Categorizing Facial Expressions: A Comparison of computational models

Aruna Shenoy[1], Sue Anthony[2], Ray Frank[1], Neil Davey[1]

1 Department of Computer Science,2 Department of Psychology, University of Hertfordshire, College Lane Campus, Hatfield, AL10 9AB.

Phone: + 44 (0) 1707 284 326

Fax: + 44 (0) 1707 284 156

Email: arushenoy@hotmail.com

Abstract: Recognizing expressions is a key part of human social interaction, and processing of facial expression information is largely automatic for humans, but it is a non-trivial task for a computational system. The purpose of this work is to develop computational models capable of differentiating between a range of human facial expressions. Raw face images are examples of high dimensional data, so here we use two dimensionality reduction techniques: Principal Component Analysis and Curvilinear Component Analysis. We also preprocess the images with a bank of Gabor filters, so that important features in the face images may be identified. Subsequently the faces are classified using a Support Vector Machine. We show that it is possible to differentiate faces with a prototypical expression from the neutral expression. Moreover we can achieve this with data that has been massively reduced in size: in the best case the original images are reduced to just 5 components. We also investigate the Effect size on face images, a concept which has not been reported previously on faces. This enables us to identify those areas of the face that are involved in the production of a facial expression.

Keywords: Expressions, Image Analysis, Classification, Dimensionality Reduction.

1. Introduction

According to Ekman and Friesen [1] there are six easily discernible facial expressions: anger, happiness, fear, surprise, disgust and sadness, apart from neutral. Moreover these are readily and consistently recognized across different cultures [2]. In the work reported here we show how a computational model can identify facial expressions from simple facial images.

Data presentation plays an important role in any type of recognition. High dimensional data is normally reduced to a manageable low dimensional data set. We perform dimensionality reduction using Principal Component Analysis (PCA) and Curvilinear Component Analysis (CCA). PCA is a linear projection technique and it may be more appropriate to use a non linear Curvilinear Component Analysis (CCA) [3]. The Intrinsic Dimension (ID) [4], which is the true dimension of the data, is often much less than the original dimension of the data. To use this efficiently, the actual dimension of the data must be estimated. We use the Correlation Dimension to estimate the Intrinsic Dimension. We compare the classification results of these methods with raw face images

---

1. Department of Computer Science, University of Hertfordshire, Hatfield, AL10 9AB, UK.

2. Department of Psychology, University of Hertfordshire, Hatfield,AL10 9AB, UK.

and of Gabor Pre-processed images [5, 6]. The features of the face (or any object for that matter) may be aligned at any angle. Using a suitable Gabor filter at the required orientation, certain features can be given high importance and other features less importance. Usually, a bank of such filters is used with different parameters and later the resultant image is a *L2 max norm* (at every pixel the maximum of feature vector obtained from each of the filters of the filter bank) superposition of the outputs from the filter bank.

Examining the effect size of each pixel allows us to show the encoding face, which shows those pixels on the face that are involved in the production of an expression.

2. Background

We perform feature extraction with Gabor filters and then use dimensionality reduction techniques such as Principal Component Analysis (PCA) and Curvilinear Component Analysis (CCA) followed by a Support Vector Machine (SVM) [7] based classification technique. We also implement the Effect size and all these methods are described below.

2.1 Gabor filters

A Gabor filter can be applied to images to extract features aligned at particular orientations. Gabor filters possess the optimal localization properties in both spatial and frequency domains, and they have been successfully used in many applications [8]. A Gabor filter is a function obtained by modulating a sinusoidal with a Gaussian function. The useful parameters of a Gabor filter are orientation and frequency. The Gabor filter is thought to mimic the simple cells in the visual cortex. The various 2D receptive field profiles encountered in populations of simple cells in the visual cortex are well described by an optimal family of 2D filters [9]. In our case a Gabor filter bank is implemented on face images with 8 different orientations and 5 different frequencies as it captures all the features necessary that may be significant [10].

Recent studies on modeling of visual cells [11] suggest a tuned band pass filter bank structure. Formally, the Gabor filter is a Gaussian (with variances S_x and S_y along x and y axes respectively) modulated by a complex sinusoid (with centre frequencies U and V along x and y axes respectively) and is described by the following equation:-

$$g(x,y) = \frac{\exp\left[-\frac{1}{2}\left(\left(\frac{x}{S_x}\right)^2 + \left(\frac{y}{S_y}\right)^2\right) + 2\pi j(Ux+Vy)\right]}{2\pi S_x S_y} \qquad (1)$$

The variance terms S_x and S_y dictates the spread of the band pass filter centered at the frequencies U and V in the frequency domain. This filter has real and imaginary parts.

