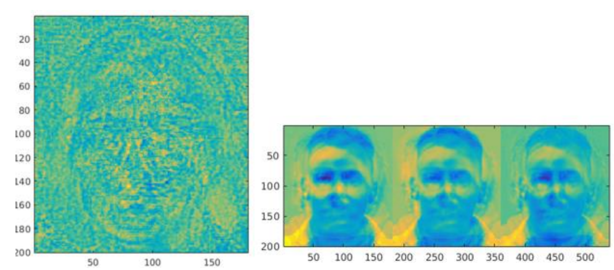

Fig. 4. Left: female eigenface 341, associated with the lowest eigenvalue. Right: Male eigenface 3, created from the side-by-side RBG images.

Once the eigenfaces were created, the test and training data were mean adjusted, then projected onto the span of different subsets of *n* eigenvectors, from 1 to 60 of the top eigenvectors. The result of projection is a vector of weights $\omega_1,\omega_2,...\omega_n$ representing the coordinates of the projection onto the span of the subset of *n* eigenvectors. We reverse the projection with another matrix multiplication and adding the appropriate mean vector back in. The better the subset of eigenvectors chosen represent the vectors, the more accurate the re-projection is. The eigenvector class that minimizes norm(original image – re-projected image) is selected as the correct class. As an example, here is an original grayscale image of a female, the reconstruction using the top 5 female eigenvectors, and the reconstruction using the top 5 male eigenvectors. It is clear the projection into the female face basis better preserves the image when re-projected into Euclidean space, as shown in Fig. 5

Fig. 5. Left: original female image. Middle: image re-projected from basis of 5 female eigenvectors. Right: image re-projected from basis of 5 male eigenvectors.

It is easy to see the one key weakness of having only 76 individuals comprise the subjects for all the training and all the test data. Some individuals seem to have been imprinted very strongly on the eigenvectors. Here is the eigenface in the first five that appears to be strongly associated with this individual, eigenface 4, shown in Fig. 6.

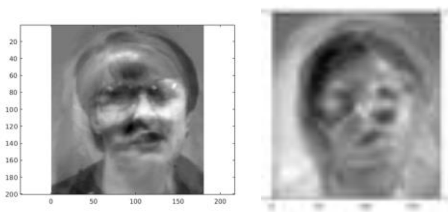

Fig. 6. Left: Female eigenvector #4, grayscale images. Right: Female eigenvector #2, grayscale images.

The vector of weights associated with the projection of this image onto the span of the first five eigenvectors, rounded to 2 decimal places, is:

[-3.22 22.34 5.88 19.80 6.09]

Indeed, the magnitude of the fourth component is very large, but the magnitude of the 2nd female eigenvector is larger. This image appears to be less correlated, at first glance, but the human eyes and mind cannot grasp all the information conveyed by the eigenvectors.

## IV. Results

When the SVD algorithm was applied to just the grayscale images the accuracy was 69.8% overall. The accuracy increased to 87.32% when applied to the sober filtered images. Both fear and neutral expressions were the most difficult for the algorithm to classify, when using grayscale image. The accuracy rating increased by 25.71% by filtering the image first. This is a significant increase that can be achieved quickly (Fig. 7).

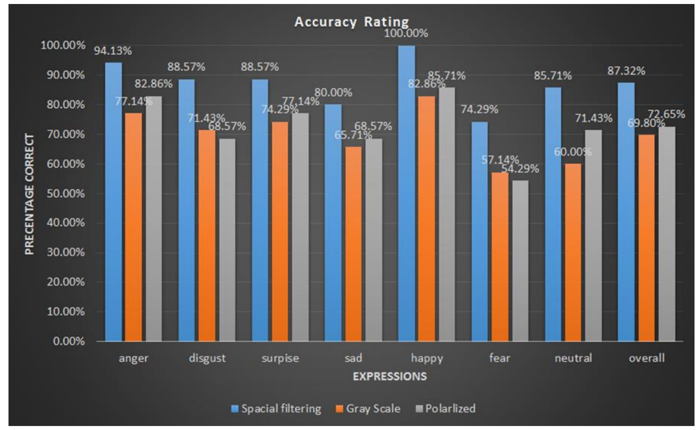


Fig 7. Accuracy rating for classifying each image by expression.

