Principal Component Analysis of Gender, Ethnicity, Age, and Identity of Face Images

Samarasena Buchala[1], Neil Davey[1], Tim M. Gale[1,2], Ray J. Frank[1]
[1] Department of Computer Science, University of Hertfordshire, College Lane, Hatfield, AL10 9AB, UK
{S.Buchala, N.Davey, R.J.Frank, T.Gale}@herts.ac.uk
[2] Department of Psychiatry, QEII Hospital, Welwyn Garden City, AL7 4HQ, UK

***Abstract*—Principal Component Analysis (PCA) has been widely used for efficient representation of face images data in a low dimensional subspace. In this study, we use PCA to analyse different properties of faces, such as gender, ethnicity, age and identity. Using Linear Discriminant Analysis (LDA), we show that PCA efficiently encodes information related to different properties, different components of PCA encode different information, and there may be components which encode information related to multiple properties.**

I. Introduction

The human face is considered to be special in terms of its biological and social roles, and is rich in information. Faces have multiple properties based on which they can be categorized at different levels of specificity, such as gender, ethnicity, age, expression, identity, degree of attractiveness, typicality and distinctiveness, and so on. Due to the fundamental importance of face recognition and categorisation in every-day life, this is one of the most researched topics in the fields of Psychology and Computer Science. While the research in Psychology is largely related to face perception, Computer Science research is related more to face detection and recognition in a Biometric scenario. There are also psychologically motivated studies that use Computer Science methodologies.

Face images are very high dimensional and usually contain redundancies. According to the "curse of dimensionality" [1], an impractically large number of examples would be needed for analysis of such data. To overcome this problem, feature extraction is usually applied on face images before any further task. Feature extraction efficiently transforms the data into a lower dimensional subspace by reducing the redundancy in the data. Feature extraction also makes sense from a neurophysiological point of view as there is evidence that redundancy reduction is an important part of sensory processing in human brain [2].

There have been various feature extraction methods used in the literature of face recognition: Principal Component Analysis (PCA) [3], Independent Component Analysis [4], Elastic Bunch Graph Matching [5], and recently Non-negative Matrix Factorization [6]. PCA is the one most widely used in face recognition [7-17]. In all of these studies, PCA has been shown to encode efficiently the face properties of interest. Many of these studies [12-17] also suggest that PCA encodes face information in a psychologically plausible manner. However, all these studies use small datasets with less variation, and except for [16], analyse data with respect to one or two properties of faces. In this paper we use much larger number of faces and test if PCA encodes properties such as gender, ethnicity, age, and identity efficiently. Using Fisher's Linear Discriminant Analysis (LDA), we also analyse how these different properties vary on the different components of PCA.

The main findings, with respect to the above aims, of this paper are

1. PCA encodes face image properties such as gender, ethnicity, age, and identity efficiently.
2. Different components of PCA encode different properties of faces. Very few components are required to encode properties such as gender, ethnicity and age and these components are predominantly amongst the first few components which capture large part of the variance of the data. Large number of components are required to encode identity and these components are widely distributed.
3. There may be components which encode multiple properties.

The remainder of the paper is organised as follows: A brief overview of the literature of face recognition using PCA is given in the next section. Sections III and IV presents a brief description of PCA and LDA methods. Section V presents the experimental results. We conclude in Section VI.

II. Principal Component Analysis and Face Recognition

Sirovich and Kirby [18] showed that PCA can be applied for efficient representation of high dimensional face images data in a lower dimensional subspace. Turk and Pentland [3] extended this to apply for face recognition. Since then PCA has become a basis, and also a benchmark, for numerous face recognition algorithms [19]. Studies in Psychophysics [10-16] have also shown keen interest in PCA. For example, PCA is shown to account for distinctiveness effects of face perception (where distinct faces can be recognized easily compared to typical faces) [14] and [15], "other-race effect" (where faces of different race from ones own race are difficult to recognise) [20], [21], Dimensional-based model of facial expression (where different expressions are thought to be in a continuum rather than in distinct and independent categories.). Recent research, [16] and [17], also posits an optimistic view that PCA can be used to account for some aspects of the perceptual functions of face recognition proposed by [22].

III. Principal Component Analysis

The aim of the PCA is a linear transformation of a D dimensional data X into an uncorrelated d dimensional data Y, where $d \leq D$.

Hypothetically, the first step of the PCA is to find a linear function Y_1 accounting for the maximum possible variance in the data such that

$$Y_1 = \sum_{i=1}^{D} w_{1i} x_i \tag{1}$$

Where w is a weight vector.

