Gender Recognition via Face Area Analysis

Vladimir Khryashchev, *Member, IAENG*, Andrey Priorov, Lev Shmaglit and Maxim Golubev

Abstract—An automatic gender recognition algorithm based on machine learning methods is proposed. It consists of two stages: adaptive feature extraction and support vector machine classification. Both training technique of the proposed algorithm and experimental results acquired on a large image dataset are presented.

Index Terms—**face detection, machine learning, adaptive feature generation, automatic image analysis, gender classification, support vector machine**

I. INTRODUCTION

THIS research is devoted to the development of an automatic system capable to distinguish people's gender by analyzing their faces on digital images. Such systems can find application in different fields, such as robotics, human-computer interaction, demographic data collection, video surveillance, online audience measurement for digital signage networks and many others. In addition, gender recognition can be applied as a preprocessing step for face recognition. Machine learning techniques [1] used in gender classification have universal character which allows to apply the solutions and knowledge in gender recognition to any other image understanding [2] or object classification task.

There are comprehensive surveys written on face detection [3, 4], face recognition [5] and facial expression analysis [6]. Gender recognition has been studied less. The comparative analysis of lately proposed gender classification algorithms has been presented in paper [7]. The researches have proposed algorithms based on artificial neural networks [8], on a combination of Gabor wavelets and principal component analysis [9], on independent component analysis and linear discriminant analysis (LDA) [10, 11]. In paper [12] the authors utilize genetic algorithm for feature selection and support vector machine (SVM) [13] for classification. In paper [14] an algorithm based on local binary patterns in combination with AdaBoost [15] classifier was proposed. Experiments with radial basis function (RBF) networks and inductive decision trees are described in [16].

Vladimir Khryashchev, PhD, Image Processing Laboratory, P.G. Demidov Yaroslavl State University, 14 Sovetskaya, Yaroslavl, Russia, 150000 (phone: +74852797775; e-mail: vhr@yandex.ru).

Andrey Priorov, PhD, Image Processing Laboratory, P.G. Demidov Yaroslavl State Ouniversity, 14 Sovetskaya, Yaroslavl, Russia, 150000 (phone: +74852797775; e-mail: andcat@yandex.ru).

Lev Shmaglit, Graduate Student, Image Processing Laboratory, P.G. Demidov Yaroslavl State University, 14 Sovetskaya, Yaroslavl, Russia, 150000 (phone: +74852797775; e-mail: Lev_shmaglit@yahoo.com).

Maxim Golubev, PhD, Image Processing Laboratory, P.G. Demidov Yaroslavl State University, 14 Sovetskaya, Yaroslavl, Russia, 150000 (phone: +74852797775; e-mail: maksimgolubev@yandex.ru).

A new gender recognition algorithm, proposed in this paper, is based on non-linear SVM classifier with RBF kernel. To extract information from image fragment and to move to a lower dimension feature space we propose an adaptive feature generation algorithm which is trained by means of optimization procedure according to LDA principle. In order to construct a fully automatic face analysis system, gender recognition is used in connection with AdaBoost face detection classifier which selects candidates for analysis [15, 17]. Detected fragments are preprocessed to align their luminance characteristics and to transform them to uniform scale.

The rest of the paper is organized as follows. The scheme of the proposed algorithm is described in section 2. Section 3 considers training methodology and experimental setup. In section 4, the results of the proposed algorithm comparison with state-of-the-art gender classification methods is presented. Section 5 concludes the paper.

II. THE PROPOSED ALGORITHM

Classifier is based on Adaptive Features and SVM (AF-SVM). Its operation includes several stages, as shown in Fig. 1.

