GLOBAL FEATURE BASED FEMALE FACIAL BEAUTY DECISION SYSTEM

H. İrem TÜRKMEN irem@ce.yildiz.edu.tr
Zeyneb KURT zeyneb@ce.yildiz.edu.tr
M. Elif KARSLIGİL elif@ce.yildiz.edu.tr

Computer Engineering Department
Yıldız Technical University, İSTANBUL/TURKEY

ABSTRACT

This paper presents an automated female facial beauty decision system based on Support Vector Machine (SVM). First, we constructed manually two classes of female faces with respect to their facial beauty, by requesting personal opinions of people. As the second step, Principal Components Analysis (PCA) and Kernel PCA(KPCA) were applied to each class for extracting principal features of beauty. Support Vector Machine (SVM) was used for judging whether a given face is beautiful or not. Since judging the beauty is subjective, the decision results of our system were evaluated by comparing the system generated decision results with the corresponding ones made by the persons. Based on this criteria, our results showed that KPCA with a success ratio of 89% outperformed PCA with a success ratio of 83%.

1. Introduction

Beauty is the quality that pleases senses or lifts up the mind or spirit. Although facial beauty is a subjective concept that has different meaning for everyone, there is a universal agreement that a beautiful face should have a pleasant proportion. This proportion can be measured with a mathematical ratio called "golden proportion" which equals 1:1.608(fi number [1]).

There are few implementations of automated facial beauty determination. Eisenthall, Dror and Ruppin [2] studied PCA and k-Nearest Neighbourhood to propose a system to learn and analyze the mapping from two dimensional facial images to corresponding attractiveness scores, as rated by humans. H. Güneş, M.Y. Karslıgil [3] proposed a method for the analysis of facial beauty by golden ratio.

In this paper we propose a global feature based facial beauty decision system. The system is composed of two major parts: Learning facial beauty and judging facial beauty of a given face.

First, 100 female face pictures were shown to 50 persons for classifying them as beautiful or not in training phase. Then faces were grouped into two classes according to their votes of majority. People determine facial beauty by looking at the whole face globally; instead of analyzing local characteristics such as width of the nose. Therefore PCA and KPCA has been applied in face recognition for extracting the global characteristic features of the faces. Then, Support Vector Machine (SVM) was applied to train and decide the facial beauty. In the testing phase, 80 random female faces, which are not included in the constructed database, were submitted to the system and were shown to the same 50 persons from the beginning of the initial step for beauty evaluation. The decisions of the system and those 50 persons, were compared to measure the performance of the proposed system.

The outline of the paper is as follows: Section 2 presents preprocessing steps of the system. Section 3 describes extraction of the characteristic features of images using PCA and Kernel PCA. In section 4, training and decision steps of the system are described. Finally in section 5, experimental results of our implementation are given, for every individual step. Moreover, the performance of the proposed system is discussed.

2. The Preprocessing

The purpose in the preprocessing phase is not only to make face images standard, but also to ready them for feature extraction. The preprocessing steps are applied both to the training and testing images. In our system, 8-bit gray level still images were studied.

Firstly, potential noises in the images were eliminated. Face regions of all the images should be overlapped. To achieve this, in all of the images, coordinates of the midpoints between two eyes, were fetched to the same position. Thus, all images were aligned.

Sizes of sample faces may vary too. Images are resized by making the distance to fixed 40 pixels between the eyes.

In addition to this, a mask was applied to eliminate the outside of face region, from the image. The applied mask is an ellipse which is tangent to the points that are equally distant from the centers of the eyes. The outside region of the mask was set to zero.

Feature extraction is based on the gray level information of the image. Thus an object's two different images which were taken under different lighting conditions, could be interpreted as two different images and also two different objects. To overcome this problem, histogram equalization is applied to the images, resulting in uniform gray level distribution [4]. Figure 1 shows preprocessing steps applied to an image.