The next step is to find a linear function Y_2 which is orthogonal to the first one and accounts for the next most possible variance in the data. The D^{th} linear function would be Y_D which accounts for the D^{th} maximum variance in the data and is orthogonal to the first $D - 1$ linear functions. However, it is hoped that only the first d linear functions would account for most of the variance in the data. In matrix notation Equation 1 can be written as

$$Y = W^T X \tag{2}$$

Mathematically, PCA can be achieved by estimating the Eigenvectors and Eigenvalues of the covariance matrix

$$\frac{1}{N}(X - \overline{X})(X - \overline{X})^T \tag{3}$$

where, $\overline{X}$ is the mean of the dataset X and N is the number of datapoints in the dataset. The Eigenvectors are the weight matrix W. The Eigenvalues, λ, characterize the variance accounted by the corresponding eigenvectors and signifies their importance in defining the data. The variance of an Eigenvector W_i can be calculated as

$$\sigma_i^2 = \frac{\lambda_i}{\sum_{j=1}^{D} \lambda_j} \tag{4}$$

We refer to the variables Y and W of Equation 2 as Principal Components (PCs) and Eigenvectors or Eigenfaces respectively.

For further details of PCA, see [23].

IV. Linear Discriminant Analysis

The usual choice for selecting the components to define the data without significant information loss is to order the components according to their importance in accounting the variance of the data and consider the first few components which account for some percentage, usually above 80 %, of the cumulative variance. However, if the properties of interest of the data are encoded by the last few components, even if they are not significant in defining the data, this method is disadvantageous.

Data usually have multiple properties. For example, the data shown in Fig. 1 have different properties and can be classified by colour – red, blue, by shape – circle, square, or both colour and shape. PCA on the data would indicate maximum variance in the direction of W_1. This first PC encodes only shape information and it fails if the property of interest is colour, whereas the second component in the direction of W_2, though accounts lesser variance, would be successful in this regard. This shows that the selection of the components should be based on its importance for a given task, rather than its importance in accounting the total variance. This example would be apt for face data which have multiple properties, such as identity, gender, ethnicity, age, expression and so on.

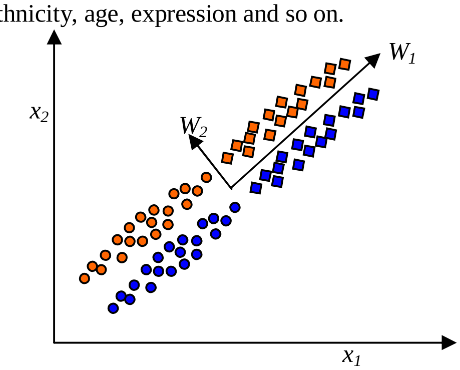


Figure 1. An illustration of PCA on a dataset with multiple properties. The first component in the direction of W_1, though accounts for the maximum variance, fails to encode information regarding the colour of the data, whereas the second component in the direction W_2, would be successful.

The problem discussed above can be resolved using Fisher's LDA. It takes into account both the between-class scatter as well as within-class scatter of the data.

The between-class scatter, S_B, is given by

$$S_B = \sum_{i=1}^{C} (M_i - M)(M_i - M)^T \tag{5}$$

where C is the number of classes, and M_i and M are the means of class i and the whole data.

The within-class scatter, S_W, is given by

$$S_W = \sum_{n \in C_1} (X^n - M_1)(X^n - M_1)^T + \sum_{n \in C_2} (X^n - M_2)(X^n - M_2)^T \tag{6}$$

The class separation ability (information encoding power) of a representation can be estimated using various choices. One of the examples is given by

$$J_W = Tr\{J\} \tag{7}$$

where $Tr\{M\}$ denotes the trace of the matrix M and

$$J = \frac{S_B}{S_W} = S_W^{-1} S_B . \tag{8}$$

Since we are only interested in the information encoding abilities of individual components, rather than whole representation obtained by PCA, we estimate the encoding power of the individual components by the following

$$P_i = \frac{(Diag\{S_B\})_i}{(Diag\{S_W\})_i} \tag{9}$$

where $Diag\{M\}$ denotes the diagonal matrix of M. We refer to P as the *Encoding Power*.

The above framework is used in Fisher's LDA where the original data are transformed in LDA space to achieve better class separability. The transformation matrix is determined by the Eigenvectors of Equation 8, by considering the largest corresponding Eigenvalues. This approach was used for face recognition in [24]. Due to problems related to the high dimensionality of face images and the usual availability of few face examples, [25] and [26] have applied PCA for dimensionality reduction before applying LDA.

V. EXPERIMENTS

A. Dataset

We use a subset of FERET [27] face image database for this research. The FERET database has a large number of face images and is rich in variety, with different categories of gender, ethnicity, age, and identity being well represented. The face images also vary in pose, background lighting, presence or absence of eye glasses, slight changes in expression. Due to these properties, FERET has become a standard database to compare the performance of different face recognition algorithms [27-30]. In this study of analysis of different face properties we select the dataset, a subset of FERET, such that it contains 2670 grey scale frontal face images. An overview of the representation of different categories of gender, ethnicity, age and identity properties in the dataset is given in Table I.

TABLE I. REPPRESENTATION OF DIFFERENT PROPERTIES IN THE DATASET

Property	**No. Categories**	**Categories**	**No. Faces**
Gender	2	Male	1603
		Female	1067
Ethnicity	3	Caucasian	1758
		African	320
		East Asian	363
Age	5	20 – 29	665
		30 – 39	1264
		40 – 49	429
		50 – 59	206
		60+	106
Identity	358	Individuals with 3 or more Examples	1161

Each image is preprocessed to a 65×75 resolution, cropped such that little or no hair is visible. Faces are aligned with each other based on their eye locations and histogram equalization is applied to reduce the lighting effects that may result from different lighting conditions. A few examples of the dataset are shown in Fig 2.