AF-SVM algorithm consists of the following steps: color space transform, image scaling, adaptive feature set calculation and SVM classification with preliminary kernel transformation. Input image $A_{Y \times Y}^{RGB}$ is converted from RGB to HSV color space and is scaled to fixed image resolution $N \times N$. After that we calculate a set of features $\left\{ AF_i^{HSV} \right\}$, where each feature represents the sum of all rows and columns of element-by-element matrix product of an input image and a coefficient matrix C_i^{HSV} with resolution $N \times N$, which is generated by the training procedure:

$$AF_i^{HSV} = \sum_N \sum_N A_{N \times N}^{HSV} \times C_i^{HSV}.$$

The obtained feature vector is transformed using a Gaussian radial basis function kernel:

$$k(z_1, z_2) = C \exp\left(\frac{-\left\| z_1 - z_2 \right\|^2}{\sigma^2} \right).$$

Kernel function parameters C and σ are defined during training. The resulted feature vector serves as an input to linear SVM classifier which decision rule is:

$$f(AF) = \text{sgn}\left(\sum_{i=1}^{m} y_i \alpha_i k(X_i, AF) + b \right).$$

The set of support vectors $\left\{ X_i \right\}$, the sets of coefficients $\left\{ y_i \right\}$, $\left\{ \alpha_i \right\}$ and the bias b are obtained at the stage of classifier training.
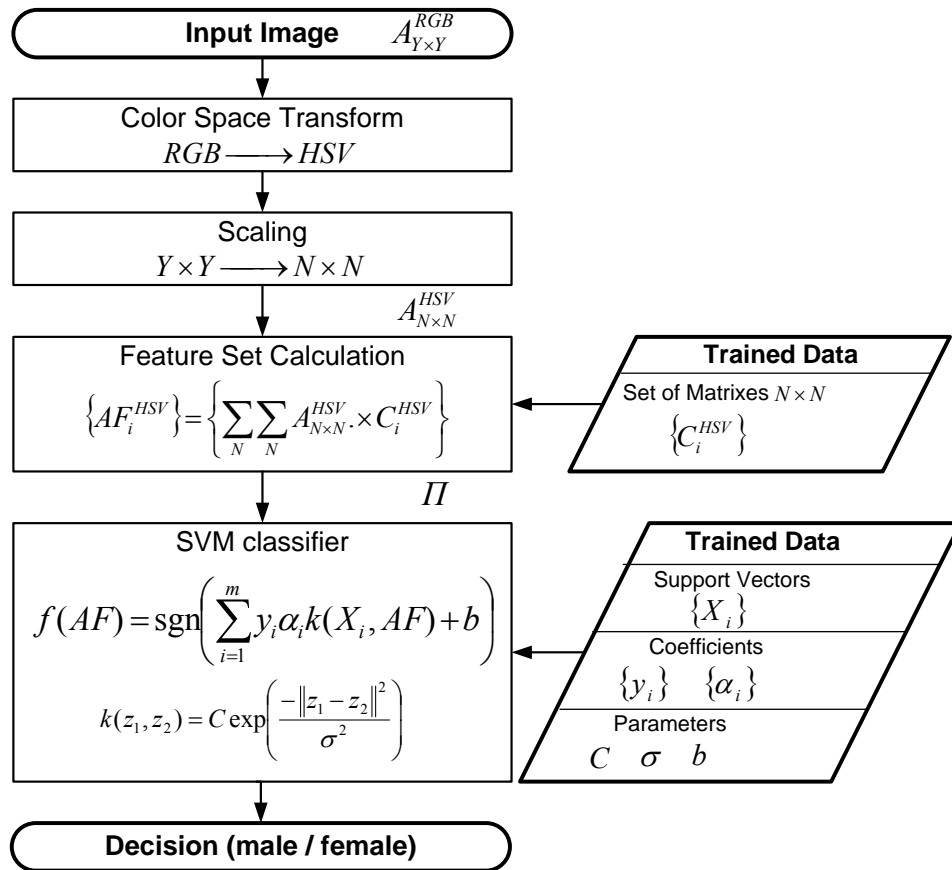
Fig. 1. The scheme of the proposed gender classification algorithm.

III. TRAINING AND TESTING SETUP

Both gender recognition algorithm training and testing require huge enough color image database. The most commonly used image database for the tasks of human faces recognition is the FERET database [18], but it contains insufficient number of faces of different individuals, that's why we collected our own image database, gathered from different sources (see Table I).

