Eye Center Localization on a Facial Image Based on Multi-Block Local Binary Patterns

Anatoly Nikitin, Vladimir Khryashchev, Olga Stepanova
Yaroslavl State University
Yaroslavl, Russia
anatolynikitinyar@gmail.com, vhr@yandex.ru,
olga1stepanova@yandex.ru

Igor Kosterin
Fire and Rescue Academy
Ivanovo, Russia
kosteriniv@gmail.com

Abstract—In this paper an eye center localization algorithm based on multi-block local binary patterns is described. Performance of the suggested algorithm is compared to another methods based on Bayesian approach and image gradients. Visual examples of eye center localization results are provided.

I. INTRODUCTION

Being probably the most expressive and salient features of the human face, eyes are very important sources of information for face analysis. Accurate and efficient eye locating in a given face image is very important for wide range of practical problems in computer vision, including face recognition, face alignment, gaze and pose estimation, expression analysis, etc [1-4].

It is not uncommon to mistakenly think that eye localization is a simple task since eyes are just a simple structure in the face. However, eyes have its unique geometric, photometric and motion characteristics, and changes due to these three characteristics would lead to a very complicated nonlinear manifold of eyes. In particular, the following factors have significant influence on the states of the eyes [5]:

- wide variety of colors and shapes of eyes;

- emotions and facial expressions: for example, eyes of laughing person can be almost closed;

- different occlusions: eyes are often occluded by hair, myopia glasses or sunglasses;

- pose: it is possible that one eye is completely occluded in a profile face;

- imaging condition and quality: environment factors, such as lighting (varying in spectra, source distribution, and intensity), may change the appearance of the eyes in different ways. Moreover, the commonly seen factors in the real world, such as low resolution, blurring or detailed texture missing, may also lead to poor image quality.

The great challenges of eye localization under the uncontrolled conditions illustrated on images from Yale [6] and Robotics [7] databases (Fig. 1).

Eye localization can be successfully used as a preprocessing stage for face recognition [1]. Many authors [8-11] observed that the accuracy of eye localization has a significant effect on face recognition accuracy, which urge the need for developing robust and accurate eye localization techniques in real-life scenarios.

In the recent thirty years, problem of eye localization has gained increasing attention from both the academic and industrial communities. However, this problem is still far from being resolved. Most approaches suggested by researches in the recent time can be divided into four groups depending on the way of building of eye model [5]:

- measuring eye characteristics: this type of method exploits the inherent features of eyes: distinct shapes, strong intensity contrast, etc;

- learning statistical appearance model: this type of method tries to extract useful visual features from photometric appearance, eye model is then learned from a large set of training images;

- exploiting the spatial structure of the face: this approach explores the geometrical regularity between eyes and other facial features in the face context:

- describing the local structure of the image: this approach is based on image representation as a low-dimensional vector using local structure of the image (local binary patterns, for example).

The major contribution of this paper is to present an eye center localization algorithm based on local binary patterns (LBP). The structure of the algorithm is explained in Section II. Section III is devoted to methodology of evaluation. The results of testing of the algorithm in comparison to other popular methods are presented in Section IV.

II. EYE CENTRE LOCALIZATION APPROACH BASED ON MULTI-BLOCK LOCAL BINARY PATTERN

The basic idea of the LBP is to avoid image representation as a high-dimensional vector which contains a lot of redundant information, and describe the local structure of an image. Extracted features will have lower dimension.

Fig. 1. An illustration of the great challenges of eye localization under the uncontrolled conditions

The original LBP was introduced by Ojala [12]. This theoretically very simple yet efficient multiresolution operator is defined for each pixel by thresholding the 3×3 neighborhood pixel value with the center pixel value. In this way, it can give us a binary sequence defined by local structure of an image.

LBP is computationally efficient because it works with integer arithmetic (it can achieve real-time performance in some tasks), and it is invariant to changes in the brightness of the image caused by shooting in different lighting conditions [13].

