

Comprehending and Transferring Facial Expressions Based on Statistical Shape and Texture Models

Pengcheng Xi¹, Won-Sook Lee¹, Gustavo Frederico¹, Chris Joslin², Lihong Zhou³

¹ School of Information Technology and Engineering, University of Ottawa, ON., Canada
{pxi086, wslee, gfred006}@uottawa.ca
<http://www.site.uottawa.ca/~wslee>

² School of Information Technology, Carleton University, Ottawa, ON., Canada
Chris_joslin@carleton.ca

³ Information Security Center, Southeast University, Nanjing, China
lihzhou@gmail.com

Abstract. We introduce an efficient approach for representing a human face using a limited number of images. This compact representation allows for meaningful manipulation of the face. Principal Components Analysis (PCA) utilized in our research makes possible the separation of facial features so as to build statistical shape and texture models. Thus changing the model parameters can create images with different expressions and poses. By presenting newly created faces for reviewers' marking in terms of intensities on masculinity, friendliness and attractiveness, we analyze relations between the parameters and intensities. With feature selections, we sort those parameters by their importance in deciding the three aforesaid aspects. Thus we are able to control the models and transform a new face image to be a naturally masculine, friendly or attractive one. In the PCA-based feature space, we can successfully transfer expressions from one subject onto a novel person's face.

1 Introduction

There have been many techniques developed for image compression by catching statistical information with encoders. Representative methods are fractal, wavelet, Differential Pulse Code Modulation and the most recent version of JPEG for still imagery [1]. However, almost all conventional methods captured the information in only one particular image. When we have a set of similar images, PCA should be an excellent technique in dimension reduction as well as in novel-image creation [2] [3]. After conducting a statistical analysis on the image dataset, the dimensions of the images are greatly reduced by mapping each image onto particular eigen-vectors with eigen-decomposition. Thus shape and texture models can be represented as a sum of average vectors and weighted variances along eigen-vectors. The weights, i.e., the parameters of the models, are normally referred to as "modes". Modifications on these modes help learn relations between the variances of parameter values and changes in facial image expressions.

It is of great interest to find the reason for those preferable faces. Do viewers' physical conditions affect their judgment? What kind of features can be observed from one's face, hatred, love, or happiness? Are there any relations between these features and facial attributes? Current solutions to these questions can be divided into two categories: geometry-based and psychology-based approaches.

The geometry-based methods are mainly concerned with geometric measurements, such as the distance between the eyes, the eye width and the context location of the nose. In the representative literature of [4], fractal geometry analysis is conducted on facial images and a female "attractive" face is created according to fractal rules. Artists such as Leonardo and Dürer also provide some criteria in this field [4]. Moreover, expression ratio images were proposed to map one person's expression details to a different subject's face. Given the feature point motions of an expression, this method requires an additional input of a different person's image with the same expression [5]. For situations where expression ratio images are unavailable, a geometry-driven facial expression synthesis approach based on feature point positions of an expression is more applicable [6].

Psychology-based approaches mainly argue the relations between attractiveness and symmetry, averageness as well as nonaverage sexually dimorphic features, i.e., the hormone markers [7]. They also analyze the relationship between female viewers' menstrual shift and their choice of male faces for dominance and short-term mates [8]. And certain researchers find that "shape normalized faces" and "texture normalized faces" are more attractive and younger than original faces [9].

In [10], a technique for defining facial prototypes supports transformations along quantifiable dimensions in "face space". Shape and color information are used to perform predictive gender and age transformations. Flexible models are also introduced for identification and coding of facial images [11]. In [12], linear object classes are learnt from a basis of 2D prototypical views and then used for 2D image synthesis.

Based on the above literatures, we try to combine geometry-based and psychology-based approaches as well as to imbibe ideas from [11] and [12] for our research. The basic idea is to utilize PCA as applied in [13] for building statistical shape and texture models with a facial image database. Utilizing these models, we can thus create more novel images with different expressions and poses. After some reviewers mark these images in terms of masculinity, friendliness and attractiveness, we can thus analyze the relations between parameter values and the facial expressions with certain feature selection scheme. This knowledge can thus help us transform a new face image to be more masculine, friendly or attractive. We can also introduce a neutral face image into the eigen-space and transfer different expressions among faces.