A Gabor filter can be described by the following parameters: The S_x and S_y of the Gaussian define the shape of the base (circle or ellipse), frequency (f) of the sinusoid, orientation (θ) of the applied sinusoid. Figure 1 shows examples of various Gabor filters. Typically in face processing, 5 scales and 8 orientations are used [10]. Figure 2 b) shows the effect of applying a variety of

Gabor filters shown in Figure 1 to the sample image shown in Figure 2 a). Note how the features at a particular orientation are exaggerated.
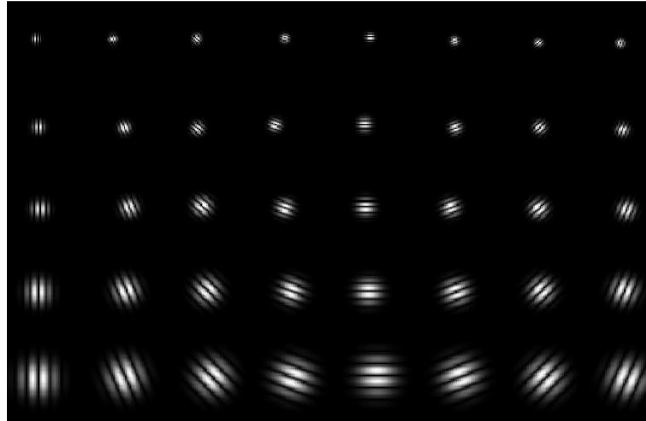

Figure 1: Gabor filters - Real part of the Gabor kernels at five scales and eight orientations

An augmented Gabor feature vector is created of a size far greater than the original data for the image. Every pixel is then represented by a vector of size 40 and demands dimensionality reduction before further processing. So a 63×63 image is transformed to size $63 \times 63 \times 5 \times 8$. Thus, the feature vector consists of all the information extracted from different frequencies, orientations and from all locations, and hence could be useful for expression recognition.
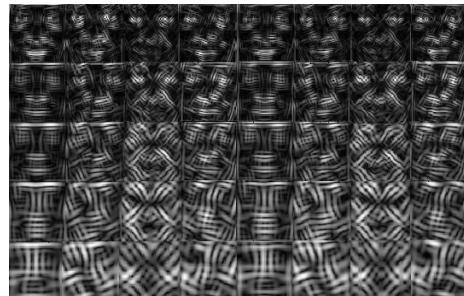
(a)   (b)

Figure 2: a) Original face image 63×63 (3969 dimensions). b) Forty Convolution outputs of Gabor filters.

Once the feature vector is formed, it can be handled in various ways. We reduce the image to its original size by taking the *L2 max norm*, over the 40 convolutions, for each pixel in the feature vector. So that the final value of a pixel is the maximum value found by any of the filters for that pixel.

The *L2 max norm* Superposition principle is used on the outputs of the filter bank and the Figure 3 b) shows the output for the original image of Figure 3 a).
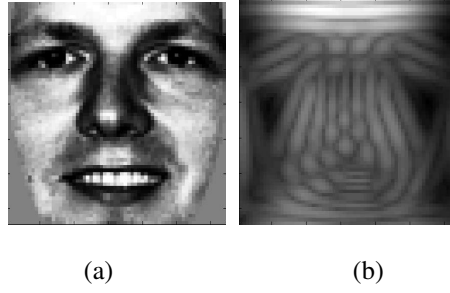
(a)   (b)

Figure 3 a): Original Image used for the Filter bank b) Superposition output (*L2 max norm*)

2.2 Curvilinear Component Analysis

Curvilinear Component Analysis (CCA) is a non-linear projection method that preserves distance relationships in both input and output spaces. CCA is a useful method for redundant and non linear data structure representation and can be used in dimensionality reduction. CCA is useful with highly non-linear data, where PCA or any other linear method fails to give suitable information [3]. The D – dimensional input X should be mapped onto the output d - dimensional space Y. Their d - dimensional output vectors $\{y_i\}$ should reflect the topology of the inputs $\{x_i.\}$ In order to do that, Euclidean distances between the x_i's are considered. Corresponding distances in the output space y_i's is calculated such that the distance relationship between the data points is maintained.
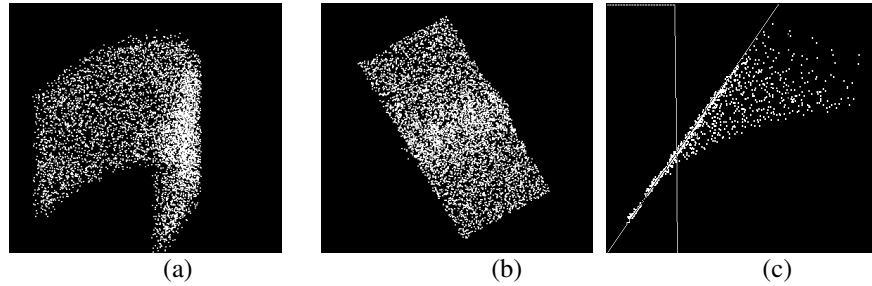

(a)   (b)   (c)

Figure 4: (a) 3D horse shoe dataset (b) 2D CCA projection (c) *dy-dx* plot.