In the recent years a number of approaches based on Multi-Block LBP (MB-LBP) has gained increasing attention from researches in the field of image processing [14-17]. The basic idea of MB-LBP is to process rectangle regions of an image with classic LBP operator. To encode the rectangles, the MB-LBP operator is defined by comparing the central rectangle's average intensity with those of its neighborhood rectangles. The following steps are offered to apply MB-LBP to the problem of eye center localization:

A. Collecting the eye training samples:

The samples of eyes and eye-like negative patterns are collected in three different scales:

a) larger scale – samples contain both eye and eyebrow information for better processing of eyes with myopia glasses or sunglasses. The geometry of these eye training samples is shown in the left of Fig. 2;

b) standard scale – samples only contain the eye itself but exclude the eyebrow, nose and others. The geometry of these eye training samples is shown in the center of Fig. 2;

c) smaller scale – samples only contain the central part of eye itself. The geometry of these eye training samples is shown in the right of Fig. 2.

The training is made for every scale independently.

The training set is represented as pairs of values $(x_1, y_1) \ldots (x_N, y_N)$, where $y_i \in \{-1,1\}$ is the class label of the example $x_i \epsilon R^n$.

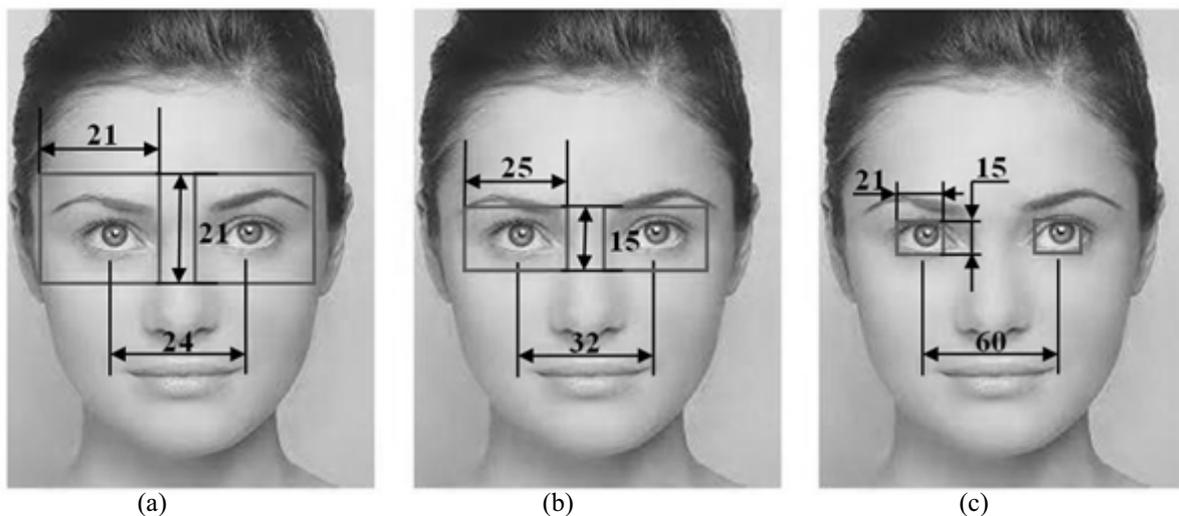(a)   (b)   (c)

Fig. 2. The geometry of eye training samples: a – larger scale; b – standard scale; c – smaller scale

B. Choosing the start weights w_i for samples:

The start weights are chosen according to the number of eye samples and eye-like negative patterns. In case of equal number of eye samples and eye-like negative patterns the start weights could be chosen as following:

$$w_i = \frac{1}{N}. \tag{1}$$

In our experiments we selected larger weights for eye samples, since the number of eye samples was lower than the number of eye-like negative patterns. In any case the start weights should be normalized:

$$\sum_{i=1}^{N} w_i = 1. \tag{2}$$

C. Constructing of integral image:

Integral image is used for rapid computation of MB-LBP operators. The integral image at pixel location (x, y) is constructed as follows:

$$ii(x, y) = \sum_{x' \le x, y' \le y} i(x', y') \tag{3}$$

where $ii(x, y)$ is the integral image, $i(x, y)$ is the original image. Using the following pair of recurrences the integral image can be computed in one pass over the original image:

$$\begin{aligned}s(x, y) &= s(x, y-1) + i(x, y) \\ ii(x, y) &= ii(x-1, y) + s(x, y)\end{aligned}, \tag{4}$$

where $s(x, y)$ is the cumulative row sum, $s(x, -1) = 0$, $ii(-1, y) = 0$.