We organize our paper as follows. Section 2 is for building parameterized shape and texture models, where section 2.1 is to introduce the image database, section 2.2 for triangulation of facial images, 2.3 to present the PCA, 2.4 to build the models, and 2.5 is to interpret the effects on facial expressions after changing parameter values. In section 3.1, we present the survey results as well as feature selection scheme. In 3.2, we utilize the feature space to transfer expressions from one subject onto a novel face image. And we conclude our paper and mention some future work in section 4.

2 Parameterized Facial Models

In this section, we will introduce how to build statistical shape and texture models. We will also interpret the changes on facial expressions with parameter adjustments.

2.1 Database of Facial Images

We need parameterized facial models which can represent great differences in shape and texture vectors among images; therefore in our experiment, we take facial photos from 8 males of different ethnicities as shown in first row of Fig. 1. For each person, we take his photos with different expressions and poses. From all these images, we select 112 ones into our final database which can represent the maximum variances in facial expressions and poses.

All the face images are taken at the same lighting condition. Since the facial parts in the originally taken images are different in their scale, poses and locations, firstly we need to perform similarity transformation so as to get normalized images [14].

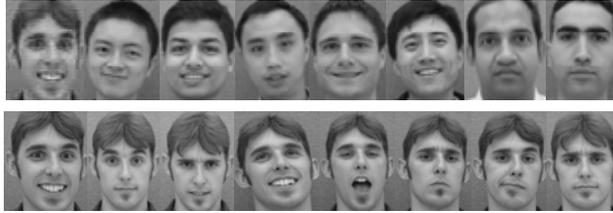


Fig. 1. Our database: different subjects with different expressions

2.2 Triangulation of Facial Images

To get texture vectors of faces, the first step is to get feature points of face contours. We define 44 control points for the human face, such as nose tip, eye and lip contours, middle chin, and so on. These points are the solid dots in Fig. 2. Besides these control points, we also define another 70 points among the solid dots to describe the detailed contours. These points well represent the main action units of expressions and are marked as hollow dots in Fig. 2 [15] [16].

Based on similarity transformations, we acquire the transformed coordinates of these 70 points in all the images of the database to reduce manual work of registering points. Therefore, for each image in the database, only 44 control points are located manually but locations of other 70 points can be decided automatically [17] [18].

With the 114 control points, we triangulate the average human face into 209 small triangles, as can be found out in Figure 2. In the eye, nose, and mouth portions, we create more triangles than we make in other parts so as to represent facial appearances in greater detail. The point selection is better than MPEG-4 in creating new faces from our experiments [19].

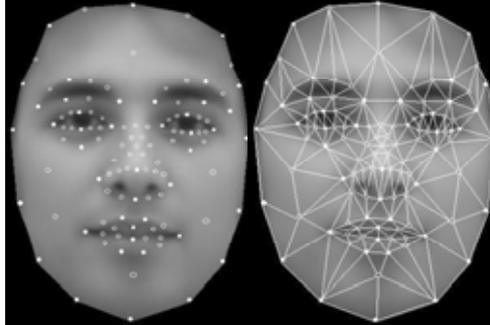


Fig. 2. Feature points on the average face as well as the triangulations

2.3 Principal Components Analysis

Principal components define a projection that catches the maximum amount of variation in a dataset and is orthogonal to the previous principle component [2].

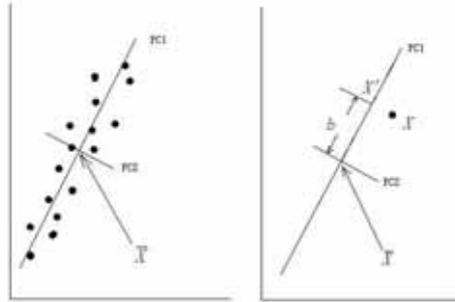


Fig. 3. Principal components analysis

After conducting eigen-decomposition in Fig. 3, we find two orthogonal vectors, i.e. principal component I (PC1) and principal component II (PC2). We observe that the main difference in the dataset is distributed along PC1, while less obvious differences lie on PC2. Therefore, data X can be approximated by $X' = \bar{X} + b \cdot PC_1$.

2.4 Parameterized Shape and Texture Models

After we get the shape and texture vectors for each training image, we build the parameterized shape and texture models in equation 1:

$$S = \bar{S} + \Phi_s b_s, \quad T = \bar{T} + \Phi_t b_t \quad (1)$$

where \bar{S} and \bar{T} is the average shape and texture vectors, Φ_s and Φ_t are eigenvectors for shape and texture models respectively. And b_s and b_t are shape and texture parameters [19]. 31 out of 228 principal components selected represent 99.23% variances in the original dataset. And we keep 93.95 of texture variance with the PCA.

