An Improvement of Face Detection Algorithm for Color Photos[1]

M. A. Egorova, A. B. Murynin, and I. V. Safonov

Dorodnicyn Computing Centre, Russia Academy of Sciences, ul. Vavilova, 40, Moscow, 119333 Russia
e-mail: marta.egerova@gmail.com; AMur@ccas.ru; ilia.safonov@gmail.com

Abstract—The problem of automatic detection of regions of interest (ROI) on color photos is considered. The efficiency of Viola-Jones face detector implemented in Intel OpenCV library is analyzed. The image size dependence of the algorithm characteristics is studied. It is shown that there are 2 significant drawbacks of this face detection algorithm as applied to the task mentioned. Some modifications are made for elimination of these drawbacks. To reduce false positives it is proposed to modify the algorithm with color segmentation and human skin tones analysis. To reduce the processing time the algorithm of downsizing preprocessing is proposed.

Key words: face detection algorithm, efficiency for practice applications.

DOI: 10.1134/S1054661809040105

INTRODUCTION

Face detection methods are widely used in different applied problems:

—real-time video surveillance systems based on biometric technologies [1, 2];

—detection of region of interest at images;

—image quality improving, for example, before printing, for collage creating [3] or photobook compositing, etc.;

—video compression.

Choice of one or another face detection algorithm depends on its efficiency in specific conditions of particular practical application. These conditions are determined the methods for algorithm efficiency estimation, and also the choice of image database used for test of algorithms [4, 5]. In our case determinative conditions are the following:

—the requirement to detect faces of different size in wide range of input image size;

—the requirement of efficient face detection in the images with complex background.

One of the acknowledged algorithms for face detection is the one developed by Viola and Jones [6]. There are three key features:

—image presentation as an integral image;

—a small number of critical visual features;

—a method for organizing cascade of classifiers.

The Intel OpenCV library provides an efficient implementation of the Viola-Jones face detector. This library is used in some research works [7, 8].

In this paper we analyze computational efficiency of this algorithm under various conditions and propose a number of modifications to improve the efficiency.

1. METHOD OF EVALUATING ALGORITHM EFFICIENCY

We are testing and analyzing the following kinds of face detection algorithms:

base algorithm—implementation of Viola-Jones algorithm in OpenCV library without any modifications;

modified algorithms—algorithms obtained as a result of applying one or more modifications described in this paper.

Algorithm investigation is conducted using special prepared test data set (a set of images with predefined characteristics). It consists of typical images of typical sizes, as noted in [9]. The data set contains images with faces and without ones. Faces are presented at ~50% of the images, at which the number of faces vary from 1 to 26 (number of faces per image is distributed as shown in table). Eye distance varies in the interval from 5–7 to ~675 pixels. The most part of the faces is close to frontal, however, face position can differ from frontal in some cases.

Faces distribution between images

Faces count	0	1	2	3–5	>5
Image count, %	54.1%	28.2%	9.1%	6.1%	2.5%

[1] The article is published in the original.

Received March 1, 2009

Work with the test data contains three stages: *preprocessing* stage, *data processing* stage and the stage of characteristics *analysis* for the algorithm under consideration.

The images are marked in specific manner at the *preprocessing* stage (before testing): faces regions are selected manually on source images. Coordinates of these regions are registered in configuration file (file of accompanying information), which is used during algorithm testing to get sample values of faces position on an image.

At *data processing* stage the following operations are performed using the algorithm being tested:

—firstly accompanying information for current image is read from configuration file;

—then face detection algorithm is applied. Processing time and the results of face detection algorithm (coordinates of the detected faces) are registered. All results and algorithm characteristics are written to intermediate data file to subsequent analysis.

At the *analysis* stage the data from intermediate and accompanying file are compared, corresponding events of different type are registered and the characteristics of tested algorithm are calculated.

