 The Open Automation and Control Systems Journal, 2015, *7*, 1796-1801

The Research on the Multiple Kernel Learning-Based Face Recognition in Pattern Matching

Shan Feng*

School of Computer, Hubei Polytechnic University, Huangshi, China

Abstract: The paper analyzes the multiple kernel learning-based face recognition method in pattern matching area. Based on the analysis of the basic theory of multiple kernel SVM, this thesis focuses on the multiple kernel SVM algorithm based on semi-infinite linear program (SILP), including SILP based on column generation (CG) and SILP based on chunking algorithm (CA). The two SILP improved algorithms are applied to several classification problems, including UCI binary classification problem datasets and multi-classification problem datasets. Furthermore, the two SILP improved algorithms are applied to the actual problems of face recognition. The experiment data shows that with the multiple kernel learning-based method, the performance of face recognition can be obviously improved.

Keywords: Face recognition, Kernel Functions, Multiple kernel Learning-based, Pattern Recognition, Semi-infinite linear program (SILP).

1. INTRODUCTION

Automatic Face Recognition (AFR) is challenging in image processing and analyzing [1]. By AFR, we mean that people attempt to endow computers with the ability to analyze the human face image, to extract the valid individual information, and to identify him/her. Such a type of theory and technology is not only greatly desired in the theoretical research but also has significant potential in applications [2, 3].

After the nearly ten years' research, AFR technology has already obtained considerable advances especially in the past several years. The existing AFR commercial systems can basically meet the application need under certain strict constraints. Thus, it is far from the true trend practical level in the non-ideal situation. Human face recognition is one of the most valuable biometric identification methods, which has drawn a lot of attention for its immense potentials in many applications and has been the hot topic over the past decades. Meanwhile, it has got extensive attention in the development and construction of society and attracted a large number of researchers and institutions engaged in the related research. Although face recognition has already made considerable achievements, many researchers who have developed numerous algorithms for face recognition and they have become more and more mature, still faced lots of challenges [4]. The paper presents the studies of face verification and face recognition in terms of application.

To avoid the disadvantages of the client specific linear discriminant analysis (CSLDA) and kernel method CSKDA, the client specific multiple kernel learning (CSMKL) method is studied in this paper [5]. The theory of CSMKL is proposed based on the combination of multiple kernels.

In solving complex visual learning tasks, multiple kernel learning has been a feasible way for improving performance, which can characterize the data more precisely and then provide more robust discriminant information for classification. While CSKDA does not consider the effect of different kernels in feature extraction, to solve the learning and client specific subspace, which can fully make use of the advantages of MKL and client specific subspace, firstly, we need the framework of MKL to reduce the dimensionality and obtain more accurate image features. Then we obtain client specific subspace in MKL subspace to describe the diversity of different faces. The experimental results obtained on the internationally recognized facial database XM2VTS using the Lausanne protocol show the effectiveness of the proposed client specific multiple kernel learning method.

A novel scheme for feature extraction for face recognition by fusing local and global contextual constraints based on linear discriminant analysis (LGCCLDA) is studied. The facial changes due to variation of pose, illumination, expression, *etc*., often appeared only in some regions of the whole face image. Therefore, global features extracted from the whole image fail to cope with these variations. To address these problems, our method first divides the original images into modular sub-images and then CCLDA is utilized to each of these sub-images as well as to the whole image to extract local and global discriminant features respectively [6]. Moreover, CCLDA can take the contextual constraints in images into account, which can provide useful information for classification. Experimental results obtained on various

*Address correspondence to the author at the Hubei Polytechnic University, 16th Guilin North Road, Huangshi Hubei, China; Tel: 13597628435; E-mail: 63824998@qq.com

1874-4443/15

face databases show that our method can not only improve the performance in recognition, but also get desired results in less training samples.

In the paper, we introduced an improved notion of subspace learning from image gradient orientations for face recognition. A fundamental problem of the majority of traditional subspace learning techniques (both linear and nonlinear) for face recognition is that they are not robust when the images are affected by illumination, noise and so on. Most methods are usually based on linear correlation of pixel intensities which fails very often to model reliably visual similarities. Research shows that replacing pixels intensities with gradient orientations, to some extent, a remedy to this problem [7]. To improve the robust and face recognition rate of the method, we adopt image gradient orientations in combination with subspace methods and apply it to face recognition. Meanwhile, to prove that the proposed method is more robust and can significantly outperform some traditional subspace methods, lots of experiments have been carried out on different databases and the results show the effectiveness of our method in solving the illumination of images.