Here we separately adjust shape and texture parameters to create new shapes and textures. Then we morph the new textures into new shapes with thin-plate splines [19].

2.5 Interpretation of Parameter Values

In this session, we interpret the effects of adjustments in shape and texture parameters. This analysis helps create images of different appearances for psychological analysis.

We increase the first element value in the b_s from equation (1) and keep other elements to be zero. Then we warp the average texture T to the new shape contour and acquire the smoothly transformed image sequence in first row of Fig. 4.

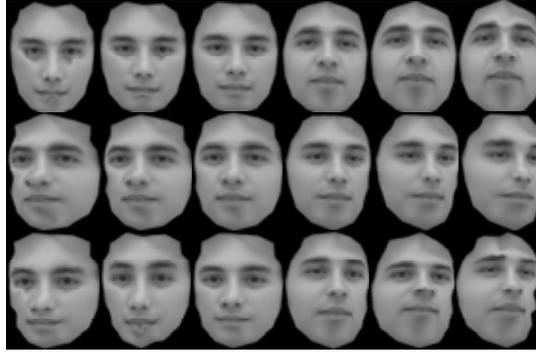


Fig. 4. Changes in expression and pose by changing parameters

The most obvious difference among images of first row in Fig. 4 is in the pose. With the changes of head pose, the proportion of upper face size to the lower size also differs. Another less significant change in this sequence is the mouth shape.

Second row of Fig. 4 shows the transformation of facial shapes by changing the second element in b_s . Though the viewing direction changes, the proportion in these faces keeps the same.

Changing $b_s(1)$ and $b_s(2)$ simultaneously can lead to transformations in third-row images of Fig. 4, where face poses and proportions are simultaneously modified because of the alternation in coordinates of both the outer and inner control points.

Therefore we find that pose transformations with changes in different modes equal to the sum of transformations in each mode transformation. Therefore, after we learn more relationships, we can combine them together so as to perform more complicated transformations on faces.

Now another series of experiments is to increase the value of elements in the texture parameter vector b_t from equation (1).

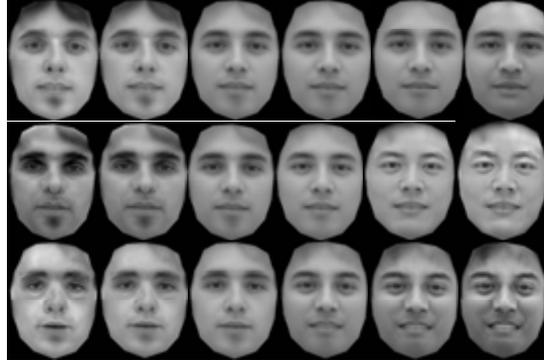


Fig. 5. Expression changes with the variance on first three texture parameters

In first row of Fig. 5, the main shift lies in the facial hair, the cheek, the beard as well as the mouth. Basically the general change is from a bonny boy with dark eyelid, thick eyebrow, black beard and small nose to a fleshy man with light eyelid, thin eyebrow, no beard and big nose. And a light smile keeps on every face in this sequence.

In second row of Fig. 5, generally the face becomes more Asian alike during the process. Also the mouth opens gradually during the changes of the parameter, which leaves an impression of light smiling.

In the third row of Fig. 5, main difference lies in the mouth area besides changes in eyebrow, skin color and extent of flesh cheeks.

Then how about changing the elements in shape and texture parameter vectors at the same time? Fig. 6 shows the effect of facial expression changes after we increase values of $b_s(I)$ and $b_t(I)$ simultaneously.



Fig. 6. Changes with the increment of first texture parameter

The alteration in Fig. 6 represents the combinatorial effects, i.e. the pose change and cheek variance.

These analyses can thus direct the creation of novel faces we need if only faces in our dataset are as different as possible. This is due to the fact that our models can represent most of the differences in shape and texture and can produce novel images which are outside of our dataset.

3. Manipulation on Novel Faces

In this part, we are addressing applications based on the information we learnt from the previous reconstructions and analysis.

3.1 Making Chubby and Skinny Faces

Manipulating faces to be chubby or skinny ones is widely applied in facial image processing as well as in face recognition.

Suppose our datasets are in a Gaussian distribution, then adding a new data into the space will not change the distribution. Thus we can replace the mean shape and texture vectors in equations (1) by those from a novel image. Then we perform transformations on our new facial images by changing the model parameters.