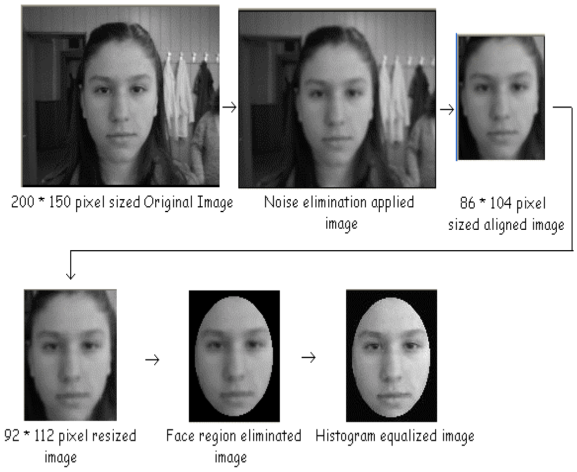


Figure 1: Preprocessing steps of proposed system

3. Feature Extraction

Humans usually identify faces looking at the whole face globally; instead of interpreting local characteristics such as length of the nose. Matthew and Turk [5] have applied PCA in face recognition for extracting the characteristic features of faces. In this study PCA and Kernel PCA methods were used for reducing dimensionality of original data while expressing most important characteristic features, and the performances of PCA and Kernel PCA were compared.

3.1 Principal Component Analysis

PCA finds basis vectors for a subspace, which:

- maximizes the variance retained in the projected data
- or (equivalently) gives uncorrelated projected distributions
- or (equivalently) minimizes the least square reconstruction error.

As in face identification, humans score facial beauty, looking at the whole face rather than measuring face proportion.

Each of the p number of images whose size is n*m is assumed to be [n*m]*1 sized column vectors that correspond to the original data.

$$\Omega = X * X^T \tag{1}$$

The covariance matrix (Ω) is defined by the (1) where X is the mean matrix subtracted vector whose size is [n*m]*p, where there are p image in training set. Because of the multidimensionality, (Ω) is calculated using (2) instead of (1).

$$\Omega = X^T * X \tag{2}$$

The eigenvalues and associated eigenvectors are computed as (3) where **Ω** is the covariance matrix**,** λ is the set of eigenvalues and V is the associated eigenvectors.

$$\Omega * V = \lambda * V \tag{3}$$

The eigenvectors are sorted in ascending order according to their eigenvalues and the first 60% of them are kept. Both the training and testing images are projected into the eigenspace by multiplying mean centered image with selected eigenvectors. These features in eigenspace are called as eigenfaces.

$$X' = V^T * X \tag{4}$$

3.2. Kernel Principal Component Analysis

Kernel PCA can be used for nonlinear feature extraction. Schölkopf, Smola, and Muller (1998) introduced Kernel PCA as a nonlinear generalization of PCA[6]. Kernel PCA is based on the simple idea of performing Principal Components Analysis in the feature space of a kernel [7]. In kernel PCA, the kernel trick is used for finding the principal components of the samples. The kernel trick is simply an efficient way of computing inner products in high-dimensional spaces. Kernel PCA utilizes the kernel trick to perform feature space operations using only dot-product between data points in the feature space. In simpler terms, Kernel PCA requires only the similarity between all pairs to perform nonlinear dimension reduction. Kernel PCA first maps the data into some feature space F via a (usually nonlinear) function and then performs linear PCA on the mapped data. This is performed through a mapping function Φ by the following equation where x is original input space (5).

$$\Phi: X \rightarrow F \tag{5}$$

Kernel PCA employs gaussian, polynomial or linear kernels instead of carrying out the mapping.

The kernel trick or with different words the replacement of all inner products in the algorithm with a kernel function is:

$$K(\underline{x}, \underline{y}) = (\Phi(\underline{x}) \,.\, \Phi(\underline{y})) = f(\underline{x}, \underline{y}) \tag{6}$$

The kernel can be any function which satisfy MERCER's theorem, thus it should be a semi positive definite, symmetric function. Two examples of a Mercer's kernel are illustrated in (7) and (8):

$$K(\underline{x}, \underline{y}) = \underline{x} \,.\, \underline{y} \tag{7}$$

Equation (7) is the first order polynomial kernel, and it can be called standard PCA.

$$K(\underline{x}, \underline{y}) = \exp(- \| \underline{x} - \underline{y} \|^2 / 2\sigma^2) \tag{8}$$

Equation (8) is the Gaussian Radial Basis Function (RBF) kernel. It is a mapping to infinite dimensions, that can be used successfully in dimension reducing and feature extracting. [8]

4. Training and Testing by SVM

In this study, learning facial beauty and testing the system are accomplished by using SVM. SVM is a learning, method introduced by V.Vapnik et al [9]. The SVM's objective can also be justified by structural risk minimization: the empirical risk (training error), plus a term related to the generalization ability of the classifier, is minimized. In other words, the SVM loss function is analogous to ridge regression.

In the learning phase, eigenfaces of beautiful and not beautiful classes are trained with SVM. Then the eigenfaces of the test images are classified using SVM.