Kernel learning method is widely used in the field of pattern recognition. There are some disadvantages in traditional kernel methods, such as poor generalization ability, long iteration time, and poor robustness. Multiple kernel support vector machine (SVM) as an extension and development of single kernel SVM has a better generalization capability and better robustness [8]. Based on the analysis of the basic theory of multiple kernel SVM, this thesis focuses on the multiple kernel SVM algorithm based on semi-infinite linear program (SILP), including SILP based on column generation (CG) and SILP based on chunking algorithm (CA). The two SILP improved algorithms are applied to several classification problems, including UCI binary classification problem datasets and multi-classification problem datasets. Furthermore, the two SILP improved algorithms are applied to the actual classification problems, including handwritten recognition problems such as the United States postal service CUSPS) dataset and the Modified National Institute of Standards and Technology (MNIST) dataset, and biomedical problems such as electroencephalogram (ECG) problem [9] and Electroencephalogram (EEG) problem. The results show that two algorithms are feasible and effective.

2. THE FRAMEWORK OF FACE RECOGNITION

Human has an excellent appetence on face recognition, it's our dream to make a machine that has the same intelligent recognition ability. Original dream and curiosity drive people conduct the continuing research on automatic face recognition. As the development of the modern information technologies, automatic face recognition has attached importance to broad fields such as military, commercial, security, in virtue of its good applicability and a non-intrusive property Face recognition has become one of the most representative and challenging research content of pattern recognition domain. Face recognition is an old but young academic problem, about which people have thought across three centuries [10].

As early as 1888, Nature magazine published the first academic paper on face recognition, starting people's exploration on face recognition. 120 years later, *i.e.* 2008, Science magazine published a technical comment on "100% automatic face recognition accuracy, which pointed out that the highly accurate face recognition in the real world application, is still an ambitious goal and its solution requires me to continue the innovative research.

Face recognition is a special complex pattern recognition problem, whose particularities lie in its high feature dimensionality versus small sample size, the large within-class variations versus the small between-class variations, and so on. These particularities make the mature pattern classification theory that cannot be applied to solve face recognition problem, which qualify the feature extraction as the determinant of the accuracy level. Unfortunately, traditional feature extraction methods often start from a specific object function, which makes the algorithm adaptable to the specific variations contained in a certain data set, and thus achieve high recognition accuracy. However, they cannot solve the real world recognition problem in complex conditions.

Feature extraction is a key problem in the study of face recognition. Face images obtained by image acquisition equipment construct the high-dimensional face observation space, and discriminative features of human face that lie on the low manifold (subspace) formed by face samples. The feature extraction is the process of searching the low-dimensional discriminative subspace (manifold) from the high-dimensional observation space. During this process, illumination processing and dimensionality reduction are the two crucial steps. The distribution of face sample data in the observation space is essentially sparse, while that of each class is highly overlapping with each other, resulting in that the differences of face samples within classes are larger than that between classes. Processing the illumination of face images effectively can adjust the distribution form of face samples in the observation space, reducing the differences within classes, increasing those between classes, and thus laying a good foundation for the followed dimensionality reduction (feature extraction). However, face images, after handling lighting, still lie in the observation space, and the high-dimensional property of vectors will incur Curse of Dimensionality and measure concentration, in which case it is hard to achieve good results by using classifiers to classify the samples directly. Therefore, it becomes the central task in the study of face recognition how to design effective algorithms of reducing dimensionality to extract the low-dimensional face features from the high-dimensional observation space with the most powerful discriminative.

The kernels of Gabor wavelet are similar to the response of the two-dimensional receptive field profiles of the mammalian simple cortical cell, and exhibit the desirable characteristics of capturing salient visual properties such as spatial localization, orientation selectivity, and spatial frequency selectivity. Therefore, the Gabor features of images are insensitive to the variations of illumination, pose and expression. Based on this property of Gabor wavelet, we first made a Gabor transform to face images, with the facial Gabor features obtained; and then proposed a locality preserving based supervised manifold learning algorithm designed for the high-dimensional Gabor features, called Supervised Neighborhood Preserving Embedding, by which the dimensionality is reduced effectively. Experimental results on the two face databases of Yale and ORL show the feasibility of the proposed algorithm.