D. Calculation of all possible MB-LBP values for the selected window size:

MB-LBP is a generalized form of local patterns that processes rectangle regions of an image. To encode the rectangles, the MB-LBP operator is defined by comparing the central rectangle's average intensity g_c with those of its neighborhood rectangles $\{g_0, \dots g_8\}$. In this way, it can give us a binary sequence. An output value of the MB-LBP operator can be obtained as follows [14]:

$$MB - LBP = \sum_{i=1}^{8} 2^i s(g_i - g_c), \tag{5}$$

where s is sigma function:

$$s(x) = \begin{cases} 1, & x \ge 0 \\ 0, & x < 0 \end{cases} \tag{6}$$

g_c is the average intensity of the center rectangle, g_i are intensities of its neighborhood rectangles (see Fig. 3 for an example).

In general, 256 various patterns may be obtained for a 3x3 window size, some of them are presented in Fig. 4.
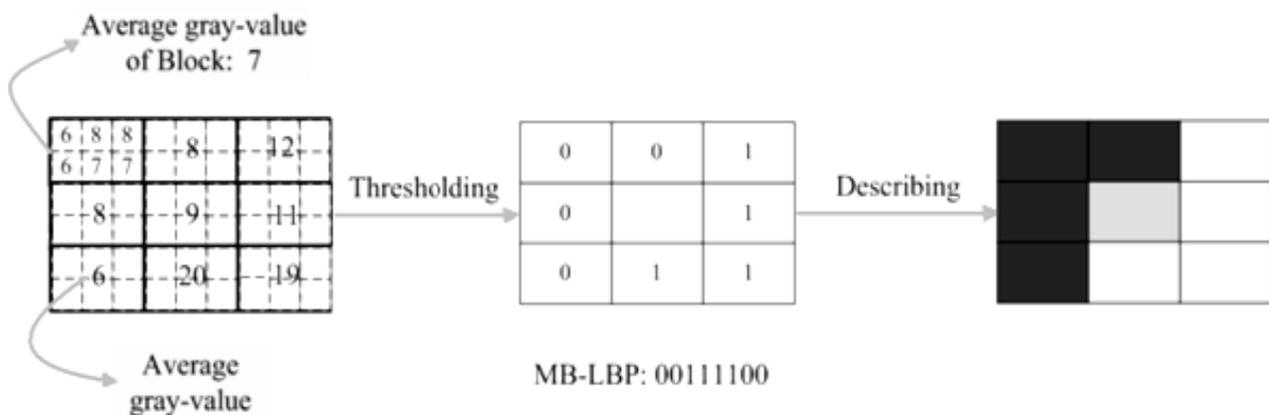

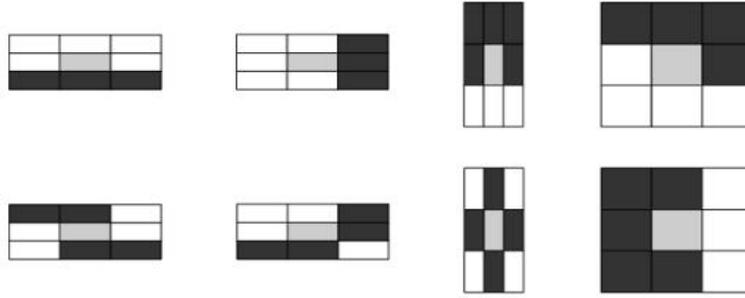Fig. 3. Example of counting of MB-LBP. MB-LBP in this case is equal to 00111100

Fig. 4. A randomly chosen subset of the MB-LBP features

E. Choosing weak classifiers

For each MB-LBP feature, we adopt multi-branch tree as weak classifiers. The multi-branch tree totally has 256 branches, and each branch corresponds to a certain discrete value of MB-LBP features.

The weak classifier can be defined as:

$$f^k(\mathbf{x}) = f^k(x^1,\ldots,x^k,\ldots,x^S) = \begin{cases} a_1, & if\ x^k = 1 \\ \ldots \\ a_j, & if\ x^k = j \\ \ldots \\ a_{256}, & if\ x^k = 256, \end{cases} \quad (7)$$

where a_j is defined as follows:

$$a_j = \frac{\sum_i^N w_i y_i \delta(x_i^k = j)}{\sum_i^N w_i \delta(x_i^k = j)}. \quad (8)$$

Feature vector elements $x^1,\ldots,x^k,\ldots,x^S$ for images from the training set are values of all possible MB-LBP. The general number s of all possible MB-LBP features is defined by chosen window size.