The following operations are performed for each image using intermediate and accompanying file:

—each detected object is checked for real presence of a face in the image by comparison of object size and position with the data extracted from the accompanying file. If a face is actually absent then the event “false positive” is registered;

—if face is present in the image then its coordinates are compared with sample values from the accompanying file: if the coordinates’ difference lies in acceptable range then event of correct detection is registered. After that the next face is considered;

—when all faces are matched with sample data of preprocessed stage, *false negatives* are calculated. Integral characteristics of investigated algorithm are calculated in terms of considered set of conditions after all test images are processed. We define this integral characteristic below.

We register event “false negative” if the algorithm has not found a face for some region stored in configuration files inside acceptable range. Number of false negatives is defined as:

$$N_{\mathrm{FN}} = N_{\mathrm{count}} - (N_{\mathrm{algo}} - N_{\mathrm{FP}}), \tag{1}$$

where N_{FN}—number of false negatives, N_{count}—faces count in the image according to the configuration file, N_{algo}—detected faces count, N_{FP}—number of false positives.

Difference ($N_{\mathrm{algo}} - N_{\mathrm{FP}}$) shows number of faces detected correctly.

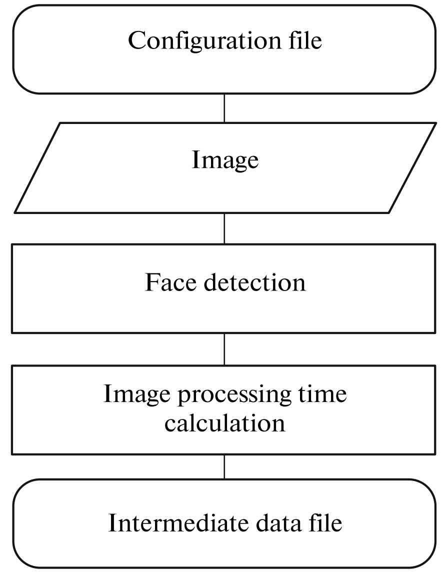

Fig. 1. Data processing stage.

To estimate probability of second kind error we employ the characteristic *FNR* (*False Negatives Rate*). It is defined as a percentage ratio of the number of false negatives to the number of true faces in the image:

$$\mathrm{FNR} = N_{\mathrm{FN}}/N_{\mathrm{count}}. \tag{2}$$

Probability of the error of first kind is estimated using average number of false positives $\langle N_{\mathrm{FP}}\rangle$. This probability is estimated as number of false positives per image:

$$\langle N_{\mathrm{FP}}\rangle = \frac{\sum_i N_{\mathrm{FP}}^i}{N}, \quad i = 1\ldots N, \tag{3}$$

where N – number of images in whole data set. Process of analysis stage is shown in Fig. 2.

2. BASE ALGORITHM ANALYSIS

Characteristics changes are the basis for further estimation of base algorithm modification. Measured dependence for average processing time versus linear image size is shown in Fig. 3.

It is evident that time for the face detection on an images with the linear size of more than 1500 pixels along the largest dimension is more than 5 s. As the image size increases, there is a significant increase in the processing time. Note that for the images with linear size less than 1000 pixels face detection time is less than 2–3 s.

This dependence on image size can be explained by the initial purpose of the algorithm. Originally Viola-Jones algorithm was developed for face detection in video bit stream. In this case frame size is quite small.