TABLE I
THE PROPOSED TRAINING AND TESTING IMAGE DATABASE PARAMETERS

Parameter	Value
The total number of images	8 654
The number of male faces	5 250
The number of female faces	5 250
Minimum image resolution	640×480
Color space format	RGB
Face position	Frontal
People's age	From 18 to 65 years old
Race	Caucasian
Lighting conditions, background and facial expression	No restrictions

Faces on the images from the proposed database were detected automatically by AdaBoost face detection algorithm. After that false detections were manually removed, and the resulted dataset consisting 10 500 image fragments (5 250 for each class) was obtained. This dataset was split into three independent image sets: training, validation and testing. Training set was utilized for feature generation and SVM classifier construction. Validation set was required in order to avoid the effect of overtraining during the selection of optimal parameters for the kernel function. Performance evaluation of the trained classifier was carried out with the use of the testing set.

The training procedure of the proposed AF-SVM classifier can be split into two independent parts: feature generation, SVM construction and optimization. Let's consider the feature generation procedure. It consists of the following basic steps:

- RGB → HSV color space transform of the training images (all further operations are carried out for each color component independently);
- scaling training images to fixed image resolution $N \times N$;
- coefficient matrix C_i^{HSV} random generation;
- feature value AF_i^{HSV} calculation for each training fragment;
- the utility function F calculation as a square of a difference between feature averages, calculated for "male" and "female" training image datasets, divided by the sum of feature variances [19]:

$$F = \frac{\left(\left\langle \left\{ AF_i^{HSV} \right\}_M \right\rangle - \left\langle \left\{ AF_i^{HSV} \right\}_F \right\rangle \right)^2}{\sigma \left\{ AF_i^{HSV} \right\}_M + \sigma \left\{ AF_i^{HSV} \right\}_F};$$

- iteratively in a cycle (until the number of iterations exceeds some preliminary fixed maximum value): random generation of coefficient matrix $\widetilde{C}_i^{HSV}$ inside the fixed neighborhood of matrix C_i^{HSV}, feature value $A\widetilde{F}_i^{HSV}$ calculation for each training

fragment, calculation of the utility function $\widetilde{F}$, transition to a new point ($F \to \widetilde{F}$, $C \to \widetilde{C}$), if $\widetilde{F} > F$;

- saving the matrix C_i^{HSV} after exceeding the maximum number of iterations;
- return to beginning in order to start the generation of the next (i+1) feature.

An optimization procedure, described above, allows to extract from an image only the information which is necessary for class separation. Besides, features with higher utility function value have higher separation ability. The feature generation procedure have the following adjusted parameters: training fragments resolution (N), the number of training images for each class (M), maximum number of iterations (T). The following values, as a compromise between reached separation ability and the training speed, were empirically received:

$$N = 65; \qquad M = 400; \qquad T = 10^5.$$

1000 features have been generated for each color component. At the second stage of training these features have been extracted from training images and were then used to learn an SVM classifier. SVM construction and optimization procedure included the following steps:

- calculation of the feature set, generated on the first stage of training, for each training fragment;
- feature normalization;
- learning an SVM classifier with different parameters of the kernel function;
- recognition rate (RR) calculation using validation image dataset;
- determination of optimal kernel function parameters (maximizing RR);
- learning a final SVM classifier with the found optimal kernel function parameters.

The goal of SVM optimization procedure is to find a solution with the best generalization ability, and thus with the minimum classification error. The adjusted parameters

are: the number of features in a feature vector (N2), the number of training images for each class (M2), the kernel function parameters σ and C.

Grid search was applied to determine optimal kernel parameters: SVM classifier was constructed varying $C = 10^{k1}$ and $\sigma = 10^{k2}$, where $k1$ and $k2$ – all combinations of integers from the range [-15 … 15]; during the search recognition rate was measured using validation image dataset. The results of this procedure are presented in Fig. 2. Maximum recognition rate (about 80%) was obtained for $C = 10^6$ and $\sigma = 10^8$.