Figure 2. Examples of the dataset

B. PCA of Face Images

The rows and columns of each image are concatenated into a single string vector. The dataset, containing 2670 images of 65×75 resolution, gave rise to 2670 vectors of length 4875. PCA on this data would give 2670 Eigenvectors of length 4875. Usually, only a few Eigenvectors account for most of the variance in the data. Only 350 Eigenvectors accounted for 90% of the variance. Each face can thus be efficiently represented, without significant loss of information, using just 350 components instead of 4875 dimensions. As these Eigenvectors act as a basis upon which each face's variation is captured, they appear face like when visualized. Hence, they are also termed as Eigenfaces. Fig. 3 shows some of the Eigenfaces of the dataset.

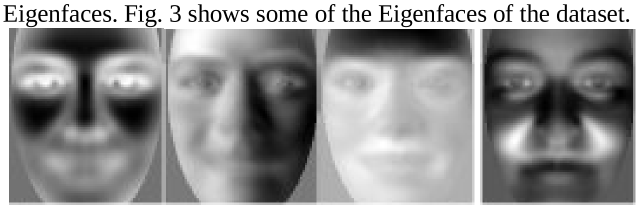


Figure 3. Eigenfaces 1 to 4: from left to right

1) Gender

Due the statistical nature of PCA, the first few components encode information common to most faces and the last components encode information common to least faces. This lead O'Toole et al [14] to claim that information related to the gender property is shared by most of the faces and hence it is encoded in the first few components. And information related to identity is shared by only a few faces and hence is encoded in the last components. In Fig. 4 the top row, left to right, shows a female face and its reconstructed images using the first 50 components and components from 51 to 350. Similarly, the bottom row images are of a male. The gender of both faces can be more easily judged from the reconstructed images which use the first 50 components, while the last components seems to encode finer information of the face and hence may be more useful for identity. This idea of different sets of components encoding gender and identity information also supports the Bruce and Young's functional model of face recognition [22], which proposes that gender and identity are handled by different perceptual components of the human cognitive system.

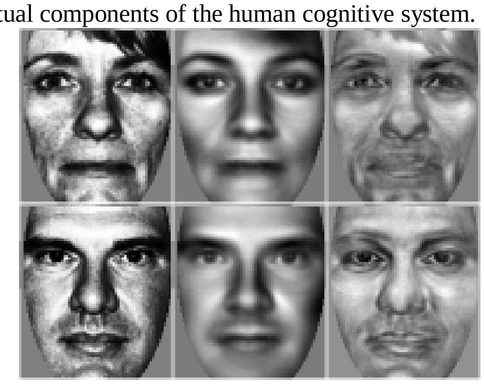


Figure 4. Top row from left to right: A female face and its reconstructions using the first 50 components and components from 51to 350. Bottom row from left to right: A male face and its reconstructions using the first 50 components amd components from 51 to 350.

Do all initial components carry high gender information? To find this, we estimate the gender encoding power of different components using (9). Fig. 5 shows a plot of the gender encoding power of the first 50 components (the rest of the components are not found to be significant and hence are not shown). It can be seen that the third component carries highest gender encoding power, followed by the fourth component. O'Toole et al [14] reports that the second component, which largely encodes presence or absence of hair, on their dataset to be the most important for gender classification. In our previous study [31], on a dataset which includes hair information, we found that the most important component for gender is the second component, which captures the hair information.

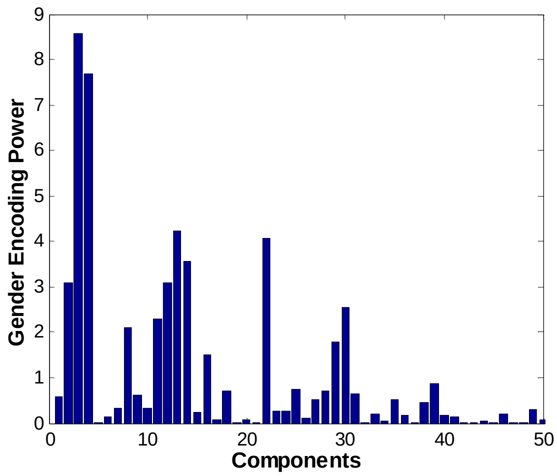
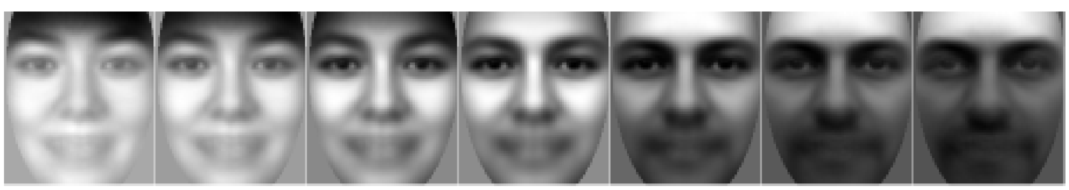
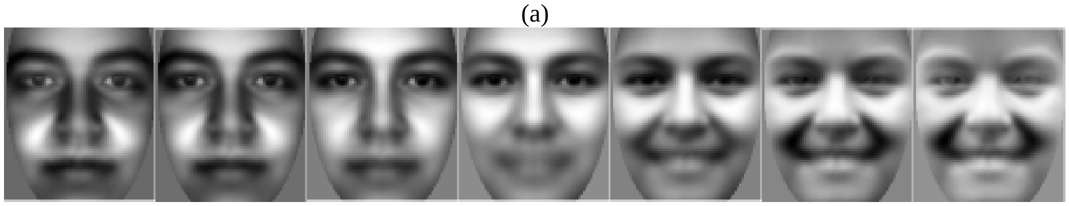