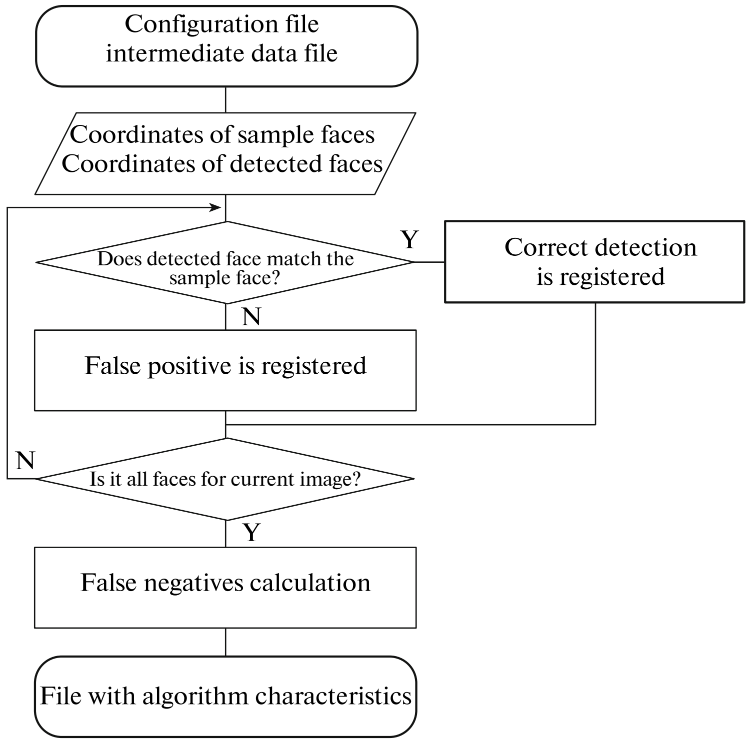

Fig. 2. Steps of data analysis stage.

Therefore algorithm does not work optimal for quite large images.

Dependence $\langle N_{\text{FP}} \rangle$ versus linear image size is shown in Fig. 4.

The average number of false positives is more than 3 for the large images. It means that 3 wrong face regions have been detected by the algorithm.

Graph for FNR against the linear image size in case of base algorithm is of the form of one shown in Fig. 5.

The graph shows that number of false negatives is acceptably low for the linear image sizes over 700 pixels. If the linear image size is less than this value the number of false negatives increases dramatically. It is explained by the following circumstance: as the image gets smaller, so does the face. As a result, the face becomes less detectable by the algorithm. Thus, it is necessary to reduce algorithm working time for large images. Also it is necessary to reduce the number of false positives keeping the number of false negatives constant.

3. ALGORITHM MODIFICATIONS AND THEIR ANALYSIS

We have carried out some modifications based on the previous analysis to improve algorithm efficiency in practical applications. The algorithm modifications

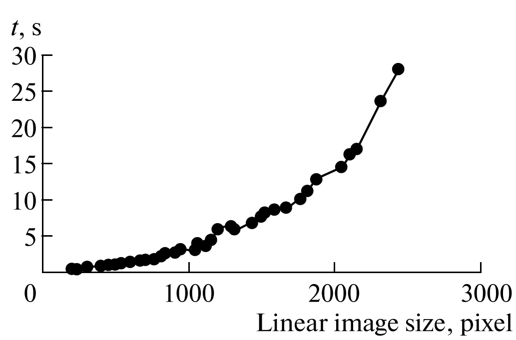

Fig. 3. Graph of average processing time against linear image size for base algorithm.

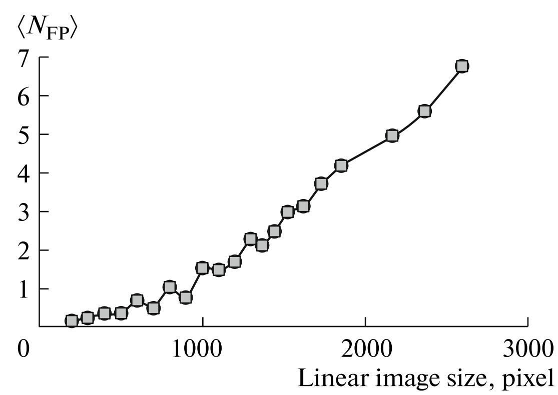

Fig. 4. Graph $\langle N_{\text{FP}} \rangle$ against linear image size.