Let m-dimensional inputs x_i (i=1..m) belongs to two classes $y \in \{-1,1\}$. The aim is finding the best decision boundary hyperplane f(x) that maximizes the sum of distances to the closest positive and negative training examples. This hyperplane intents to classify, not only train vectors but also the test samples successfully. SVM maximizes the class separability and classifier potential points with in same class. The hyperplane f(x) is defined by the following transformation where k is kernel function and b is bias value.

$$f(x) = \sum_{i=1}^{M} y_i \alpha_i \, k(x, x_i) + b \tag{9}$$

The sign of f(x) shows the class membership of x. x_i with nonzero α_i values are called support vectors of the hyperplane.

5. Experimental Results

To train and test the proposed system, 170 female face images were used. Some beautiful face images were taken from magazines. We shuffled those pictures and constructed two classes with respect to the facial beauty, by requesting the personal views of 50 persons of different cultural background, age and gender. 10 images upon which no general agreement about their beauty could be reached were left out and the system was trained by 90 images which were divided equally between two classes.

In Figure 2 and Figure 3 some beautiful and not beautiful faces from the train set are illustrated.

Figure 2: Beautiful Face Samples

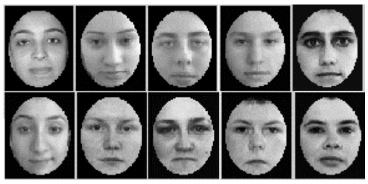


Figure 3: Not Beautiful Face Samples

The mean image of beautiful faces is shown in Figure 4a and the mean image of not beautiful faces, is shown in Figure 4b.

(a) (b)

Figure4: (a) Mean image of beautiful faces, (b) Mean image of not beautiful faces

Golden proportion leads to control the decision result of the system. Under this scheme mean image of beautiful faces and mean image of not beautiful faces are interpreted from the point of view of golden proportion. In Figure 5a, the golden proportion of the mean image of beautiful faces, in Figure 5b the golden proportion of the mean image of not beautiful faces are shown.

(a) (b)

Figure 5: (a) The Golden Proportion of mean image that belongs to beautiful faces , (b) The Golden Proportion of mean image that belongs to not beautiful faces

The abbreviations of golden rate proportion are shown in Table 1.

Table 1: Golden Rate Proportions and Abbreviations

Proportion	Abbreviation
Pupil to lip / Length of nose	I
Length of head/ Pupil to chin	II
Length of lips / Width of nose	III
Width of nose / Distance between nostril	IV
Distance between pupils/ Distance between eyebrow	V

Table 2: The estimation of beautiful and not beautiful faces mean image according to Golden Rate:

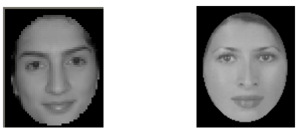
	Mean image of beautiful faces	Mean image of not beautiful faces
I	52/31=1.63	57/33=1.73
II	118/75=1.57	118/79=1.49
III	31/19=1.63	37/26=1.42
IV	19/11=1.58	26/13=2
V	40/24=1.67	40/16=2.5

As it is evident from Table 2, the mean image of beautiful faces' golden proportions are close to "fi" number 1.608.

Nevertheless the mean image of not beautiful faces' golden proportions are not close to "fi" number. Since the mask applied to face may cut off some of the face images, "Face height / Face width" ratio is taken away the expected ratio.

In test phase, 80 images were used which do not belong to the train set. Like the train set images; the test set images were labeled as beautiful or not beautiful according to majority of the people's opinions.

Figure 6a shows an example that is determined as beautiful by PCA which was judged as not beautiful by people and Kernel PCA. Also Figure 6b shows an example that is determined as not beautiful by PCA where as it was determined as beautiful by people and KPCA.



(a) (b)

Figure 6: (a)The image that PCA determines as beautiful, however people and Kernel PCA determines as not beautiful, (b)The image that PCA determines as not beautiful, however people and Kernel PCA determines it as beautiful

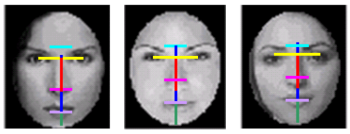
Table 3 shows persons and our system's decision results related to the test faces. Persons and PCA have given same decisions for 33 images as beautiful and 33 images as not beautiful. However persons and KPCA have given same decisions for 37 images as beautiful and 34 images as not beautiful.