Similar to the Gabor wavelet, Non-subsampled Contourlet Transform (NSCT) is also a multi-scale and multi-direction 2D wavelet transform. The difference from the Gabor wavelet is that the basic functions of NSCT are orthogonal to each other, with the least redundant information contained by the sub-bands of NSCT. In consideration of this good feature of NSCT, we transformed, based on the Retina illumination model, the face image into NSCT domain, used adaptive threshold to filter each high-frequency sub-band, and obtained the illumination invariant of face images by using inverse NSCT. Experimental results indicated that the proposed algorithm can very effectively reduce the effect of illumination variation on face images, improving the recognition rate of the algorithm to a large extent.

Manifold learning has a core idea of preserving the local geometric properties of the data space during the process of reducing the dimensionality, which is now a mainstream nonlinear subspace method and wildly applied to face recognition.

There are three traditional algorithms for face recognition.

1) Adaptive Supervised Locality Optimal Preserving Projection (ASLOPP). After investigating some drawbacks of existing supervised LPP algorithms in classification problem, ASLOPP method determines the neighborhood size of each data in sample space by using adaptive neighborhood algorithm, adds constraint conditions, and employs iterative method to optimize the objective function, which leads to orthogonal basis vectors in low-dimensional embedding, decreasing the information redundancy in the embedding while increasing the discriminative power. We conducted experiments on Extended Yale B and CMU PIE face databases, with good results improving the face recognition rates obviously.

2) Supervised Locality Preserving Projection (SLPP). This algorithm is also based on LPP, which first employs adaptive neighborhood method for determining the sample neighborhood size in the process of constructing the Eigen map of sample space, in terms of the distribution of facial feature space. Then the prior class label information of samples is used to construct within-class and between-class maps, respectively, by which the distribution of each class is represented. And the objective function absorbs the idea of LDA, which allows the optimized embedding not only maintain the local geometry of the original sample space, but also minimize the within-class variance while maximize the between-class one, greatly enhancing the discriminative of the embedding.

3) Supervised neighborhood Preserving Embedding. On the basis of the Neighborhood Preserving Embedding (NPE), this algorithm determines the link mode between samples by the class label in constructing the neighborhood map of the sample space. By optimization of the objective function, the embedding holds the local linearity of the sample space optimally and also greatly improves the discriminative power.

3. IMPROVEMENT OF KERNELS FUNCTION

Face Tracking is an essential stage in those applications, such as video-based human face recognition, video-based monitoring. The Camshift algorithm has strong robustness regarding the object tracking, but it has the drawbacks such as Tracking-window must be initialized manually, and the robustness to those skin color-like regions is unsatisfactory. We propose a real-time face tracking algorithm through making improvements to Camshift, named AM-Camshift, which can initialize the track window and implement the multi-objects tracking automatically, and enhance the robustness to skin color-like area.

The linear discriminant analysis (LDA), especially the Fisher linear discriminant analysis (F-LDA), which can distinguish different face classes through maximizing the ratio between S_B, n_v and S_{wrf}, is another significant research result after the Principal Component Analysis (PCA) in the AFR research history. But because of the Small Sample Size, problem will become singular and cannot be solved. In the fourth chapter, this paper proposes the Adaptive Linear Discriminant Analysis (ALDA) algorithm through adjusting the Fisher criterion and making the improvement to the Fredman thought. Using the complement space (B) of between-class scatter matrix, the algorithm avoids the inverse operation of within-class scatter matrix (SW) and adaptively changes the parameter according to the sample information of each class. The experimental result shows that the A-LDA algorithm can resolve the SSS problem of FR effectively.

Accurate Face Key Feature Localization is essential to effective face representation. For instance, eyes, nose and mouth are the most important feature areas of a human face. The eyes' positions and the relative distance to each other remain a constant overwhelmingly for the majority people. So the eye's position plays a very important role in the face-normalization, moreover it is advantageous of further locating the other landmarks. Therefore, Eye Detection is the key step of the AFR system. In the fifth chapter, we propose a novel Eye Detection algorithm based on the eye's Hoar-like features, the RSVM Cascaded-Classifier and the pupil location.

Efficient face representation is the important premise of a highly effective FR core-algorithm. With the fine spatial local feature extraction and the orientation selection, the Gabor wavelet has strong robustness to the image illumination and geometry change, which makes it a good face representation method. However, the excessively high dimension of the Gabor-based feature vector is the main bottleneck in FR application, therefore we proposed a new Gabor feature extraction algorithm based on the key feature region of the face and another Gabor-feature extraction algorithm based on the valid face area. The former extracts the Gabor-feature sub-vectors of those areas, such as eyes, nose and mouth as well, which can fully represent the face individual characteristics, and combines these sub-vectors into a whole Gabor-based feature vector. The latter extracts the Gabor-based features of those valid face areas which is masked by a mask-template, and discards those areas outside the mask area which do not have any contributions to recognition task. Both above algorithms reduce the time and spatial complexity of the following FR core-algorithm to a great extent. The experimental results show that both methods remain the same robustness as the traditional methods while reducing the face feature vector dimension.