According to previous analysis, the 1st and 3rd texture parameters represent an effect of changing chubby faces. We also learn that modifications of the 14th and 17th shape modes play a significant role in getting skinny faces. By increasing these parameter values for a novel face, we achieve the transformed images as in Fig. 7.

As observed from Fig. 7, the transformed face (g) looks skinnier than (d), which is chubbier than (a), the previously neutral face.



Fig. 7. (a) Neutral face, (b ~ d) chubby and (e ~ g) skinny faces

3.2 Comprehension on Expression

The parameterized model introduced above brings us a good approach for creating facial images with different contours as well as expressions. These images are also a good source for image understanding. We discuss the relations between parameters and viewers' understanding on masculinity, friendliness and attractiveness.

3.2.1 Collection of Survey Results

We applied the parameterized shape and texture model to create 58 images in different expressions. Then we presented these pictures to 59 volunteers from 8 different ethnicities, half of which are females. They were asked to mark each image on masculinity, friendliness and attractiveness. After deleting those obviously improper results, finally we get 1004 data on each aspect.

The first step in processing these datasets is to standardize them so as to reduce personal differences in scales of scoring.

The next step, also the key process, is to find the sequence of those parameters' importance in deciding the created images on masculinity, friendliness and attractiveness. Here we use feature selection to solve this problem.

3.2.2 Feature Selection

Feature selection aims at picking out some original input features (i) for performance issues by facilitating data collection as well as compressing storage space and processing time, (ii) to perform semantics analysis for understanding the problem, and (iii) to improve prediction accuracy [20].

According to [21], [22], and [23], feature selection approaches can be divided into three categories: filters, wrappers and embedded approaches. The filter model relies on general characteristics of the training dataset to select some features without involving any learning algorithm [24]. The wrapper model requires one predetermined learning algorithm in feature selection and uses its performance to evaluate and determine which features are selected [24]. The embedded approaches simultaneously determine features and classifier during the training process [20].

Wrapper models tend to find features better suited to the predetermined learning algorithm resulting in superior learning performance, but it is also computationally expensive compared to the filter model [25].

Concerning the above analysis, we apply ReliefF, a filter model, to be the attribute evaluation approach for our dataset. ReliefF is an extension to Relief, which is to estimate the relevance of features according to how well their values distinguish between the instances of the same and different classes that are close to each other [24]. It randomly samples a number of instances from the training set and updates the relevance estimation of each feature based on the differences between the selected instance and the two nearest instances of the same and opposite classes [24].

For search method, we utilize Ranker to discover the final sequence of model parameters. Ranker, used in conjunction with attribute evaluators, attributes by their individual evaluations [26].

In our experiments, we use the WEKA software for feature selection [26]. This free software is a collection of machine learning algorithms for data mining tasks: data-processing, classification, clustering, association and data visualization. It provides us a complete scheme: ReliefF for attribute selection, and Ranker for searching. After processing our datasets with WEKA, we achieve the ranking sequence of model parameters in determining masculinity, friendliness and attractiveness as in Table 1, where s_i represents the i^{th} shape parameter, while t_i denotes the i^{th} texture parameter.

Table 1. Ranking sequences of model parameters in determining the three aspects

Mascu.	s_2	t_2	s_3	s_1	t_1	t_5	s_6	t_3	s_5	s_4	t_6	s_8	s_7	t_7	t_4
Friend.	t_3	s_1	s_2	t_5	t_6	s_4	t_2	s_3	t_1	s_6	t_4	s_5	t_7	s_8	s_7
Attrac.	s_2	t_3	s_4	s_6	t_2	t_5	s_5	t_6	s_1	s_7	t_1	s_8	t_7	t_4	s_3

3.2.3 Face Transformation Results

By changing the first 15 shape (s_i in Table 1) and texture parameter values (t_i in Table 1) according to the feature selection result in Table 1, we create the transformed images based on a neutral boy's face shown in Fig. 8,

After presenting these four images to our volunteers again, 92.1% of them agreed on masculinity in the second face, 89.3% volunteers believed the third image is friendlier than the first one, and 83.6% viewers considered the last one to be more attractive.



Fig. 8. Neutral, masculine, friendly and attractive faces

3.2.4 Transferring Facial Expressions

We notice that a feature space has been constructed after we perform PCA on our dataset. To make use of this space, we also attempt to transfer expression changes from one person's faces onto another subject. This can be explained by Fig. 9.