CCA puts more emphasis on maintaining the short distances than the longer ones. Formally, this reasoning leads to the following error function:

$$E = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(d_{i,j}^X - d_{i,j}^Y \right)^2 F_\lambda \left(d_{i,j}^Y \right) \quad (2)$$

where $d_{i,j}^X$ and $d_{i,j}^Y$ are the Euclidean distances between the points i and j in the input space X

and the projected output space Y respectively and N is the number of data points. $F_\lambda \left(d_{i,j}^Y \right)$ is the

neighbourhood function, a monotonically decreasing function of distance. In order to check that the distance relationship is maintained a plot of the distances in the input space and the output space (*dy-dx* plot) is produced. For a well maintained topology, *dy* should be proportional to the

4

value of *dx* at least for small values of *dy*'s. Figure 4 shows CCA projections for the 3D data horse shoe data. The *dy-dx* plot shown is good in the sense that the smaller distances are very well matched [3].

2.3 Intrinsic Dimension

One problem with CCA is deciding how many dimensions the projected space should occupy, and one way of obtaining this is to use the *Intrinsic Dimension (ID)* of the data manifold. The *ID* can be defined as the minimum number of free variables required to define the data without any significant information loss. Due to the possibility of correlations among the data, both linear and nonlinear, a *D*-dimensional dataset may actually lie on a *d*-dimensional manifold ($d \leq D$). The *ID* of such data is then said to be *d*. There are various methods of calculating the *ID*; here we use the correlation Dimension [8] to calculate the *ID* of face image dataset.

2.4 Classification Using Support Vector Machines

A number of classifiers could be used in the final stage for classification. We have concentrated on the Support Vector Machine since this is commonly used in the face processing recognition [12]. Support Vector Machines (SVM) are a set of related supervised learning methods used for classification and regression. SVM's are used extensively for many classification tasks such as: handwritten digit recognition [13] or Object Recognition [14]. A SVM implicitly transforms the data into a higher dimensional data space (determined by the kernel) which allows the classification to be accomplished more easily. We have used the LIBSVM tool [7].

The SVM is trained in the following way:

1. Transform the data to a format required for using the SVM software package – LIBSVM –2.83 [7].

2. Perform simple scaling on the data so that all the features or attributes are in the range [-1, +1].

3. Choose a kernel. We used the RBF kernel,
   $$k(x, y) = e^{-\gamma(x-y)^2}.$$

4. Perform fivefold cross validation with the specified kernel to find the best values of the cost parameter C and γ.

5. By using the best value of C and γ, train the model and finally evaluate the trained classifier using the test sets.

2.5 Effect Size and Encoding face

Effect size is a way of expressing the difference between two groups. Here two groups: a prototypical expression and *Neutral* are used. The Effect size is obtained by finding the difference between the means of the two groups and dividing it by the standard deviation of either group and is denoted by *d* [15]. This can be explained by Equation (3).

$$d = \frac{\mu_1 - \mu_2}{\sigma} \tag{3}$$

5

μ_1 and μ_2 the means of two groups and σ is the standard deviation of the whole population. The 'Encoding face' is obtained by finding the Effect size of each pixel in the two sets of images. In other words for example, which pixels discriminate most between the smiling and neutral faces can be seen and the result of this analysis is reported here.

3. Experiments

3.1 Dataset description

We experimented on 616 faces (308 female and 308 male face images) of 88 individuals with seven basic expressions: happiness, angry, sadness, surprise, fear, disgust and neutral. Apart from neutral all other expressions were selected with the highest degree of intensity for that expression. The classification was done between neutral and one of the expressions at a time. For example: the model classified a test face image as neutral or happy if the classifier was trained for neutral and happy face image classification. Considering one of the six basic expressions (say for example angry) along with neutral, the dataset of 176 images (88 images of *angry* and 88 images of *neutral* set) was divided into 4 equal subsets of 44 images, balanced in terms of gender and expression. The SVM classifier was then trained with 3 subsets at a time and the left out set was used as the test set. A total of 22 male and 22 female face images was used in each set and was balanced, i.e., a person pictured in the neutral set was also present in the angry expression set. Hence at any time, the training set had 132 images. The accuracy was obtained by calculating the average of the classification accuracy for all four subsets used as test sets (when three subsets were used for training, the left out set was used as test set).