Besides, we investigated the dependence of classifier performance from the number of features extracted from each color component – N2, and from the number of training images for each class – M2 (see Fig. 3).

The analysis shows that each feature has essential separation ability, and at $N2 = 30$ recognition rate reaches 79.5%. At the same time the growth of RR is observed both with the growth of N2 and M2 due to the accumulation of information about considered classes inside the classifier. Thus, to obtain a compromise between quality and speed the following parameters were chosen: $N2 = 30$; $M2 = 400$.

IV. EXPERIMENTAL RESULTS

In this section we present the results of the proposed AF-SVM algorithm comparison with state-of-the-art classifiers: SVM [13] and KDDA (Kernel Direct Discriminant Analysis) [19].

Classifier AF-SVM was trained according to a technique, given above. SVM and KDDA classifiers have far less adjustable parameters as they are working directly with image pixel values instead of feature vectors. To construct these classifiers the same training base, as for AF-SVM classifier, was used. The following conditions also were identical for all three considered classifiers: the number of training images for each class, training fragments resolution
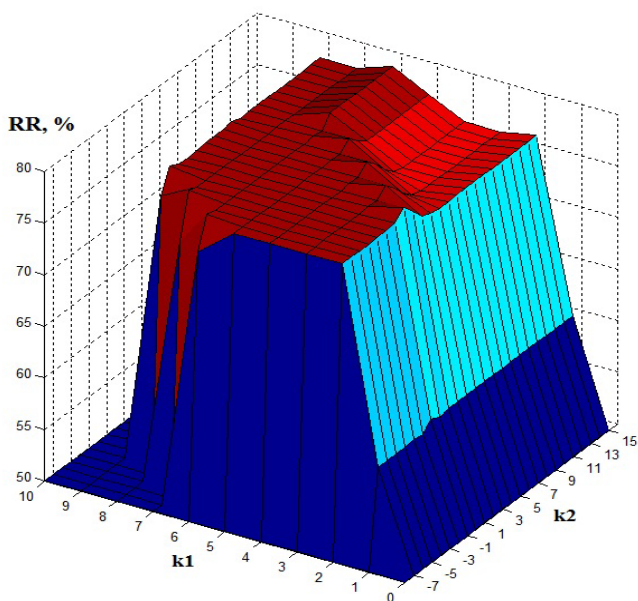

Fig. 2. The dependence of RR from kernel function parameters $C = 10^{k1}$ and $\sigma = 10^{k2}$.
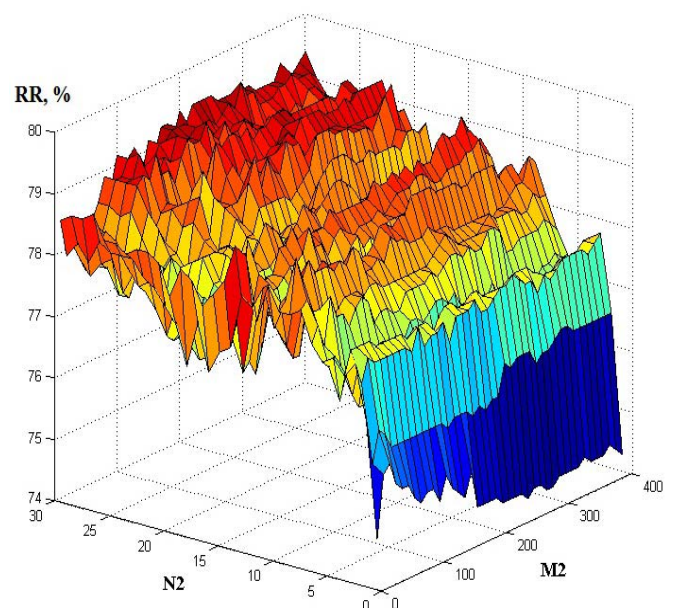

Fig. 3. The dependence of recognition rate from training procedure parameters.

and image preprocessing procedure. Optimization of SVM and KDDA kernel function parameters was held using the same technique and the same validation image dataset as used in case of AF-SVM classifier. Thus, equal conditions for independent comparison of considered classification algorithms, using testing image dataset, were provided.