In step t, the weak classifier $f_t(\mathbf{x})$ is chosen so as to minimize the weighted squared error:

$$f_t(\mathbf{x}) = \min_{f^k \in F} \sum_{i=1}^N w_i (y - f^k(\mathbf{x}_i))^2. \quad (9)$$

After that the weights w_i are updated:

$$w_i' = w_i e^{-y_i f_t(\mathbf{x}_i)}. \quad (10)$$

F. Constructing strong classifier:

The strong classifier is constructed as a superposition of weak classifiers:

$$F(x) = \sum_{i=1}^T f_i(x). \quad (11)$$

For three scales of eye patterns three strong classifiers $F'(x)$, $F''(x)$, $F'''(x)$ are constructed according (11).

To achieve high localization precision we used a probabilistic cascade (P-Cascade) [15]. Firstly, we search m_1 points on face image in larger scale with the highest probability values

$$p' = \frac{e^{F'(x)}}{e^{F'(x)} + e^{-F'(x)}}. \quad (12)$$

These points are rough assessment of eye position. For selected m_1 points and their neighbor pixels in standard scale we choose m_2 with the highest probability

$$p'' = \frac{e^{F''(x)}}{e^{F''(x)} + e^{-F''(x)}}. \quad (13)$$

The final eye coordinate is the point with the highest probability p''', calculated similar to (12) and (13) for smaller case. For facial images with low resolution only coarse classifier is used. Similarly, for images with high resolution the bigger number of classifiers can be constructed. Thus, the algorithm can be adapted for images of different quality.

III. EVALUATION

To illustrate the performance of the proposed method, we conduct experiments on the BioID database [18] and FERET database [19]. The left and right eye centres are

annotated and provided together with the images in this databases. Samples of images from the BioID and FERET are shown in Fig. 5.

In experiments, the proposed method based on MB-LBP is compared with state-of-the-art methods: the simple Bayesian approach [20] and algorithm based on image gradients [21].

The FERET images are all taken indoors, with good resolution, image quality, and limited variation in lighting. Pose of the faces in these images is typically very close to frontal. In the experiments, we divide 3,363 images from FERET into two parts, one for training the other for testing. Ground truth eye positions are provided for all images. MB-LBP and Bayesian algorithms are trained on 1000 of 3,363 images while training for gradient detector is not required. Of the rest 2,363 frontal images, the 2,350 images for which the face detector detected a face correctly were retained for testing.

The BioID database consists of 1,521 grey level images of 23 different subjects and has been taken in different locations and at different daytimes, which result in variable illumination conditions comparable to outdoor scenes. Of the 1,521 images, the 1,469 images for which the face detector detected a face correctly were used as the test set for all algorithms.

The normalized error [22] is used to evaluate the error between the localized eye positions and the ground truth:

$$err = \frac{\max(\left\| l - l_g \right\|, \left\| r - r_g \right\|)}{\left\| l_g - r_g \right\|}, \qquad (14)$$

where l_g and r_g are the ground truth positions of the left and right eye respectively; l and r are the eye positions localized by an algorithm.

(a) (b)

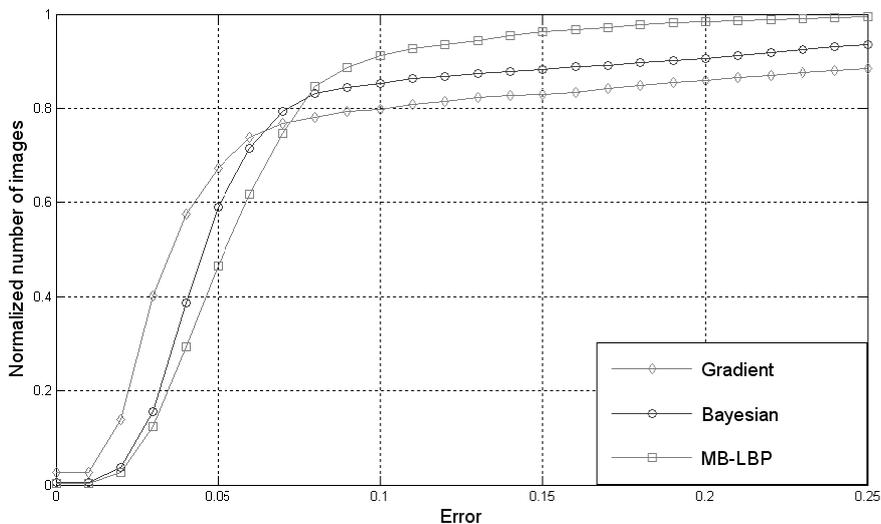Fig. 5. Samples of images from databases: a – FERET; b – BioID