Figure 5. Gender encoding power of the first 50 components.

It is often difficult to discern what information each Eigenface encodes, as is the case with the Eigenfaces shown in Fig. 3. In order to determine the information encoded by the third and fourth components, which are found to be important for gender (from Fig. 5), we generate a series of reconstructed images. We first estimate the average of the components of all faces and then produce a reconstructed image from these average components. The resultant image is an average face. We then alter the component of interest, third or fourth, while keeping the other components of the average components unchanged, to see what changes it makes to the average face. We add or subtract, progressively, more quantities of Eigenfaces 3 or 4 to capture its effects. In Fig 6 (a) on the extreme left is a reconstructed face with 6 S.D of the third component removed from it and the extreme right face is a reconstructed face with 6 S.D of the third component added to it. The extreme left face of Fig. 6(a) (-6 S.D) appears more feminine, while the extreme right face (+6 S.D) appears more masculine. The feminine face complexion looks lighter than the masculine face (this may be due to the cosmetics used by females) and the masculine face is much darker around the mouth region. The masculine face has a longer nose and its forehead is lighter compared to the feminine face (this is due to the reason that most females in the dataset having hair falling on their forehead). More specifically, the third component encodes information related to the complexion, length of the nose, presence or absence of hair on the forehead. Similarly, Fig. 6(b) shows a series of images for the fourth component. This component encodes information related to the eyebrow thickness. Eyebrow information is important for gender recognition, as many females tend to have thinner eyebrows. The fourth component also encodes information related to presence or absence of smiling expression. One reason that this might differentiate males and females of the dataset is due to the artefact that many females, compared to males, of the dataset have a smiling pose, when their picture is taken. This artefact, however, is not unique to our dataset. Social Psychology research widely reports that women, across cultures, smile more often than men[32, 33].



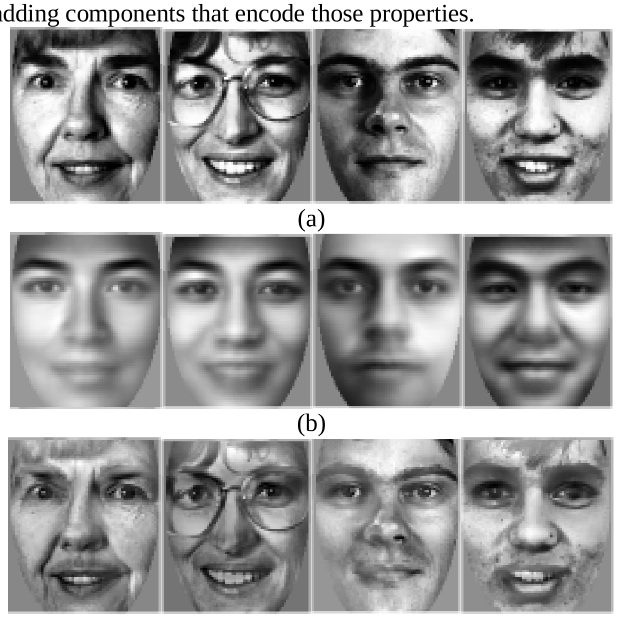
(a)



(b)

Figure 6. Reconstructed images from the altered components (a) third and (b) fourth components. The components are progressively added by quantities of -6 S.D (extreme left) to +6 S.D (extreme right) in steps of 2 S.D.

Fig. 7 shows a few face examples and a series of their reconstructed images using 20 gender important components (components with highest gender encoding power), all but the 20 gender important components. In Fig. 7 (b) the gender of the reconstructed faces, using 20 gender important components, is apparent while its identity is not. And in the case of Fig. 7(c), which uses all but the 20 gender important components, the gender identification is difficult, while its identity is still possible. It can also be noticed from Fig. 7(c) that properties which do not carry gender information are de-emphasized. For example eye glasses are de-emphasized. This also shows that invariances to properties can be built by selectively removing or adding components that encode those properties.



(a)

(b)

(c)

Figure 7. (a) Face examples with the first two being female and the next two being male faces. (b) Reconstructed faces of (a) using the top 20 gender important components. (c) Reconstructed faces of (a) using all components, except the top 20 gender important components.