Table 3: Decision results of facial beauty for PCA and Kernel PCA implementation

	PCA		KPCA	
	Beautiful Faces	not Beautiful Faces	Beautiful Faces	not Beautiful Faces
Persons	42	38	42	38
System	37	43	39	41
Agreement	33	33	37	34
Hit ratio	%78	%87	88%	89%

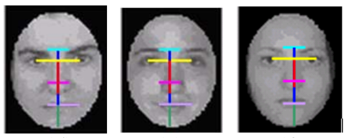
Kernel PCA can be used to discover nonlinear correlations in data that may otherwise can not be found using standard PCA. The information generated about a data set using Kernel PCA captures nonlinear features of the data. As a result KPCA outperfoms PCA in success ratio. We could also point out that the success ratio of not beautiful face determination for both of KPCA and PCA is better than success ratio of beautiful face determination.

Test images were evaluated according to both our system and golden proportion to compare and discuss the decision results. Figure 7 and Figure 8 show golden proportion of the six test examples. Also Table 4 gives the values of golden proportion metrics of these instances



(a) (b) (c)

Figure 7: The Golden Proportion of three samples that were determined as beautiful by the system according to both of PCA and KPCA results.



(d) (e) (f)

Figure8: The Golden Proportion of three samples that were determined as not beautiful by the system according to both of PCA and KPCA results

Table 4 shows the proportions of the face images. As is evidence, the images which were estimated as beautiful by the system were close to the golden proportion. On the contrary, the proportion of the face images which were estimated as not beautiful by the system were far away from golden proportion.

Table 4: The estimation of sample test images above (*Figure 7 and 8*) according to Golden Rate:

Proportion	a	b	c
I	51/31=1.64	51/31=1.64	45/28=1.61
II	118/73=1.61	118/74=1.59	118/74=1.59
III	31/19=1.63	30/20=1.50	26/16=1.62
IV	19/11=1.58	20/12=1.67	16/10=1.60
V	40/25=1.60	40/24=1.67	40/23=1.73
Proportion	**d**	**e**	**f**
I	49/33=1.48	51/30=1.70	46/25=1.85
II	118/77=1.53	118/80=1.48	118/75=1.57
II	33/21=1.57	33/21=1.57	29/19=1.53
IV	21/10=2.10	21/12=1.75	19/12=1.58
V	40/18=2.22	40/18=2.22	40/19=2.10

6. Conclusion

This paper presents an automated female facial beauty decision system which is inspired by the idea, that persons decide beauty by looking at the face as a whole. By setting up a system, which takes the principle components of the faces into consideration rather than performing measurements on the face images, it is made possible to achieve very close results to those of humans, in both classification and decision steps. In this paper, we used PCA and Kernel PCA to extract the principal components of facial images as the first step. Next, SVM was used to train and test the system. Although the proposed system doesn't measure the face proportions, the consistency in golden proportions of results at hand indicates that the recognition success is satisfactory. Consequently, this system could support systems like HCI (Human-Computer Interaction) by adding human beauty as a new property. Also the system could be used in plastic surgery operations to evaluate the look and beauty of reconstructed faces.

References

[1] M. Livio, (2002). *The Golden Ratio: The Story of Phi, The World's Most Astonishing Number*. New York: Broadway Books.

[2] Eisenthal, Y., Dror, G., Ruppin, E., 2004. *Learning Facial Attractiveness*. Israel Tel-Aviv University, School of Computer Science.

[3] H. Gunes, M.Y. Karsligil,"Measuring Female Facial Beauty by Calculating the Proportions of the Face", *Proc. of ISCIS XVII the Seventeenth International Symposium on Computer and Information Sciences,* Florida, USA, October 28-30, 2002, pp. 71-75.

[4] Gonzales, R. C., Wintz, P., *Digital Image Processing,* 1987.

[5] M. Turk and A. Pentland. Eigenfaces for recognition. J. Cog. Neurosci., 3(1), 1991.

[6] Schölkopf B., Mika S., Smola A.J., Rätsch G., and Müller K.-R.. Kernel PCA pattern reconstruction via approximate pre-images. Proceedings of the 8th International Conference on Artificial Neural Networks, Perspectives in Neural Computing, 1998.

[7] Schölkopf B., Smola A.J. and Müller K.-R.: Nonlinear Component Analysis as a Kernel Eigenvalue Problem,1998

[8] Tan J., Yang R., Kafatos M., Kernel Pca Analysıs For Remote Sensıng Data, 18th Conference on Climate Variability and Change, January 2006.

[9] Vapnik, V. N. An overview of statistical learning theory, *IEEE Transactions on Neural Networks*, 1999