The images are from the BINGHAMTON dataset [16] of 100 subjects and some examples are shown in Figure 5. Each row corresponds to the expressions of one of the subjects. They are displayed from left to right in the order: neutral, happy, angry, fear, sad, surprise and disgust.
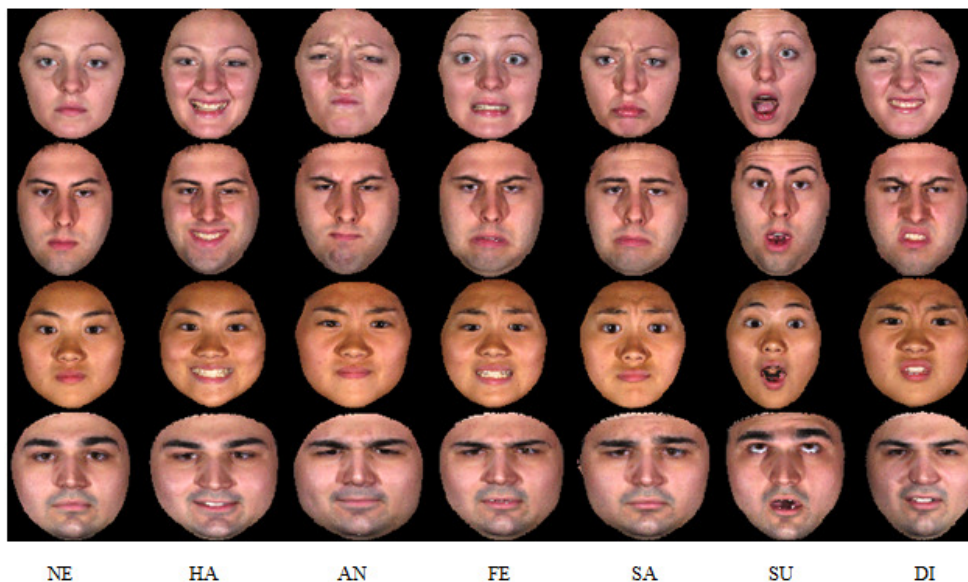

Figure 5: Examples face images from the BINGHAMTON BU-3DFE dataset. Each row is a subject showing various expression (left to right) neutral (NE), happy (HA), angry (AN), fear (FE), sad (SA), surprise (SU) and disgust (DI).

To make images suitable for the experiments, these images had to be reduced from their original size to size 64 x 64 using an image editing tool named Irfanview [17]. The images were also converted into grey scale to lower the computational complexity.

3.2 Method

A total of six experiments were performed with six computational models. Each experiment involved two expressions: one of them was neutral and the other was one of the six basic expressions. The six models that were tested are:

- **RAW**: - Raw face images without any pre-processing or dimensionality reduction
- **RAW+PCA**: - Raw face images without any pre-processing but reduced in dimensionality with PCA.
- **RAW+CCA**: - Raw face images without any pre-processing but reduced in dimensionality with CCA.
- **GAB**: - Gabor pre-processed face images with no dimensionality reduction.
- **GAB+PCA**: - Gabor pre-processed face images reduced by PCA.
- **GAB+CCA**: - Gabor pre-processed face images reduced by CCA.

All these models used an SVM for classification.

4 Results

4.1 Principal Component Analysis

For PCA reduction we always select the first principal components which collectively account for 95% of the total variance of the data, and project the data onto these principal components - we call this is our standard PCA reduction. Figure 6 shows the Eigenfaces obtained by the PCA technique with raw faces (Happy with Neutral set and Angry with Neutral set).
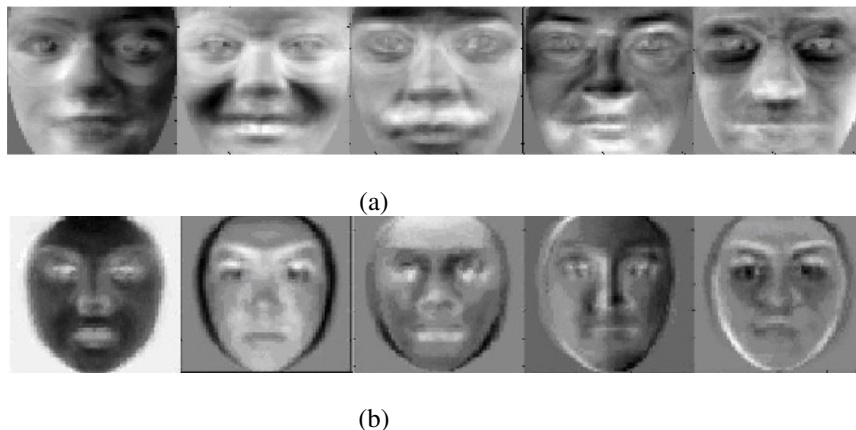

(a)

(b)

Figure 6: a) The first 5 eigenfaces of the neutral and happy data set. B) The first 5 eigenfaces of the neutral and angry data set.