For the representation of classification results we utilized the so called ROC-curve (Receiver Operator Characteristic) [20]. As there are two classes, one of them is considered to be a positive decision and the other – a negative. ROC-curve is created by plotting the fraction of true positives out of the positives (TPR = true positive rate) vs. the fraction of false positives out of the negatives (FPR = false positive rate), at various discrimination threshold settings. The advantage of ROC-curve representation lies in its invariance to the relation between the first and the second error type's costs.

The results of AF-SVM, SVM and KDDA testing are presented in Fig. 4 and in Table II. The computations were held on a personal computer with the following configuration: operating system – Microsoft Windows 7; CPU – Intel Core i7 (2 GHz) 4 cores; memory size – 6 Gb.
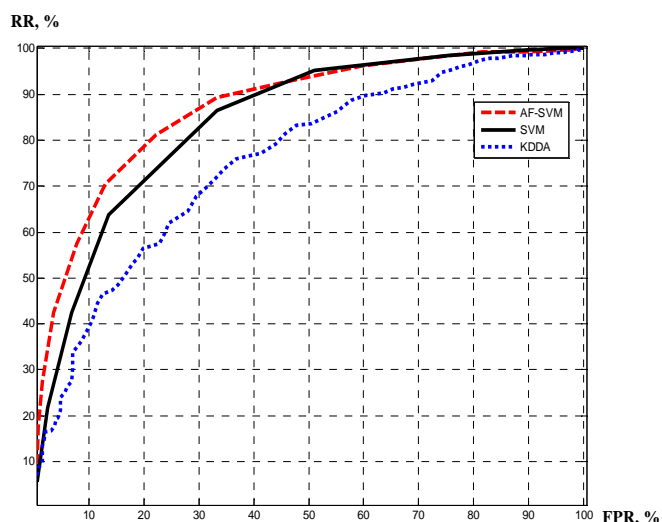
and KDDA approximately by 50%.

Such advantage is explained by the fact that AF-SVM algorithm utilizes a small number of adaptive features, each of which carries a lot of information and is capable to separate classes, while SVM and KDDA classifiers work directly with a huge matrix of image pixel values.

Let's consider the possibility of classifier performance improvement by the increase of the total number of training images per class from 400 to 5000. Experiments showed that SVM and KDDA recognition rates can't be significantly improved in that case. Besides, their computational complexity increases dramatically with the growth of the training dataset. This is explained by the fact that while the number of pixels, which SVM and KDDA classifiers utilize to find an optimal solution in a high dimensional space, increases it becomes harder and even impossible to find an acceptable solution for the reasonable time.

In the case of AF-SVM classifier the problem of the decrease of SVM classifier efficiency with the growth of the training database can be solved by the use of the small number of adaptive features, holding information about a lot of training images at once. For this purpose we suggest that each feature should be trained using a random subset (containing 400 training images per class) from the whole training database (containing 5000 images per class). Thus each generated feature will hold the maximum possible amount of information, required to divide the classes, and a set of features will include the information from each of the 10 000 training images.

On the stage of feature generation 300 features were trained according to the technique described above. After that an SVM classifier, utilizing these features, was trained similarly as before. Besides, we preserved the number of training images for SVM construction equal to 400, and thus the operation speed of the final classifier remained the same as in previous experiments – 65 faces processed per second.