If these two kernels are combined linearly, the new mixed kernel function is:

$$K = \alpha K_1 + (1-\alpha)K_2 \tag{1}$$

K_1 is the Gaussian kernel function, K_2 is the Polynomial kernel function. If $\alpha = 1$, K is the Gaussian kernel function. If $\alpha = 0$, it is Polynomial kernel function.

The sample collection is

$$\{x_i, y_i\}_{i=1}^{N} \tag{2}$$

Establish the regression model to fit sample collection:

$$y(x) = \hat{y}(x) + e(x) = \sum_{i=1}^{M} \omega_i g_i(x) + e(x) \tag{3}$$

x is the variable of m-dimension. $y(x)$ is the actual output. $y(x)$ is the approximate output of regression model. $e(x)$ is the model error. ω_1 is the weighting factor of sub-regression. M is the several of regression models. $g_i(x)$ is the return item.

$$g_i(x) = \alpha_i k_{1i}(x) + (1-\alpha)k_{2i}(x), \alpha \in (0,1) \tag{4}$$

α is the right combination weighting factor of mixed kernel function.

$$\vec{y} = [y_1, y_2, y_3, \ldots, y_N]^T \tag{5}$$

$$\vec{e} = [e_1, e_2, e_3, \ldots, e_N]^T \tag{6}$$

$$\vec{\omega} = [\omega_1, \omega_2, \omega_3, \ldots, \omega_N]^T \tag{7}$$

$$\vec{G} = [g_1, g_2, g_3, \ldots, g_N]^T \tag{8}$$

$$g_k = [g_k(x_1), g_k(x_2), g_k(x_3), \ldots\ldots, g_k(x_N)]^T \tag{9}$$

Based on these functions,

$$\vec{y} = \vec{G}\vec{\omega} + \vec{e} \tag{10}$$

$$\vec{G} = \vec{P}\vec{A} \tag{11}$$

$$\vec{y} = \vec{P}\vec{\theta} + \vec{e} \tag{12}$$

$$\vec{\theta} = [\theta_1, \theta_2, \theta_3, \ldots, \theta_N]^T \tag{13}$$

$$\vec{\theta} = \vec{A}\vec{\omega} \tag{14}$$

The MSE is

$$J = \vec{e}^T\vec{e}/N = \vec{y}^T\vec{y}/N - \sum_{i=1}^{M} \vec{p_i}^T \vec{p_i}\theta_i^2 / N \tag{15}$$

So

$$J_k = J_{k-1} - \frac{1}{N}\vec{p_k}^T \vec{p_k}\theta_k^2 \tag{16}$$

After k times iteration, $J_k < \xi$. ξ is the selected threshold. Then a regression of mixed kernels function is obtained.

The direction of the coefficient is calculated using the following formula:

$$D(X) = D + \frac{\partial D^T}{\partial X}X + \frac{1}{2}X^T \frac{\partial^2 D^T}{\partial X^2}X \tag{17}$$

$$m(x,y) = \sqrt{\begin{array}{l}(L(x+1),y) - L(x-1,y))^2 \\ +(L(x,y+1) - L(x,y-1))^2\end{array}} \tag{18}$$

$$\theta(x,y) = \tan^{-1}\frac{L(x,y+1) - L(x,y-1)}{(L(x+1),y) - L(x-1,y)} \tag{19}$$

4. EXPERIMENT RESULTS

In this section, the algorithms are evaluated with the ORL face database and Yale face database. The ORL contains 400 images of 40 individuals. Some images were captured at different times and have different variations including expression (open or closed eyes, smiling or non-smiling) and facial details (glasses or no glasses). The image size of ORL image is 112 x 92; we subsample the image down to a size of 28 x 24 without any other pre-processing. We randomly took 5 images from each class as the training data, and left the rest 5 images as the probe.

The Yale face database contains 165 grey scale images of 15 individuals. The images demonstrate variations in lighting condition (left-light, center-light, right-light), facial expression (normal, happy, sad, sleepy, surprised, and wink), and with/without glasses. We also resized images to a size of 25 x 25 without any other pre-processing. A random subset with 6 images per individual was taken with labels to form the training set. The rest of the database was considered to be the testing set.