The number of components to which the PCA reduced the original data is detailed in the Table 1.

Table 1: Comparison of number of components used with PCA for raw and Gabor pre-processed face images for all expressions.

Number of components Reduced by PCA	Raw face images	Gabor pre-processed face images
Angry	97	22
Happy	100	23
Fear	99	23
Sad	96	22
Surprise	103	23
Disgust	101	23

It is interesting to note that the GABOR processed images do not need as many components for 95% variance to be captured.

4.2 Curvilinear Component Analysis

As CCA does not provide any measure of captured variance we need a way of deciding an appropriate dimensionality for the projection space. As already discussed in Section 2.3 we estimate the Intrinsic Dimensionality of both the raw faces images and the Gabor pre-processed face images. The results are given in Table 2. A large reduction in dimensionality can be achieved using CCA. The best has just 5 components required for almost all of the Gabor pre-processed face images.

Table 2: Comparison of estimated ID for raw and Gabor pre-processed face images for all expressions.

ID of Dataset	Raw face images	Gabor pre-processed face images
Angry	5	6
Happy	6	5
Fear	6	5
Sad	7	5
Surprise	6	5
Disgust	5	5

4.3 Encoding face and Effect size

The encoding face was obtained by applying the effect size to the pixels of the face image. Two classes were considered at a time: one of the basic expressions and the other being the neutral expression. The discriminating pixels for varying expressions are different. This result supports the evidence of variations in the facial appearance and movements of the facial muscle in response to the expression and in particular, emphasizes those parts of the face corresponding to each of the basic expression [18]. The colored images are shown on the right as they are clearer than their grey scales on the left. The research literature results are in the description given first followed by a comparison of these with the computational model. Figure 7 and 8 show the encoding face for expressions angry and happy respectively.

The encoding face in Figure 7 shows which pixels of the face discriminate most between the angry and neutral classes. Note the changes in the forehead, above and in between the eyebrows and changes in the lip and mouth area. Lowered eyebrows, which may be pulled together forming wrinkles in the skin of the forehead, tension in lips and mouth, all characterize the anger expression. Also, some people have their lowered eyelids tensed and the eyebrows pulled down and may have a glaring look. Others who have a closed mouth form of the angry expression will have a pushing up of the chin [19, 20]. All these areas described are indeed the parts of the angry encoding face that are highlighted showing that the effect size is emphasizing the same areas.
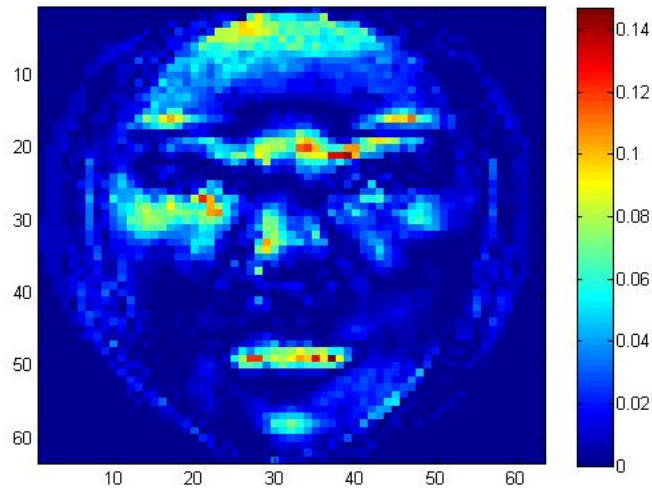