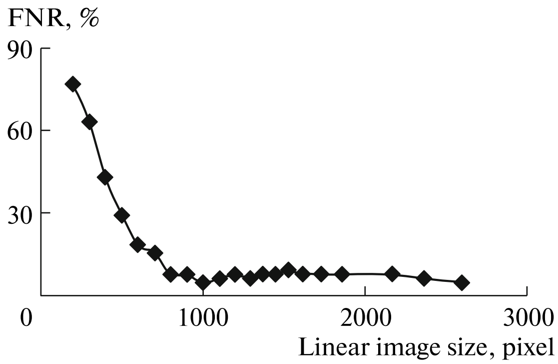

Fig. 5. Graph for FNR against linear image size.

can be separated into following main groups of procedures:

—the first part of these procedures (*preprocessing*) is performed before base algorithm applying;

—the second part of these procedures (*post-processing*) is performed after base algorithm has finished. These procedures are used to make the results of base algorithm more accurate taking into account some specific environment parameters.

Specifically, conducted modifications add up to the following:

—the optimal image downsampling is applied at *preprocessing* step, then modified image is used as entry data for base algorithm;

—optimization of search region using color information is performed after termination of base algorithm work. More generally algorithm modifications can be presented as a sequence of image processing procedures shown in Fig. 6.

We describe these modifications and some results obtained in more detail below.

3.1. Modification 1: Preprocessing with Downsampling

To improve image processing time it is evident to reduce amount of processed data. Namely it is a reduced image size. However it can influence on face detection, therefore it is important to find optimal image resizing parameters, which minimize image processing time and care about acceptable error of first and second kind level.

Image resizing parameters were determined empirically based on processing results and analysis of experimental data. As a result the following image processing approach has been chosen and taken as a principle for base algorithm modifications (see Fig. 7).

—For each image, where face detection is required, a linear image size is checked.

—If linear image size is out of range [750, 1000] pixels, then a scale factor is determined. Scale factor is the ratio of the actual linear image size to the optimal linear image size (we chose the optimal linear image size of 850 pixels).

—If linear image size satisfies the constraint of range [750, 1000] pixels, then image scaling should not be applied.

Comparative results of face detection time before and after image downsizing is shown in Fig. 8.

It is obvious that time for face detection is essentially reduced. An error of second kind remains at acceptable level corresponding to image size of approximately 850 pixels.

Reduce of false positives is important result of the modification with downsampling. The number of false positives is reduced to level of 1–2 per image that is quite acceptable.

3.2. Modification 2: Skin Tone Analysis

Skin tone analysis is major in face detection tasks. Generally color segmentation is employed to initial image to reduce region of interest used face detection algorithms and then more complex algorithms are applied to analyze selected regions.

Face searching in OpenCV library is accomplished on a grayscale image. In order to reduce the number of false positives, we use information about the color of the image part detected by base algorithm as face.

The color of pixels in such regions should be like that of human skin. Different approaches are proposed in skin tone analysis.

The main variation in skin appearance is largely due to the luminance component; the normalized *rgb*

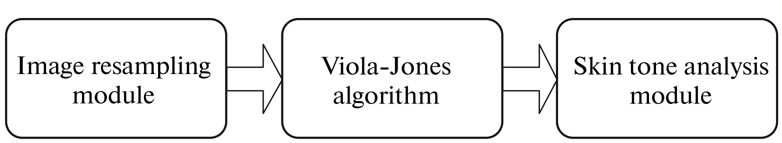

Fig. 6. Viola-Jones face detection algorithm modifications steps.

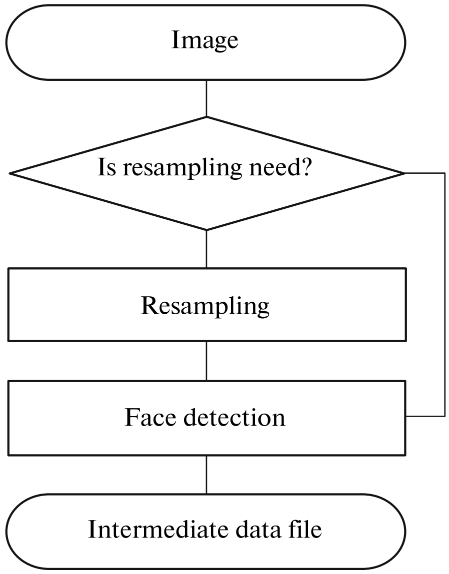

Fig. 7. Flow chart of face detection via modification 1.

color space is generally preferred to filter out the dependence on the image illumination.