2) Ethnicity

As with the analysis of gender, we perform ethnicity analysis. Fig. 9 shows a plot of the ethnicity encoding power of the first 50 components (the rest of the components, similar with the case of gender, are not found to be significant and hence are not shown). The sixth component is found to be having the highest ethnicity encoding power.

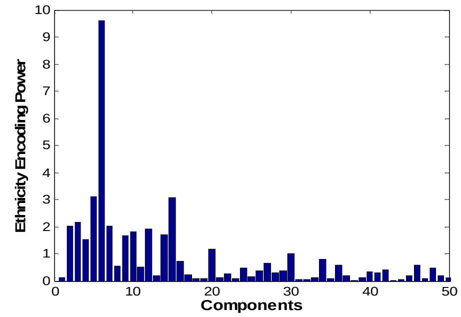
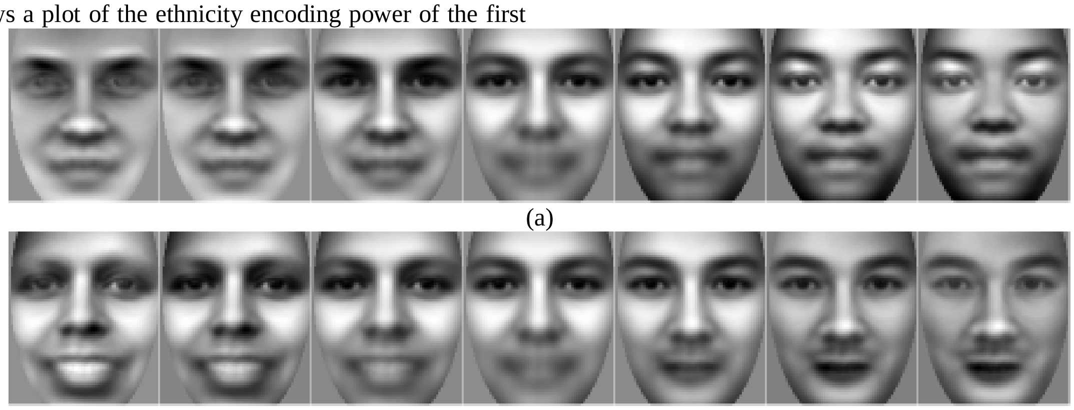


Figure 8. Ethnicity encoding power of the first 50 components.

To determine the information captured by the ethnicity important components, we produce a series of images, similar with that of the gender case. A series of the reconstructed images from altering the sixth and the fifteenth components are shown in Fig 9(a) and 9(b) respectively. The extreme left face of Fig. 9(a) (-6 S.D) appears more Caucasian like, while the extreme right face (+6 S.D) appears more African like. The African like face has a darker complexion, flatter and shorter nose compared to the Caucasian like face. Similarly, Fig. 9(b) shows a series of images for the fifteenth component. The extreme left face of Fig. 9(b) (-6 S.D) appears more African like, while the extreme right face (+6 S.D) appears more East Asian like. The African like face has a shorter nose and smiling expression compared to the East Asian like face. One reason that the smiling expression might differentiate African faces from other ethnicities is due to the fact that many African faces in the dataset had a smiling expression. This artefact is similar to that found for the female faces. However, to the best of our knowledge, there are no psychological studies, similar to that of gender [32, 33], which report that Africans smile more often than people from the other ethnicities.

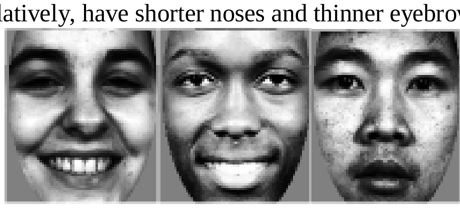


(a)

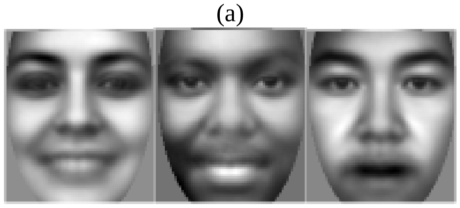
(b)

Figure 9. Reconstructed images from the altered components (a) sixth and (b) fifteenth components. The components are progressively added by quantities of -6 S.D (extreme left) to +6 S.D (extreme right) in steps of 2 S.D.

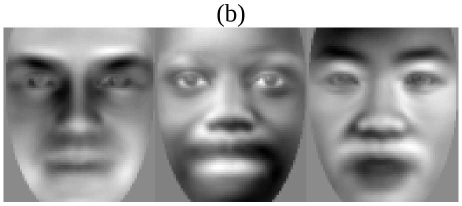
Fig. 10 shows face examples of Caucasian, African, and East Asian ethnicities, and a series of their reconstructed images using the 20 ethnicity important components, and average of each ethnicity's reconstructed faces, which uses 20 top ethnicity important components. The reconstructed faces, using 20 ethnicity important components, are shown in Fig 10(b). The ethnicities of these reconstructed faces are apparent suggesting good ethnicity encoding ability of these 20 components. To determine how faces of different ethnicities vary on these 20 ethnicity components we estimate the average of each ethnicity's reconstructed faces, using 20 ethnicity important components. To make the features, which differentiate between the ethnicities, more prominent, we stretch each of these average faces from the average face of the whole dataset. These are shown in Fig. 10(c). From Fig. 10(c), we can make the following conclusions: generally, both the Caucasian and East Asian faces have lighter complexion compared to the African faces. East Asian face has much lighter skin around the mouth region. This may be due to the fact that, relatively, very few East Asian males in the dataset have beard or moustache. The East Asian face also differs in the distance between eyelids and eyebrows, suggested by lighter pixels in this region. African faces, relatively, have shorter noses and thinner eyebrows.