Figure 7: *Angry* encoding face

Figure 8 shows the happy encoding face. Note the changes in the cheeks and the lips. A happy face is normally recognizable with the smile. There is also normally an oblique raising of the lip corners and a wrinkling and creasing of the cheeks. These are defined as the characteristics of the happy expression [19, 20]. In addition to these there is a narrowing of the eyelids, crowfeet wrinkling at the corners of the eye and a raising of the upper areas of the cheeks indicating actual happiness. It may well be that since the dataset that is used here are posed expressions and are not spontaneous expressions; these areas are not very well highlighted.
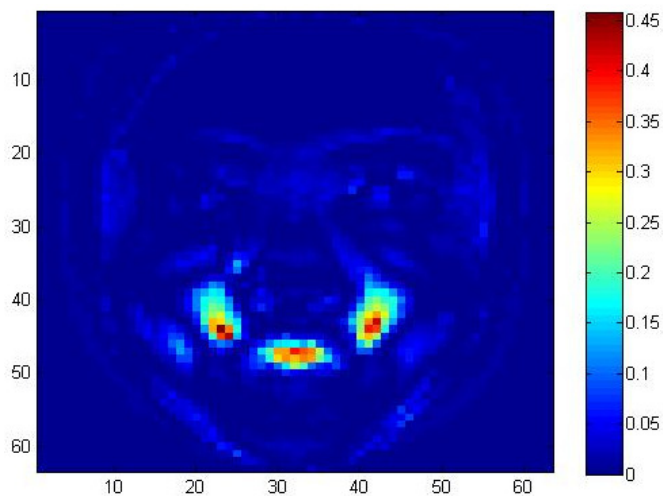

Figure 8: *Happy* encoding face

4.4 Classification by Support Vector Machines

SVM classification was performed by using a 5 fold cross validation on each of the four subsets (described earlier in Section 3.1) and the average accuracy is calculated. All the raw faces and Gabor pre-processed face images have a dimension of size 3969 (63×63); whereas the PCA and CCA dimensionality reductions have lesser dimensions, the details of which are in Tables 1 and 2.

Table 3 shows the accuracy obtained for each expression and also the average accuracy of each model across all the expressions; for example - the average accuracy for RAW models of all expressions is considered. Figure 9 plots the average classification results detailed in Table 3.

Table 3: Average SVM classification accuracy for all models across all basic expressions

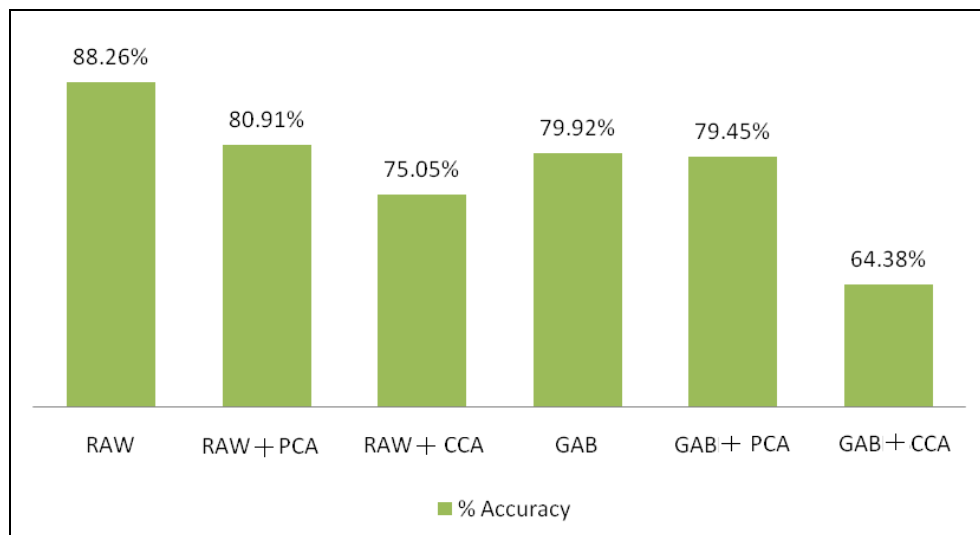| % Accuracy	Angry	Happy	Fear	Sad	Surprise	Disgust	Average
RAW	84.09%	99.43%	83.52%	77.27%	94.89%	90.34%	88.26%
RAW+PCA	70.45%	89%	82.39%	74.43%	89.20%	80%	80.91%
RAW+CCA	63.64%	87.50%	73%	62.50%	93.75%	69.89%	75.05%
GAB	75.57%	89.77%	75.00%	70.45%	95.45%	73.30%	79.92%
GAB+PCA	72.16%	86.93%	79.55%	71.02%	90.34%	76.68%	79.45%
GAB+CCA	66.48%	61.36%	55%	58.52%	84.09%	60.80%	64.38%
Average	72.07%	85.67%	74.74%	69.03%	91.29%	75.17%	

Figure 9: Average classification percentages (last column of Table 3) for each of the six models: RAW, RAW+PCA, RAW+CCA, GAB, GAB+PCA, GAB+CCA for all expressions

10

The average classification accuracy of the RAW model has an outstanding performance in comparison to the rest of them. The RAW model performs best with the *happy* dataset (99.43%) and the least with expression sad (77.27%). An average of 88.26% for the RAW model is the best in comparison with the other models; GAB+CCA being the worst (64.38%). The point here to be noted is that the RAW model did well as predicted due to the high dimensionality and no information loss, unlike other models that have undergone pre-processing (Gabor filtering) and dimensionality reduction (PCA or CCA). In all cases, the PCA reduces the dimensionality of raw faces between 96 and 103 (least for sad and maximum for surprise) whilst the CCA has the most reduction to a mere 5 components (for both *angry* and *disgust*).