The results of AF-SVM algorithm trained using expand dataset ($M = 5000$) and the initial AF-SVM classifier ($M = 400$) comparison are presented in Table III and Fig. 5.
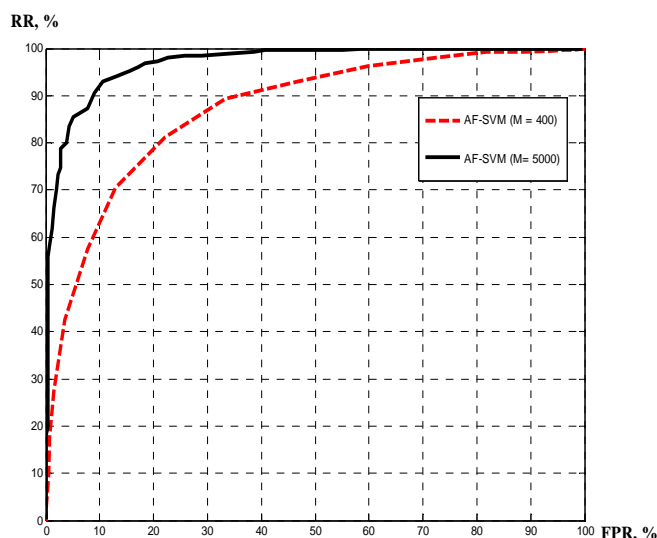

RR, %

FPR, %

Fig. 4. ROC-curves of tested gender recognition algorithms.

RR, %

FPR, %

Fig. 5. ROC-curves for AF-SVM algorithm trained on datasets of different size.

TABLE II
COMPARATIVE ANALYSIS OF TESTED ALGORITHMS PERFORMANCE

Algorithm ⟍ Parameter	SVM		KDDA		AF-SVM	
	True	False	True	False	True	False
Recognition rate						
Classified as "male", %	80	20	75.8	24.2	80	20
Classified as "female", %	75.5	24.5	65.5	34.5	79.3	20.7
Total classification rate, %	**77.7**	22.3	**69.7**	30.3	**79.6**	20.4
Operation speed, faces/sec	44		45		65	

The analysis of testing results show that AF-SVM is the most effective algorithm considering both recognition rate and operational complexity. AF-SVM has the highest RR among all tested classifiers – 79.6% and is faster than SVM

TABLE III
RECOGNITION RATE OF AF-SVM ALGORITHM TRAINED ON DATASETS OF DIFFERENT SIZE

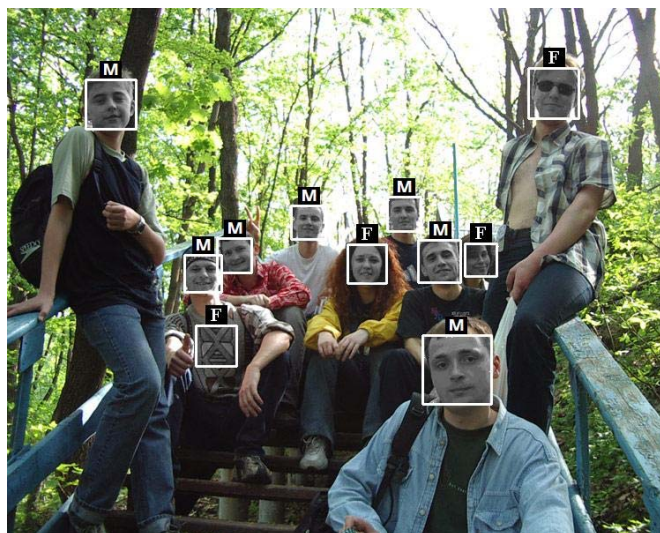| Algorithm Parameter	AF-SVM (M=5000)		AF-SVM (M=400)	
Recognition rate	True	False	True	False
Classified as "male", %	90.6	9.4	80	20
Classified as "female", %	91	9	79.3	20.7
Total classification rate, %	**90.8**	9.2	**79.6**	20.4

Fig. 6.   Visual example of the proposed automatic gender classification system.

The results show that AF-SVM algorithm together with the proposed training setup allow to significantly improve the classifier performance in case of increasing the training database size to 5000 images per class. Recognition rate of nearly 91% is achieved. Visual example of the proposed automatic gender classification system is presented in Fig. 6. Recognized classes "male" and "female" are designated by symbols "M" and "F" correspondingly (see Fig. 6).