The nearest neighbor classifier was employed for classification of two databases. Such test is run 10 times and we took the average of the recognition rates for comparison. In addition to our proposed method, we also tested the UDP and KUDP methods. The popular Gaussian kernel is involved. Average recognition rate for each algorithm is reported in Figs. (**1**) and (**2**). From these results we found that the CKUDP gives better performance.

The experiment is carried out with classic artificial data. Because of strong signal shock of artificial data, it is used to check the reliability of thinning model.

Table **1** shows Yale face recognition with the new mixed kernel function of Gaussian and polynomial kernel function and Gaussian kernel function.

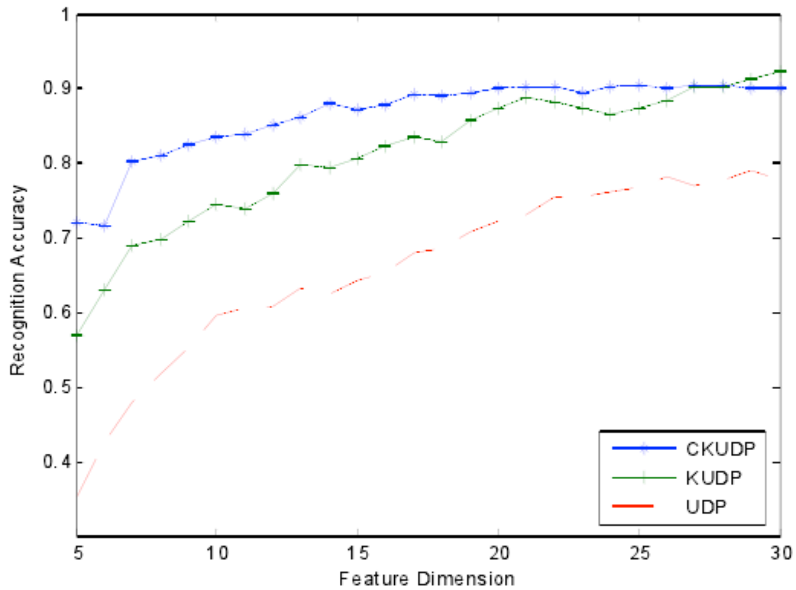

Fig. (1). Recognition rates of CKUDP, KUDP, and UDP.

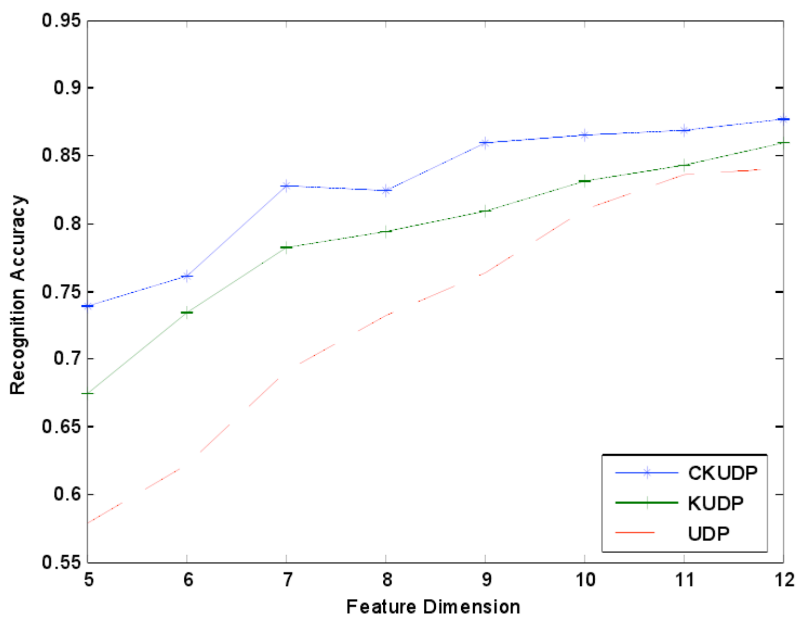

Fig. (2). Recognition rates of CKUDP, KUDP, and UDP versus the dimension on Yale database.

In the new method, data are processed directly in the feature space F avoiding the loss of significant discriminant information due to the PCA preprocessing step. The effectiveness of the proposed method has been demonstrated through experimentation using the popular face database ORL and Yale.