The other way to look at the results is to compare expressions. Here it can be seen that *surprise* has been particularly easy for all the models to identify. Conversely *sad* has been difficult for all the models to correctly classify. It is also interesting that the expression *fear* produces a wide variation in the performance of the varying models. The RAW model and the RAW+PCA model do relatively well, but the GAB+CCA model is only just above chance. In fact it appears that the problem here is probably the CCA projection – perhaps the ID of this data set has been incorrectly estimated.

4.5 Discussion

The recognition rates of this system seem to be really encouraging in comparison with other results by various researchers over the past few years [21-25]. Unfortunately, these results cannot be directly compared because there have been large differences in the datasets used, the methods adopted for feature extraction, dimensionality reduction and also the type of classifier used. The lack of literature on similar models to that used here makes the comparisons even more difficult.

The accuracy obtained by the RAW model here is the best of all the models. The accuracy obtained from other computational models such as RAW+PCA, GAB and GAB+PCA have intermediate results. All classification accuracies are above chance.

From a direct comparison of the classification accuracy, the GAB model seems to perform exceptionally well with the expression *surprise* but no so well with other expressions. Overall the classification of the surprise expression, with all models, is extremely good. Although the GAB+CCA model uses just 5 components, the accuracy result is as high as 84.09%. This expression seems to be different from others in that it can be easily detected by any of the models and with very good accuracy.

Comparing the average classification accuracy of raw face images and Gabor pre-processed images that are reduced in dimensionality with PCA, it can be noted that the PCA on Gabor images reduces the dimensionality considerably; but the classification results are similar (80.91% and 79.45% respectively). Comparing the performances of the models with respect to expression shows that surprise and happy expressions appear easier to identify than other expressions.

Figure 10 shows some of the misclassifications in each class. From these examples, it is not clear which feature has caused the misclassification. These experiments will be further extended to compare the performance of human subjects in the classification process.
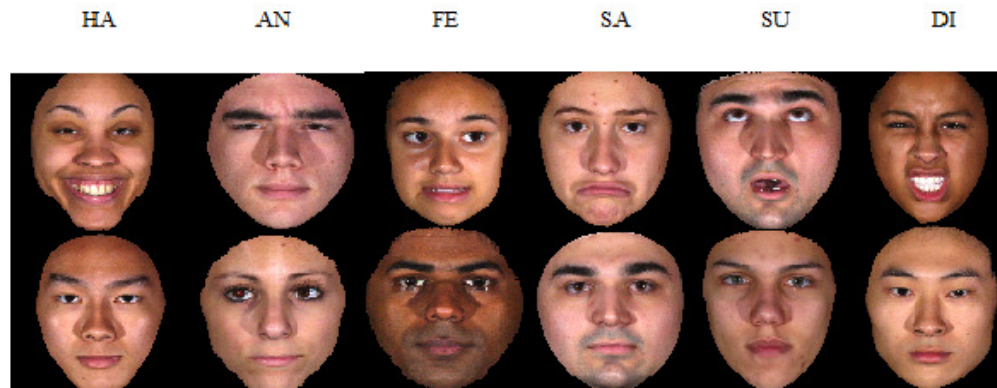
Figure 10: Examples of some of the misclassified faces. The top row shows an example from each of the prototypical expressions that were incorrectly classified as neutral. The bottom row shows neutral faces wrongly classified as the expression labeled above it. (HA - happy, AN - angry, FE - fear, SA - sad, SU - surprise, DI - disgust)

The main point to be noted is that the dimensionality reduction methods used here such as the PCA and CCA in combination with Gabor pre-processing can reduce the original image dimensions to just a few components in comparison to the RAW models. This saves a lot of computational time and also memory space when handling larger databases. Although the raw images have managed to do better in classification accuracy this should be obvious as there has been no dimensionality reduction which could result in information loss. However, when the number of images increases dimensionality reduction will be a necessity and hence methods such as CCA with Gabor pre-processing may become more useful.