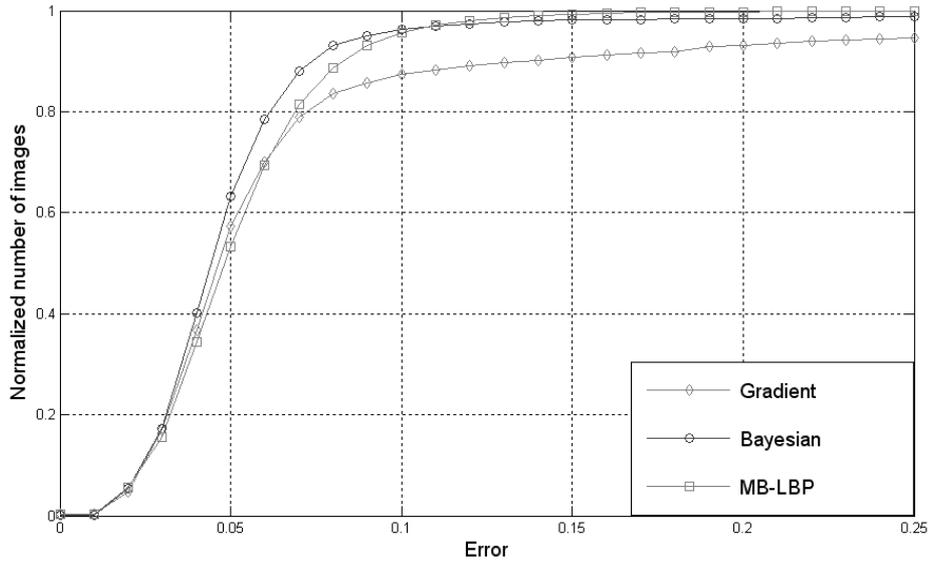Fig. 6. Results of eye center localization for the BioID database

Fig. 7. Results of eye center localization for FERET database

IV. EXPERIMENTAL RESULTS

In this section we report results of the three approaches: the simple Bayesian approach, MB-LBP and algorithm based on image gradients. The normalized error curves on FERET and BioID are shown in Fig. 6 and Fig. 7. The curves show the normalized number of images which has the normalized error rate equal to or less than corresponding value on axis of abscissae.

From the results on the BioID database (Fig. 6) we can see that MB-LBP has better precision in large error range, but poorer precision in small error range. MB-LBP localizes eye centers in 100% of images with error rate equal to or less than 0.25. It also has error rate equal to or less than 0.1 for more than 90% of images while other methods process the same number of images with much bigger error rate.

From the results on FERET database (Fig. 7) we can see that the performances of all methods are very similar due to the good quality of FERET.

The time costs for the simple Bayesian approach, MB-LBP and algorithm based on image gradients are shown in Table I. It should be noted that time costs are recorded on facial image with resolution 170x170 pixels.

TABLE I. LISTS OF TIME COSTS EVALUATED ON FACIAL IMAGE

Algorithm	Time of processing
Gradient	587 ms
Bayesian	367 ms
MB-LBP	44 ms

Visual examples of results of eye center localization made by MB-LBP algorithm are shown on Fig. 8.

V. CONCLUSION

This paper focuses on the eye center localization problem. The new algorithm based on MB-LBP is described. The results of testing show that MB-LBP algorithm has better precision in large error range especially on the challenging BioID database and outperforms the simple Bayesian approach and algorithm based on image gradients. More than that, MB-LBP is almost 10 times faster than two others state-of-the-art approaches. Thus, MB-LBP can be a good decision for the speed-accuracy trade-off in eye localization.

In the structure of Emergency Control Ministry of Russia eye localization and face recognition technology is planned for use in video surveillance systems in the hardware-software complex "Safe City". This program based to introduce the streets of Russian cities, including the major business centers crowded with people, the newest automated surveillance, monitoring and control, designed for the needs of the emergency services.

The development of this research will help to solve the problem of creating additional opportunities to optimize the existing system of monitoring the state of public safety; create and develop a system of situational analysis of the causes of destabilization and forecasting of existing and potential hazards for public safety.