(a)



(b)



(c)

Figure 10. (a) Face examples from left to right: Caucasian, African, East Aian. (b) Reconstructed faces of (a) using the top 20 ethnicity important components (c) Average of the reconstructed Caucasian, African, and East Asian faces using the top 20 ethnicity important components.

3) Age

Accurate age estimation from face images is, perhaps, the most difficult of the face recognition tasks. Though, we find very difficult to estimate the age of a face accurately, we are good at judging whether the face is young or old. For this reason, for age analysis, we only consider faces which fall in the following age groups of the dataset: 20-39 and 50-60+. We term them as *young* and *old* faces respectively. As with the analysis of gender and ethnicity properties, we perform an analysis of age information. Fig. 11 shows a plot of the age encoding power of the first 50 components (the rest of the components, similar with the case of gender and ethnicity, are not found to be significant and hence are not shown). The tenth component is found to be having the highest age encoding power.

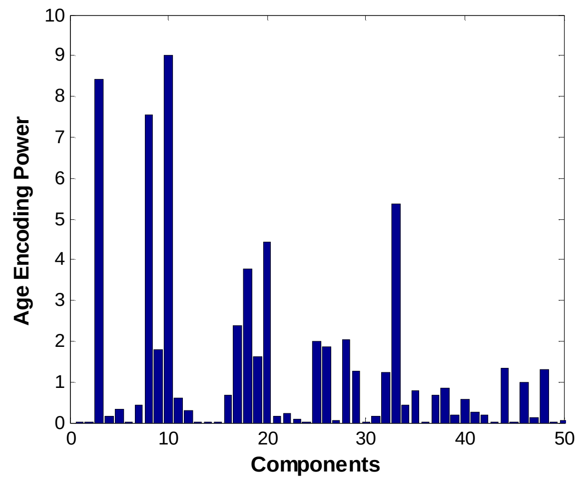


Figure 11. Age encoding power of the first 50 components.

We produce a series of images in Fig. 12, similar with that of the gender and ethnicity cases, to discern the information captured by the 10^{th} component. The extreme left face of Fig. 12 (-6 S.D) has lighter complexion compared to the extreme right face (+6 S.D). The age information captured by this component is not as obvious as that of the important components of gender and ethnicity. However, there are some subtle differences, which can be observed upon careful inspection. The region above the eyes, in the extreme left face, indicates sagging skin and the eyebrows of this face are also thinner compared to the extreme right face. This suggests that the face at the extreme left is related to the old group and the face at the other end of the series is related to the young group.

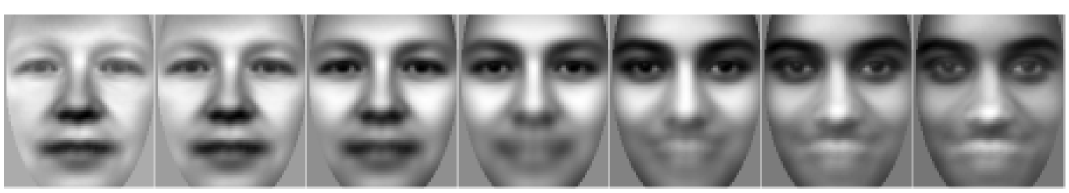


Figure 12. Reconstructed images from the altered tenth component. The component is progressively added by quantities of -6 S.D (extreme left) to +6 S.D (extreme right) in steps of 2 S.D.

4) Identity

Identity is different from gender, ethnicity, and age in the sense that the information related to the latter properties is shared by large number of faces in the population, while identity information is distinct and common to a single person or at most to a few similar looking individuals. In Fig. 13, we plot the identity encoding power of the first 100 components. In contrast to the other properties cases, many components are found to have, relatively, higher identity encoding power. These components are widely distributed and are not restricted to the first 50 components. Another contrasting feature that can be noticed from Fig. 13 is that the magnitude of these high identity encoding components is not as high as in the case of the important components of the other properties, and hence not highly significant on their own. It can be thus concluded that large number of components are needed to encode identity information.