V. CONCLUSIONS

A new classifier based on adaptive features and support vector machines, solving the problem of automatic gender recognition via face area analysis, is proposed. It shows recognition rate of 79.6%, which is 1.9% and 9.9% higher than that of SVM and KDDA correspondingly. The possibility of AF-SVM recognition rate improvement up to 91% in case of increasing the training database size to 5000 images per class is shown. The proposed algorithm allows to process 65 faces per second which is enough to use it in real time video sequence analysis applications.

The adaptive nature of the feature generation procedure allows using the proposed AF-SVM classifier for the recognition of any other object on an image (in addition to faces). For this purpose a new training database of the considered class should be constructed and a classifier should be retrained according to the described in this paper technique.

REFERENCES

[1] E. Alpaydin, *Introduction to Machine Learning*. The MIT Press, 2010.
[2] R. Szeliski, *Computer Vision: Algorithms and Applications*. Springer, 2010.
[3] D. Kriegman, M.H. Yang, N Ahuja, "Detecting faces in images: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, No 1, pp. 34-58, 2002.
[4] E Hjelmas, "Face detection: A Survey," *Computer vision and image understanding*, vol. 83, No 3, pp. 236-274, 2001.
[5] W. Zhao, R. Chellappa, P. Phillips, A Rosenfeld, "Face recognition: A literature survey," *ACM Computing Surveys (CSUR)*, vol. 35, No 4, pp. 399-458, 2003.
[6] B. Fasel, J Luettin, "Automatic facial expression analysis: A survey," *Pattern Recognition Letters*, vol. 36, No 1, pp. 259-275, 2003.
[7] E. Makinen, R. Raisamo, "An experimental comparison of gender classification methods," *Pattern Recognition Letters*, vol. 29, No 10, pp. 1544-1556, 2008.
[8] S. Tamura, H. Kawai, H. Mitsumoto, "Male/female identification from 8 to 6 very low resolution face images by neural network," *Pattern Recognition Letters*, vol. 29, No 2, pp. 331-335, 1996.
[9] M. Lyons, J. Budynek, A. Plante, S. Akamatsu, "Classifying facial attributes using a 2-d Gabor wavelet representation and discriminant analysis," *Proc. Int. Conf. on Automatic Face and Gesture Recognition*, pp. 202-207, 2000.
[10] A. Jain, J. Huang, "Integrating independent components and linear discriminant analysis for gender classification," *Proc. Int. Conf. on Automatic Face and Gesture Recognition*, pp. 159-163, 2004.
[11] Y. Saatci, C. Town, "Cascaded classification of gender and facial expression using active appearance models," *Proc. Int. Conf. on Automatic Face and Gesture Recognition*, pp. 393-400, 2006.
[12] Z. Sun, G. Bebis, X. Yuan, S.J. Louis, "Genetic feature subset selection for gender classification: A comparison study," *Proc. IEEE Workshop on Applications of Computer Vision*, pp. 165-170, 2002.
[13] C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121-167, 1998.
[14] N. Sun et al, "Gender classification based on boosting local binary pattern," *Proc. Int. Symposium on Neural Networks*, vol. 2, pp. 194-201, 2006.
[15] P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proc. Int. Conf. on Computer Vision and Pattern Recognition*, vol. 1, pp. 511-518, 2001.
[16] S. Gutta, H. Wechsler, P.J. Phillips, "Gender and ethnic classification of face images," *Proc. Int. Conf. on Automatic Face and Gesture Recognition*, pp. 194-199, 1998.
[17] L. Shmaglit, M. Golubev, A. Ganin, V. Khryashchev, "Gender classification by face image," *Proc. Int. Conf. on Digital Signal Processing and its Application (DSPA)*, vol. 1, pp. 425-428, 2012.
[18] P.J. Phillips et al, "The FERET evaluation methodology for face recognition algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, No 10, pp. 1090-1104, 2000.
[19] H. Gao, J. Davis, "Why direct LDA is not equivalent to LDA," *Pattern Recognition Letters*, vol. 39, No 5, pp. 1002-1006, 2006.
[20] T. Fawcett, "ROC Graphs: Notes and Practical Considerations for Researchers," *Pattern Recognition Letters*, vol. 27, No 8, pp. 882-891, 2004.