To transform *RGB* values into corresponding chromaticities *r*, *g*, *b* the following equations are used:

$$r = cR/I, \quad g = cG/I, \quad b = cB/I, \tag{4}$$

where $R, G, B \in [0, 255]$ and c is a constant ($c = 100$), $I = R + G + B$ is intensity.

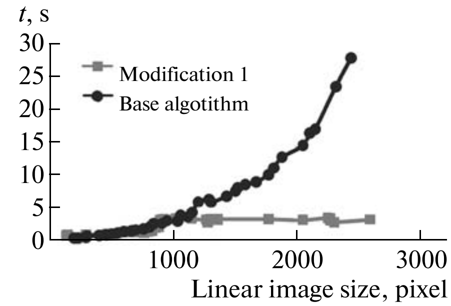

Fig. 8. Comparative graph for average face detection time against linear image size for base algorithm and algorithm with preprocessed image downsizing.

Segmentation only in *rgb* space gives bad results for dark skin-tones. Therefore a number of researches [10–14] use segmentation in *HSV* space.

We have considered different opportunities and chosen a method proposed by authors of [15] which combine the outputs of the threshold operations in both *rgb* and *HSV* color spaces. It is checked whether each pixel of face detection region belongs in the range of *rgb* color space: $r \in [38, 55]$, $g \in [25, 38]$ or in the range of *HSV* color space: $H \in [0, 50] \cup [340, 360]$, $S > 0.2$, $V > 0.35$.

Since the face detection region consists of both skin (e.g., face) and non-skin parts (e.g., hair, background detail, etc.) we consider the region to be a face if more than half of its pixels (N_{region}) are deemed to be skin like (N_{skin}).

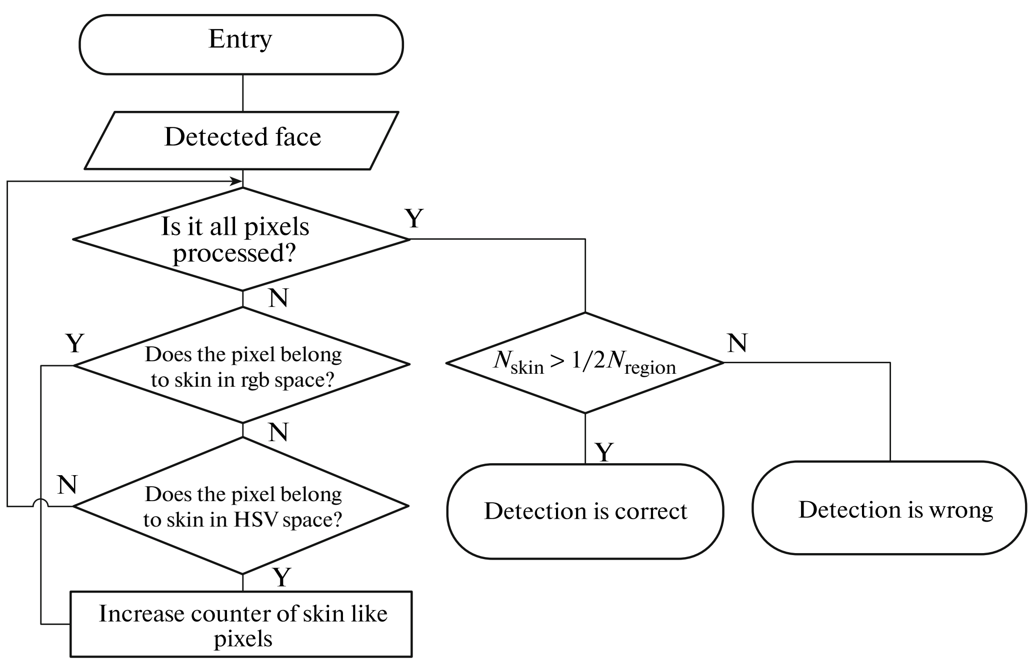

Fig. 9. Skin tone analysis.