N-Dimensional Feature Space

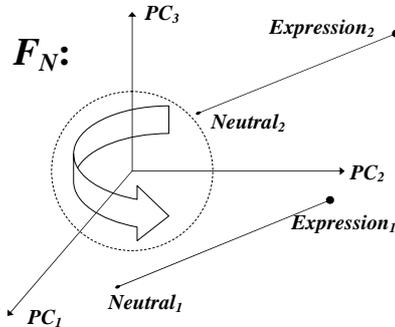


Fig. 9. Illustration of expression transfers in feature space

In Fig. 9, we simplify the explanation on N-Dimensional feature space into that in a 3-D space, where PC_1 , PC_2 and PC_3 denote the first three eigen-vectors of our dataset. "Neutral₁" represents the reference person's neutral face, and the "Expression₁" denotes the person's face with certain expression. Then the directed vector from "Neutral₁" to "Expression₁" can denote the facial expression changes.

Mappings from the directed difference vector onto each eigen-vector correspond to various weights on each feature vector. Then we can contribute these weights to

“Neutral₂”, which denotes a target neutral face. The contribution creates the final “Expression₂” as shown in Fig. 9.

Figure 10 provides one good example for this expression transfer. Besides our previously collected facial image database, we asked another volunteer to make up angry, disgust, feared, surprised and sad faces. Then we perform the calculation on contour vectors in the feature space. The novel neutral face in Figure 10 is out of our database. After performing the contribution, we reconstruct the transformed faces by warping previous face patch to the transformed shape contours.

The whole procedure is automatic except for the manual work on locating contour points. We find that better expressions can be captured by controlling the transformations on different facial parts. For example, we can increase the weights on mouth area to make the transformed face in a more surprised way.



Fig. 10. Neutral, angry, disgust, feared, surprised and sad faces

4. Conclusions and Future Work

The main role of this paper is to disclose the relations between model parameters and subjective understanding on facial images concerning masculinity, friendliness and attractiveness. Applications include transforming a novel face image to be more masculine, friendly or attractive as well as transferring expressions from one subject’s face onto another one’s.

To avoid the disadvantages of conventional approaches, we do not rely on the geometric measurements but conduct much analysis on the shape and texture models’ parameters. Specifically, we apply the face-space to leap over the complicated and normally inaccurate geometric measurements of distances and ratios on human faces.

The PCA-based model is an excellent source for creating images of many differences in their contour and texture vectors. It is also robust in image compression.

We separate shape and texture models so that the two models can create more faces. This should be an important improvement to previous approaches [14].

We received a large dataset from a survey and well analyzed the relations between parameter values and subjective understanding on the created faces. This knowledge helps us create a masculine face from the original neutral face image. Similar approaches also work for creating friendly and attractive faces out of a neutral one.

Based on the feature space, we also successfully transfer facial expressions from one subject’s face onto another face. This is not a simple adding or subtraction as it seems to be, but an important manipulation in feature space.

Although Table 1 provides the ranking of model parameters, it is not easy to decide the exact values for them. We are going to try another approach called expressional

image synthesis controlled by emotional parameters [28]. Borrowing the idea from expression transfer, we can also perform transfer on masculinity, friendliness and attractiveness if we can have such reference faces. Moreover, since mouth in the neutral face is always closed, we need to add teeth to transformed faces when the transformation leads the mouth to be widely opened.

Acknowledgement

Many thanks to colleagues at the DISCOVER laboratory of University of Ottawa for their devotion of facial images to our research. And we would like to express our gratitude to those volunteers in University of Ottawa (Canada) and Southeast University (China) who helped us to finish the survey. Also we would like to thank anonymous reviewers for their invaluable comments.