Below are graphs measuring the residues for different images that the algorithm is trying to classify. The residue for the gray scale fear test images is more sporadic when compared to the sobel filter fear test images. There is still the dip in residue in the gray image, but you can also see a few images that have their minimal value in the surprise subspace (Figs. 8,9).

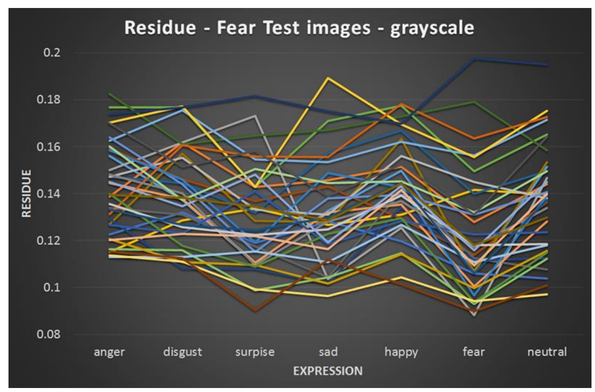


Fig. 8. Residual for grayscale fear test images.

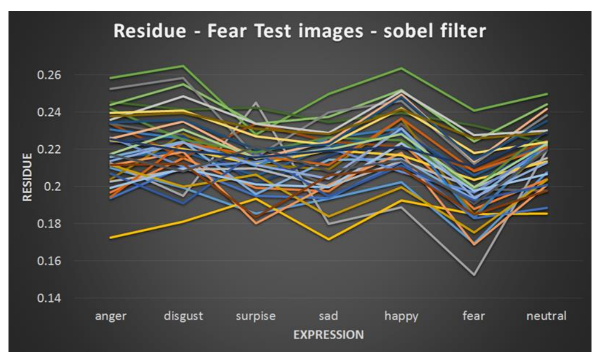


Fig. 9. Residual for sobel filtered fear test images.

The graph of the residual for happy test images shows an example of an ideal graph. All the minimal values are in the happy subspace, and the dip is easy to see (Fig. 10)

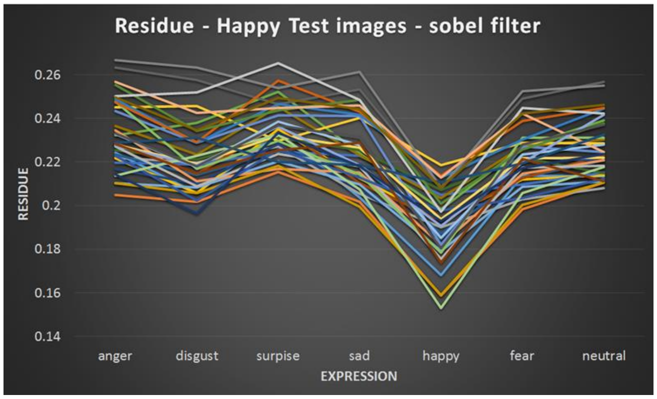


Fig. 10. Residual for sobel filtered happy test images.

The PCA accuracy in gender prediction was tested using between 1 and 60 top eigenvectors as the basis for the eigenspaces. As a baseline value, minimum Euclidean distance from the average male and average female images was tested as a classifier, for both the grayscale and side-by-side RBG images. Accuracy on test data was 76.32% using both types of image processing using this statistic. For the 683 training images, Euclidean distance yielded a 76.28% accuracy using grayscale images and a 78.18% accuracy using the side-by-side RBG images.

Eigenface classification accuracy reached 100% when using around 20 male and 20 female eigenfaces. The side by side RBG appeared to do slightly better on test and training data with a low number of eigenvectors, which may be offset by having to store eigenvectors 3 times as large as for the grayscale processed image (Figs 11-12).