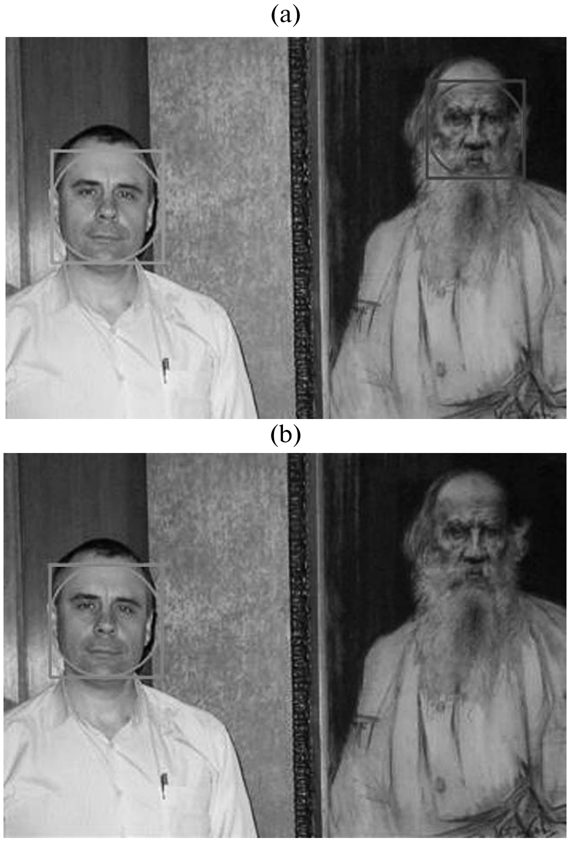

Fig. 10. Detected faces: (a) before applying Modification 2, (b) after applying Modification 2.

Figures 10a, 10b present an example of described algorithm modification:

before modification 2 the false positive is observed—the face on the picture is detected whereas it is not a face in terms of color criteria;

after applying the modification 2 there is no false positive.

A further series of experiments has been performed to build dependence between the number of false positives and the linear image size. The obtained dependence after skin tone analysis was compared to corresponding dependence without image color segmentation. Comparative graphs are shown in Fig. 11. The number of false positives for large images is considerably lower if modification 2 is applied, i.e., if the color information is being analyzed.

CONCLUSIONS

Two modifications of the base algorithm have been presented to improve its efficiency for practice applications.

—Preprocessing step with image scaling before face detection makes it possible to reduce appreciably the face detection time for images with linear size

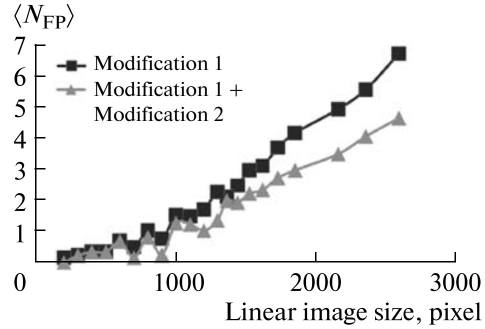

Fig. 11. Comparative graphs $\langle N_{\mathrm{FP}} \rangle$ for base and modified algorithms.

more than 2000 pixels. The reduction of face detection time amounts to about ~5 times in the case of linear image size of 3000 pixels.

—Color segmentation of found face area allows reducing a number of false positives in ~1.5 times for images with linear size more than 2000 pixels.

These modifications do not change the number of false negatives.