CONCLUSION

The normalization processing of the face recognition in this paper mainly includes the size normalization on the recognized face images. Different people may have different widths and lengths of faces. In order to enable to correctly recognize different faces after carrying out features extractions, the number of singular value vectors must be consistent with the number of singular values of face images in the database; therefore it must process the size normalization for the detected face images, which also facilitates the machine operation of feature values. In this paper, it studies two critical technical links of face detection and face recognition in the recognition of face, and conducts analyses and comparisons on some commonly used face detection and recog-

Table 1. Yale face recognition with the new mixed Kernel Function of Gaussian and Polynomial Kernel Function and Gaussian Kernel Function.

	The Mixed Kernel Function of Gaussian and Polynomial Kernel Function					**Gaussian Kernel**			
K	μ_K	α_K	θ_K	σ_K	**MSE**	μ_K	σ_K	θ_K	**MSE**
1	-0.0667	1.7524	1.9867	0.8439	0.038	0.095	1.4008	1.0494	0.0237
2	-1.76E-04	0.0746	0.7134	-0.0676	0.0332	-4.7673	1.0517	-0.2366	0.0201
3	0.8867	0.6558	0.0337	0.0022	0.0327	4.4735	0.6216	-0.1999	0.0164
4	-3.9484	0.669	0.5761	0.226	0.0282	7.451	0.9633	0.1676	0.0148
5	--	--	--	--	--	0.1248	0.7153	-0.0681	0.0144
6	--	--	--	--	--	-7.8283	0.6534	-0.8148	0.014
7	--	--	--	--	--	0.1888	0.1695	-0.0013	0.0136
8	--	--	--	--	--	1.1826	-0.0164	-0.0218	0.0131
	Training Error:0.0281					Training Error:0.0131			
	Test Error:0.0278					Training Error:0.0245			
	Run Time:3.326s					Training Error:5.057s			

nition algorithms. Through the collation of this information and the combination of related knowledge in the digital image processing, it proposes a face recognition algorithm based on the singular value decomposition. After preprocessing face images, by the use of the projection method, it obtains positions of five sense organs, and then extracts local feature values near five sense organs as the main features of faces by using the singular value decomposition. For different face images of the same individual, the matching degree of feature values will be very high. Seen from the experimental results, among five sense organs of the human face, eyes play a very important role in the determination of identities. In this experiment it adopts to look for the minimum similarity, so there is no situation of refusing to recognize. Compared with the traditional flexible template method, the speed has been greatly improved. Experimental results show that this program based on the reduction of computations enables to well-preserve features of human faces, thereby achieving a better recognition effect.

CONFLICT OF INTEREST

The author confirms that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] L. Liang, and H. Ai, "Face detections based on multi-template matching" *Journal of China's Image Graphics*, vol. 44, no. 10, pp. 623-630, 2004.

[2] M. Chen, "*Studies on the Face Image Detection and Classification System",* PhD thesis, Computer Science and Engineering Department of Shanghai Jiaotong University, Shanghai 2003.

[3] C. Lu, "*Study on Several Problems of Automatic Face Recognition and the System Implementation*", PhD thesis, Computer Academy of Sciences of Tsinghua University, Beijing 1998.

[4] G. Su, "Summary on face recognition technology", *Journal of China's Image Graphics*, vol. 5, no. 11, pp. 220-238, 2000.

[5] J. Li, "Research progress of new face recognition technology" *Journal of Computers*, vol. 31, no. 10, pp. 293- 295, 2004.

[6] L. Yuan, Z. C. Mu, Y. Zhang and K. Liu, "Ear recognition using improved non-negative matrix factorization", *International Conference on Pattern Recognition,* vol. 4, pp. 501-504, 2006.

[7] P. Motlicek, and F. ValenteValente, "Application of out-of-language detection to spoken term detection", In: *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing,* Dallas, USA, IEEE, pp. 5098-5101, 2010.

[8] J. Mantyjarvi, J. Koivumaki and P. Vuori, "Keystroke recognition for virtual keyboard", *IEEE International Conference on Multimedia and Expo*, Lausanne, Switzerland, pp. 429-432, 2002.

[9] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Hawaii, USA, pp. 511-518, 2001.

[10] T. Zhang, B. Fang, Y.Y. Tang, G. He, and J. Wen, "Topology preserving non-negative matrix factorization for face recognition", *IEEE Transactions on Image Processing*, vol. 17, no. 4, pp. 574-584, 2008.

Received: June 16, 2015 Revised: August 23, 2015 Accepted: September 11, 2015