Fig. 8. Visual examples of results of eye center localization made by MB-LBP algorithm on FERET database

REFERENCES

[1] T. Riopka, T. Boult, "The eyes have it", *in Proceedings of ACM SIGMM Multimedia Biometrics Methods and Applications Workshop*, 2003, pp. 9–16.

[2] Z. Zhu, K. Fujimura, Q. Ji, "Real-time eye detection and tracking under various light conditions", *in Proceedings of the 2002 Symposium on Eye Tracking Research and Applications*, vol. 25, 2002, pp. 139–144.

[3] Z. Zhu, Q. Ji, "Robust real-time eye detection and tracking under variable lighting conditions and various face orientations", *Computer Vision and Image Understanding* 98 (1), 2005, pp. 124–154.

[4] P. Wang, M. Green, Q. Ji, J. Wayman, "Automatic eye detection and its validation", *in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, vol. 3, 2005, pp. 164–172.

[5] F. Song, X. Tan, S. Chen, Z.H. Zhou, "A literature survey on robust and efficient eye localization in real-life scenarios", *Pattern Recognition*, 46(12), 2013 pp. 3157-3173.

[6] The Yale Face Database, Web: http://cvc.yale.edu/projects/ yalefaces/yalefaces.html.

[7] Robotics Database, Web: http://robotics.csie.ncku.edu.tw/ Databases/.

[8] J. Marques, N.M. Orlans, A.T. Piszcz, "Effects of eye position on eigenface-based face recognition scoring", *Technical Paper of the MITRE Corporation*, October 2003.

[9] Z. Cao, Q. Yin, X. Tang, J. Sun, "Face recognition with learning-based descriptor", *in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2010, pp. 2707–2714.

[10] R. Javier, V. Rodrigo, C. Mauricio, "Recognition of faces in unconstrained environments: a comparative study", *European Association for Signal Processing Journal on Advances in Signal Processing*, 2009, pp. 1–19.

[11] S. Shan, Y. Chang, W. Gao, B. Cao, P. Yang, "Curse of mis-alignment in face recognition: problem and a novel mis-alignment learning solution", *in: Proceedings of IEEE Conference on Automatic Face and Gesture Recognition*, 2004, pp. 314–320.

[12] Ojala, T., Pietikäinen M., Harwood D. "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions", *12th IAPR International Conference on Pattern Recognition (ICPR)*, vol. 1, 1994. pp. 582 - 585.

[13] Petruk, V.I., Samorodov A.V., Spiridonov A.N., "Using of Local Binary Patterns in Face Recognition", *Bulletin of the Moscow State Technical University named after Bauman*, 2011. № S. pp. 58-63.

[14] L. Zhang, R.Chu, S. Xiang, S. Liao, S.Z. Li, "Face Detection Based on Multi-Block LBP Representation", *Advances in Biometrics, Lecture Notes in Computer Science*, vol. 642, 2007, pp. 11-18.

[15] D. Yi, Z. Lei, S.Z. Li. "A robust eye localization method for low quality face images", *International Joint Conference on Biometrics (IJCB)*, 2011, pp. 15-21.

[16] P.K.Suri, Amit Verma, "Robust Face Detection Using Circular Multi Block Local Binary Pattern and Integral Haar Features", *(IJACSA) International Journal of Advanced Computer Science and Applications, Special Issue on Artificial Intelligence*, 2011.

[17] Y. Martínez-Díaz, H. Méndez-Vázquez, Y. Plasencia-Calaña, E.B. García-Reyes, "Dissimilarity Representations Based on Multi-Block LBP for Face Detection, Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications", *Lecture Notes in Computer Science,* vol. 7441, 2012, pp. 106-113.

[18] The BioID database, Web: https://www.bioid.com/About/BioID-Face-Database.

[19] The FERET database, Web: http://www.itl.nist.gov/ iad/humanid/feret.

[20] M. Everingham, A. Zisserman, "Regression and classification approaches to eye localization in face images", *Automatic Face and Gesture Recognition*, 2006. FGR 2006.

[21] F. Timm, E. Barth, "Accurate eye centre localisation by means of gradients", *in Proceedings of the Int. Conference on Computer Theory and Applications (VISAPP)*, vol. 1, 2011, pp. 125-130.

[22] Z.-H. Zhou and X. Geng. "Projection functions for eye detection", *Pattern Recognition*, 37(5): 2004, pp. 1049–1056.