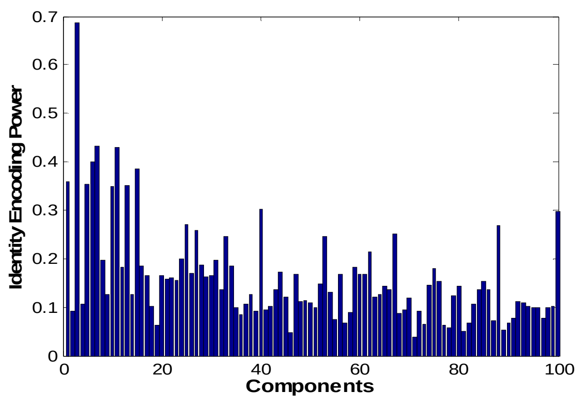
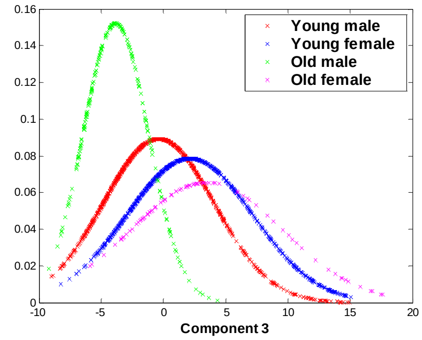


Figure 13. Identity encoding power, P, of the first 100 components.

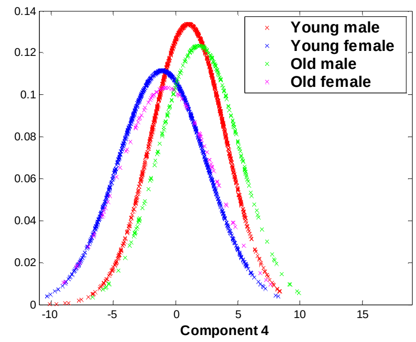
5) Can a single component encode multiple properties?

There may be features which elicit information related to more than one property. For example, a grey beard informs that the person is a male and also, most probably, old. In this section we try to find if there are any components which encode significant information related to more than one property.

The components with high encoding power for gender, ethnicity, and age are restricted to the first 50 components. Due to this, it can be noted from figures 5, 8, and 11, that a few number of components are found to be important for more than one property. One such example is the 3rd component which is found to be the most important for gender (Fig. 5) and the second most important for age (Fig. 11). This is also illustrated in Fig. 14. Fig. 14(a) shows normal distribution plots of the third component for male and female classes of young and old age groups. The distribution shows considerable class separability between males and females and also between young males and old males. However, the young and old females are nearly overlapped. To demonstrate the importance of the third component for gender and age, we show a similar normal distribution plot of the fourth component in Fig. 14(b). The fourth component is found to be the second most important for gender, but not significant for age. From the plot it can be seen that the fourth component shows near complete overlap of young males and females with their old counterparts, though it shows considerable separability between males and females.



(a)



(b)

Figure 14. Normal distribution plots of the (a) third (b) and fourth components for male and female classes of young and old age groups.

6) Classification results

To test if the information related to all the properties is efficiently represented by the PCA data, we perform classification based on different properties. For each property, except identity, 80% of the faces are used for training and 20% are used for testing. In the case of identity, we use leave-one-out strategy for classification. First, LDA is performed on the PCA data. On the resultant data, for classification, we use a simple Euclidean measure between the test data and the means of the various classes estimated from the training data.

The classification results are shown in Table II (N.B: Identity – a, Identity – b, Identity – c are related to when only individuals in the dataset with face examples >=3, >=4, and >=5 are considered). Classification performances on all properties are reasonably high and much above chance levels suggesting that information is efficiently encoded by PCA.

TABLE II. CLASSIFICATION PERFORMANCE ON DIFFERENT PROPERTIES

Property	Classification %
Gender	86.43
Ethnicity	81.67
Age	91.5
Identity – a	68.7
Identity – b	90
Identity – c	100

VI. CONCLUSION

In this paper, we presented an analysis of different properties of faces, such as gender, ethnicity, age, and identity, using PCA on face images data. Using LDA, we estimated the encoding powers, with respect to different properties, of the components obtained by PCA. Using reconstructed images, we also presented what information is captured by each important component of a property.

To summarize the main findings of the paper:

1. PCA encodes face image properties such as gender, ethnicity, age, and identity efficiently – the classification performances on all properties are reasonably high and much above chance levels.
2. Different components of PCA encode different properties of faces. Very few components are required to encode properties such as gender, ethnicity and age and these components are predominantly amongst the first few components which capture large part of the variance of the data. Large number of components are required to encode identity and these components are widely distributed.
3. There may be components which encode multiple properties – for example the third component is found to be important for gender as well as age.

REFERENCES

[1] R. E. Bellman, *Adaptive control processes: A guided tour*: Princeton University Press, 1961.

[2] H. B. Barlow, "Unsupervised learning," *Neural Computation*, vol. 1, pp. 295-311, 1989.

[3] M. Turk and A. Pentland, "Face recognition using eigen faces," presented at IEEE International Conference on pattern Recognition, Maui, Hawaii, 1991.

[4] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Transactions on Neural Networks*, vol. 13, 2002.

[5] L. Wiskott, J. M. Fellous, and C. v. d. Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 775-779, 1997.

[6] D. Guillamet and J. Vitria, "Classifying faces with non-negative matrix factorization," presented at 5th Catalan Conference for Artificial Intelligence, Castello de la Plana, 2002.