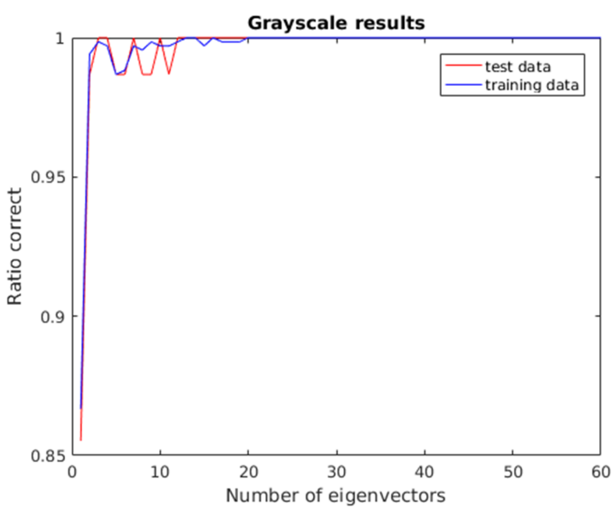


Fig. 11. PCA accuracy with grayscale images, varying number of eigenvectors.

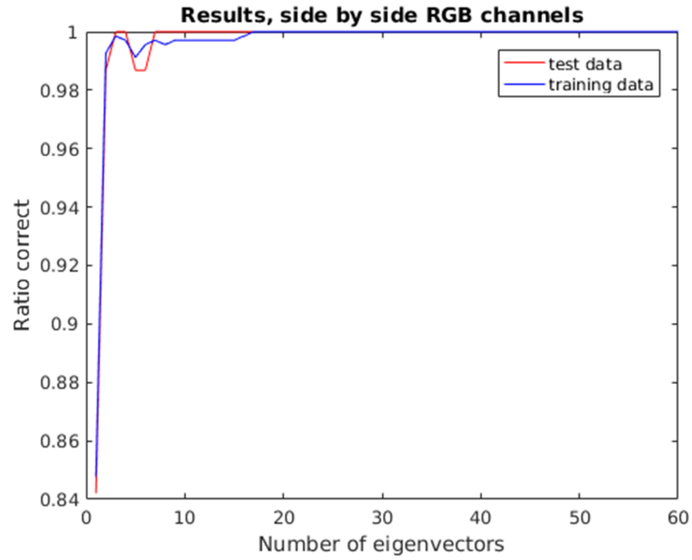


Fig. 12. PCA accuracy with RGB side by side images, varying number of eigenvectors.

## CONCLUSIONS

SVD is a very powerful tool when trying to classify images. By filtering the image beforehand it can become more robust, and give better accuracy. In this paper the sobel filter was applied resulting in an increase in the accuracy. This is a simple filter, and there could be other more complex filters that could help increase that accuracy more.

Due to time constraints we were not able to test the algorithm on half of the face. It would have been interesting to see if there was a noticeable difference between accuracy for the left and right half of the face.

The technique of recognition with eigenfaces was very successful. To be able to store 20 female and 20 male eigenvectors out of a possible 340 or 341 and still have 100 percent gender recognition reveals the power of the technique, and the need to apply it to harder data sets than the one used. Translations, rotation, scaling, and lighting differences all need to be tested. Acquiring larger data sets with more individuals is key to really honing the technique and finding its applicability.

Overall, the PCA and SVD techniques are very close in nature, and mathematically connected. Both are checking a type of residual as a classifier. As Lassiter shows, in SVD the left singular vectors form an orthonormal basis for the range of matrix A. Lassiter also shows that eigenfaces can be created by mean adjusting images, concatenating them as column vectors in a matrix, performing SVD, and the “left-singular vectors *U*, will form a basis of all facial images if given a large enough training set.” This process and the eigenvectors created with PCA are only slightly different in technique, they both form an orthonormal basis for the range of the column vectors of a matrix.

## REFERENCES


[1] Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF, CDROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9

[2] Lassiter, A. and Gugercin, S. (2013). Handwritten Digit Classification and Reconstruction of Marred Images Using Singular Value Decomposition.

[3] Turk, Matthew, and Alex Pentland (1991). Eigenfaces for Recognition. Journal of Cognitive Neuroscience Volume 3, Number 1

[4] Kaplan, David (2005). Class Term Paper Part I using Principal Component Analysis and Matlab code. Developed for UW Seattle Math 381, Fall 2005. Some basic starter code and inspiration adopted from Dr. Tom Duchamp.

[5] Kekre, H.B. Et al. (2010) Face and Gender Recognition Using Principal Component Analysis. International Journal on Computer Science and Engineering Vol. 02, No. 04.

[6] ] Spacek, Libor. (2008). Face Recognition Database, University of Essex, UK.