References

1. Marvel, L. M., Hartwig Jr, G. W.: A Survey of Image Compression Techniques and their Performance in Noisy Environments. Final report number: A409623, (May 1996–Jan. 1997) pages: 98
2. Smith, L. I.: A tutorial on Principal Components Analysis, maintained by Cornell University, U.S.A. (Feb. 26, 2002)
3. Yeung, K. Y., Ruzzo, W.: Principal Component Analysis for Clustering Gene Expression Data. *Bioinformatics* 17(9): (2001)763-774
4. Schmidhuber, J.: Facial Beauty and Fractal Geometry, Technical report. IDSIA-28-98, IDSIA, Corso Elvezia 36, 6900 Lugano, Switzerland, 1998
5. Liu, Z., Shan, Y., and Zhang, Z.: Expressive Expression Mapping with Ratio Images. *Computer Graphics, Siggraph*, (August 2001) 271-276
6. Zhang, Q., Liu, Z., Guo, B., Terzopoulos, D., Shum, H.: Geometry-Driven Photo-realistic Facial Expression Synthesis. *IEEE Trans. on Visual. and Comp. Graph.*, Volume 12, Issue 1, (2006) 48-60
7. Fink, B., Penton-Voak, I.: Evolutionary Psychology of Facial Attractiveness, *Current directions in psychological sciences*, Vol. 11, Num. 5, Blackwell Publishing, (2002) 154-158
8. Johnston, V. S., Hagel, R., Franklin, M., Fink, B., Grammer, K.: Male Facial Attractiveness Evidence for Hormone-Meditated Adaptive Design, *Evolution and human behavior* 22, Elsevier, (2001) 251-267
9. O’Toole, A. J., Price, T., Vetter, T., Bartlett, J. C., Blanz, V.: Three-Dimensional Shape and Two-Dimensional Surface Textures of Human Faces: The Role of “Averages” in Attractiveness and Age, *Im. and Vis. Comput. Journ.* 18 (1999) 9-19
10. Rowland, D., Perrett, A. D. I.: Manipulating Facial Appearance through Shape and Color, *IEEE Computer Graphics and Applications*, 15(5), (1995) 70-76

11. Lanitis, A., Taylor, C. J., Cootes, T. F.: Automatic Interpretation and Coding of Face Images Using Flexible Models, *IEEE Trans. on PAMI*, Vol. 19, No. 7, (1997) 743-756
12. Vetter, T., Poggio, T.: Linear Object Classes and Image Synthesis from a Single Example Image, *IEEE Trans. on PAMI*, Vol. 19, No. 7, (1997) 733-742
13. Cootes, T. F., Edwards, G. J., Taylor, C. J.: Active Appearance Models, *IEEE Trans. on PAMI*, Vol. 23, No. 6, (2001) 681-685
14. Cootes, T. F., Taylor, C. J.: Statistical Models of Appearance for Computer Vision, Technical report, University of Manchester, UK, (1999)
15. Tian, Y.L., Kanade, T., Cohn, J. F.: Recognizing Action Units for Facial Expression Analysis, *IEEE Trans. on PAMI*, Vol. 23, No. 2, (2001) 97-115
16. Ekman, P., Friesen, W. V.: Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto, CA, (1978)
17. Lee, W., Magnenat-Thalmann, N.: Head Modeling from Pictures and Morphing in 3D with Image Metamorphosis Based on Triangulation, *Proc. Captech98*, Geneva, (1998) 254-267
18. Goto, T., Lee, W., Magnenat-Thalmann, N.: Facial Feature Extraction for Quick 3D Face Modeling, *Signal processing: Image communication*, Elsevier Science, Vol. 17, Issue 3, (2002) 243-259
19. Xi, P., Xu, T., Zhao, Zh.: Knowledge-Based Active Appearance Model Applied in Medical Image Localization, *Proc. of IEEE International conference on Mechatronics and automation (ICMA'05)*, Niagara Fall, Canada, (2005) 637-642
20. Neumann, J., Schnorr, C., Steidl, G.: Combined SVM-Based Feature Selection and Classification, *Machine Learning*, 61, (2005) 129-150
21. Guyon, I., Elisseeff, A.: An Introduction to Variable and Feature Selection, *Journal of Machine Learning Research*, (2003) 1157-1182
22. John, G., Kohavi, R., Pfleger, K.: Irrelevant Features and the Subset Selection Problem, *Proc. of the 11th International Conference on Machine Learning*, (1994) 121-129
23. Bradley, P. S., Mangasarian, O. L.: Feature Selection via Concave Minimization and Support Vector Machines, *Proc. of the 15th International Conference on Machine Learning*, San Francisco, CA, USA, (1998) 82-90
24. Yu, L., Liu, H.: Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution. In *Proceedings of the Twentieth International Conference on Machine Learning (ICML-03)*, (2003) 856-863
25. Langley, P.: Selection of Relevant Features in Machine Learning, *Proc. of the AAAI Fall Symposium on Relevance*, AAAI press, (1994)
26. Witten, H.I., Frank, E.: *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd edition, Morgan Kaufmann, San Francisco
27. Penrose, R., A Generalized Inverse for Matrices, *Proc. Cambridge Phil. Soc.* 51, (1955) 406-413
28. Zhou, C., Lin, X.: Facial Expressional Image Synthesis Controlled by Emotional Parameters. *Pattern Recognition Letters* 26, (2005) 2611-2627