[7] M. Turk and A. Pentland, "EigenFaces for recognition," *J.Cognitive Neuroscience*, vol. 3, pp. 71-86, 1991.

[8] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 696-710, 1997.

[9] A. J. O'Toole, H. Abdi, K. A. Deffenbacher, and J. C. Bartlett, "Classifying faces by race and sex using an autoassociative memory trained for recognition," presented at Thirteenth Annual Conference of the Cognitive Science Society, Hillsdale, N.J, 1991.

[10] A. J. O'Toole, J. Peterson, and K. A. Deffenbacher, "An 'other-race effect' for categorizing faces by sex," *Perception*, vol. 25, pp. 669-676, 1996.

[11] D. Valentin, H. Abdi, B. Edelman, and A. J. O'Toole, "Principal component and neural network analyses of face images: what can be generalized in gender classification," *Journal of Mathematical Psychology*, vol. 42, pp. 175-195, 1997.

[12] H. Abdi, D. Valentin, and B. Edelman, "Eigenfeatures as intermediate level representations: The case of PCA models," *Brain and Behavioral Sciences*, vol. 21, 1997.

[13] A. J. O'Toole, H. Abdi, K. A. Deffenbacher, and D. Valentin, "A perceptual learning theory of the information in faces," in *Cognitive and computational aspects of face recognition*, T. Valentine, Ed. London: Routledge, 1995.

[14] A. J. O'Toole, K. A. Deffenbacher, D. Valentin, and H. Abdi, "Structural aspects of face recognition and the other race effect," *Memory and Cognition*, vol. 26, pp. 146-160, 1994.

[15] P. J. B. Hancock, "Face processing: human perception and principal component analysis," *Memory and Cognition*, vol. 24, pp. 26-40, 1996.

[16] A. J. Calder, A. M. Burton, P. Miller, A. W. Young, and S. Akamatsu, "A principal component analysis of facial expressions," *Vision Research*, vol. 41, pp. 1179-1208, 2001.

[17] A. M. Burton and P. J. B. Hancock, "From pixels to people: a model of familiar face recognition," *Cognitive Science*, vol. 23, pp. 1-31, 1999.

[18] L. Sirovich and M. Kirby, "Low -dimensional procedure for the characterization of human faces," *Journal of the Optical Society of America A*, vol. 4, pp. 519-524, 1987.

[19] H. Moon and P. J. Phillips, "Computational and performance aspects of PCA-based face-recognition algorithms," *Perception*, vol. 30, 2001.

[20] A. J. O'Toole, K. A. Deffenbacher, H. Abdi, and J. C. Bartlett, "Simulating the other race effect as a problem in perceptual learning," *Connection Sciences*, vol. 3, pp. 163-178, 1991.

[21] N. Furl, P. J. Phillips, and A. J. O'Toole, "Face recognition algorithms and the other-race effect: computational mechanisms for a developmental contact hypothesis," *Cognitive Science*, vol. 26, pp. 797-815, 2002.

[22] V. Bruce and A. Young, "Understanding face recognition," *British Journal of Psychology*, vol. 77, pp. 305-327, 1986.

[23] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York: Springer-Verlag, 2002.

[24] P. N. Belhumeur, J. P. Hespanha, and D. J. Kreigman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," presented at European Conference on Computer Vision, 1996.

[25] D. L. Swets and J. Weng, "Using discriminant eigenfeatures for image retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 831-836, 1996.

[26] W. Zhao, A. Krishnaswamy, R. Chellappa, D. L. Swets, and J. Weng, "Disriminant analysis of principal components for face recognition," presented at Third International Conference on Face and Gesture Recognition, 1998.

[27] P. J. Phillips, H. Wechsler, J. Huang, and P. Rauss, "The FERET database and evaluation procedure for face recognition algorithms," *Image and Vision Computing*, vol. 16, pp. 295-306, 1998.

[28] P. J. Phillips, H. Moon, P. Rauss, and S. Rizvi, "The FERET evaluation methodology for face recognition algorithms," presented at Computer Vision and Pattern Recognition 97, Los Alamitos, CA, 1997.

[29] S. Rizvi, P. J. Phillips, and H. Moon, "A verification protocol and statistical performance analysis for face recognition algorithms," presented at Computer Vision and Pattern Recognition, Los Alamitos, CA, 1998.

[30] P. J. Phillips, H. Moon, S. Rizvi, and P. Rauss, "The FERET evaluation methodology for face recognition algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1090-1104, 2000.

[31] S. Buchala, N. Davey, T. M. Gale, and R. J. Frank, "Principal component analysis of face properties," University of Hertfordshire, Technical Report --, 2005.

[32] D. K. Dodd, B. L. Russell, and C. Jenkins, "Smiling in school yearbook photos: Gender differences from kindergarten to adulthuood," in *The Psychological Record*, vol. 49, 1999, pp. 543-554.

[33] M. LaFrance and M. A. Hecht, "Gender and smiling: A meta-analysis," in *Gender and emotion: Social psychological perspectives*, A. H. Fischer, Ed. Paris: Cambridge University Press, 2000, pp. 118-142.