5 Conclusions

Identifying facial expressions is a challenging and interesting task. Our experiment shows that identification from raw images can be performed well with all six expressions used here. However, with a larger data set, it may be computationally intractable to use the raw images. It may therefore be important to reduce the dimensionality of the data. The dimensionality reduction methods do fairly well. A linear method such as PCA does not appear to be sufficiently tunable to identify features that are relevant for all facial expression characterization. However, performing Gabor pre-processing on the images increases the classification accuracy of the data after performing PCA. This, however, does not apply to images that are subjected to dimensionality reduction with CCA. Gabor pre-processed PCA data is capable of performing well in comparison to the raw images reduced with PCA. The Gabor pre-processed CCA images, however, with just 5 components does not yield such comparable results. Though the results of the classification for PCA and CCA processed raw images are comparable, it can be noted that Gabor pre-processing has managed to provide good classification with PCA reduced data and with CCA with just 23 and 6 components respectively. These experiments will now be extended to study the performance of human subjects and compared with the performance of the computational models. In these experiments, the SVM has been a preferable choice based on the performance with other classifiers [12]. It will also be interesting to pursue similar work using other Artificial Neural Networks (ANN).

6 References

1. Ekman, P. and W.V. Friesen, *Constants across cultures in the face of the emotion.* Journal of Personality and Social Psychology, 1971. **17**.

2. Batty, B. and M.J. Taylor, *Early processing of the six basic facial emotional expressions.* Cognitive Brain Research, 2003. **17**.

3. Demartines, P. and D.J. Hérault, *Curvilinear component analysis: A self-organizing neural network for nonlinear mapping of data sets* IEEE Transactions on Neural Networks, 1997. **8**(1): p. 148-154.

4. Grassberger, P. and I. Proccacia, *Measuring the strangeness of strange attractors.* Physica D, 1983. **9**.

5. Jain, A.K. and F. Farrokhnia, *Unsupervised texture segmentation using Gabor filters.* Pattern Recognition, 1991. **24**(12).

6. Movellan, J.R. (2002) *Tutorial on Gabor Filters*. **Volume**,

7. Chang, C., C. and C.J. Lin, *LIBSVM: a library for support vector machines.* 2001.

8. Zheng, D., Y. Zhao, and J. Wang, *Features Extraction using A Gabor Filter Family.* Proceedings of the sixth Lasted International conference, Signal and Image processing, Hawaii, 2004.

9. Daugman, J.G., *Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two dimensional visual cortical filters.* Journal of Optical.Society of America A, 1985. **2**(7).

10. Shen, L. and L. Bai, *Review on Gabor wavelets for face recognition* Pattern Analysis Application, 2006. **9**: p. 273-292.

11. Kulikowski, J.J., S. Marcelja, and P.O. Bishop, *Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex.* Biological Cybernetics 1982. **43**(3): p. 187-198.

12. Buchala, S., et al., *Analysis of linear and nonlinear dimensionality reduction methods for gender classification of face images.* International journal of Systems Science, 2005. **36**(14): p. 931-942.

13. Cortes, C. and V. Vapnik, *Support Vector Networks.* Machine Learning, 1995. **20**: p. 273-297.

14. Blanz, V., et al., *Comparison of view-based object recognition algorithms using realistic 3D models.* Proc. Int. Conf. on Artificial Neural Networks 1996: p. 251-256.

15. Cohen, J., *Statistical power analysis for the behavioural sciences.* Second ed. 1988, Hillsdale, New Jersey: Lawrence Earlbaum Associates.

16. Yin, L., Wei, X., Sun, Y., Wang, J. & Rosato, M. J. , *A 3D Facial Expression Database For Facial Behavior Research.* 7th International Conference on Automatic Face and Gesture Recognition (FGR06), 2006

17. skiljan, I., *Irfan View*. 2009.

18. Yacoob, Y. and L.S. Davis, *Recognizing human facial expression.* In Proceedings of Workshop on Visual form 1994.

19. Hager, J.C. *Dataface*. Human face 2006 [cited; Available from: http://www.face-and-emotion.com/dataface/general/homepage.jsp.

20. Ekman, P. and W.V. Friesen, *Unmasking the face. A guide to recognizing emotions from facial clues.* ed. W. V.Englewood Cliffs. 1975, New Jersey: Prentice-Hall.

21. Buenaposada, J.M., E. Enrique Mu˜noz, and L. Baumela, *Recognising facial expressions in video sequences.* Pattern Analysis and Application, 2008. **11**: p. 101-116.

22. Fasel, B. and J. Luettin, *Automatic facial expression analysis: a survey.* Pattern Recognition, 2003. **36**: p. 259-275.

23.   Pantic, M. and M.S. Bartlett (2007) *Machine analysis of Facial expressions*. Face Recognition **Volume**, 377-416

24.   Pantic, M. and L.J.M. Rothkrantz, *Automatic Analysis of Facial Expressions: The State of the Art*   IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000. **22**(12).

25.   Cohn, J.F., et al. *Automated face coding: a computer -vision based method of facial expression analysis*. in *European Conference on facial expression measurement and meaning*. 1997. Salzburg, Austria.