REFERENCES

1. A. B. Murynin, "Automatic Person Identification System Based on Stereo Images," Journal of Computer and Systems Sciences International **38** (1), 100–108 (1999).
2. A. A. Desyatchikov, D. V. Kovkov, V. V. Lobantsov, K. A. Makovkin, I. A. Matveev, A. B. Murynin, and V. Ya. Chuchupal, "A System of Algorithms for Stable Human Recognition," Journal of Computer and Systems Sciences International **45** (6), 958–969 (2006).
3. C. Rother, L. Bordeaux, Y. Hamadi, and A. Blake, "Autocollage," ACM Transaction on Graphics **25**, 847–852 (2006).
4. E. Hjelmas and B. K. Low, "Face Detection: A Survey," Compute Vision and Image Understanding **83**, 236–274 (2001).
5. A. Murynin, P. Bazanov, D. Buryak, W. Mun, and H. Yang, "Comparison of Gabor Wavelet and Neural Network–Based Face Detection Algorithms," in *Proc. IASTED International Conf. Signal and Image Processing* (Honolulu, 2005), pp. 178–184.
6. P. Viola and M. Jones, in *Proc. Robust Real–Time Object Detection* (Vancouver, July 13, 2001).
7. C.-C. Hsu and H. T. Chang, "An Improved Face Detection Method in Low–Resolution Video," in *Proc. IIH–MSP* (Kaohsiung, 2007), Vol. 2, pp. 419–422.
8. Q. Zhu, S. Avidan, M.-C. Yeh, and K.-T. Cheng, "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients," in *Proc. IEEE CVPR* (New York, 2006).
9. D. Wueller and R. Fageth, "Statistic Analysis of Millions of Digital Photos," in *Proc. of SPIE/IS&T* (San Jose, 2008), Vol. 6817.
10. C. Garcia and G. Tziritas, "Face Detection Using Quantized Skin Color Regions Merging and Wavelet

Packet Analysis," IEEE Trans. Multimedia **1** (3), 264–277 (1999).

11. K. Herodotou, K. Plataniotі, and A. Venetsaopoulos, "Automatic Location and Tracking of the Facial Region in Color Video Sequences," Signal Process., Image Comm. **14** (5), 359–388 (1999).

12. K. Sobottka and I. Pitas, "A Novel Method for Automatic Face Segmentation, Facial Feature Extraction and Tracking," Signal Process., Image Comm. **12** (3), 263–281 (1998).

13. S. Tsekeridou and I. Pitas, "Facial Feature Extraction in Frontal Views Using Biometric Analogies," in *Proc. EUSIPCO_98* (Rhodes, 1998).

14. Y. Wang and B. Yuan, "A Novel Approach for Human Face Detection from Color Images under Complex Background," Pattern Recognition **34** (2001).

15. B. Smolka, K. Czubin, Y. Hardeberg, et al., "Towards Automatic Redeye Effect Removal," Pattern Recognition Letters **24**, 1767–1785 (2003).

Marta A. Egorova received her MS degree in cybernetics from Moscow Engineering Physics Institute/State University (MEPhI), Russia in 2008. At present time she is a post-graduate student of MEPhI. Her current research interests include image quality estimation, image enhancement and pattern recognition. She has got 4 publications in image processing and pattern recognition.

Alexander B. Murynin (born in 1960) received his MS degree in 1984 from Moscow Institute of Physics and Technology (MIPT), Department of Physics and Power engineering. He received his PhD degree in 1991 from Scientific Industrial Corporation "Kometa." Since 1993 he works in Dorodnicyn Computing Centre of Russian Academy of Sciences, where he has a position of Senior Researcher. He has accomplished a series of projects in development of multimodal biometric algorithms for some leading software vendors, among them Samsung Research Center, He has got about 90 publications in image processing and pattern recognition, computer simulation.

Il'ia V. Safonov received his MS degree in automatic and electronic engineering from Moscow Engineering Physics Institute/University (MEPhI), Russia in 1994 and his PhD degree in computer science from MEPhI in 1997. Since 1998 he is an associate professor of Cybernetics faculty of MEPhI while conducting researches in image segmentation, features extraction and pattern recognition problems. Since 2004, Dr. I. Safonov has joint Image Enhancement Technology Group, Printing Technology Lab, Samsung Research Center, Moscow, Russia where he is engaged in photo, video and document image enhancement projects. He has got about 90 publications in image processing and pattern recognition.