Michał OCHOCKI, Dariusz SAWICKI

Warsaw University of Technology, Institute of Theory of Electrical Engineering, Measurements and Information Systems

doi:10.15199/48.2016.11.11

Facial landmarks localization using binary pattern analysis

Abstract. *In the paper an algorithm for localization of the facial landmarks is presented. A preliminary analysis in the form of identifying regions of interest has been used. In this analysis each region corresponds to a different color space. The accuracy of the algorithm was verified using the Stimuli database of face images. The proposed algorithm can locate basic 21 landmarks with a very good average accuracy of 95.62%. This result ranks it among the best algorithms published between 2005 and 2015.*

Streszczenie*. W artykule przedstawiono algorytm lokalizacji punktów charakterystycznych twarzy. Zastosowano wstępną analizę w postaci wyodrębnienia obszarów zainteresowań. Przy czym dla różnych obszarów wykorzystano różne przestrzenie barw. Dokładność algorytmu zweryfikowano wykorzystując bazę Stimuli zdjęć twarzy. Opracowany algorytm pozwala lokalizować 21 podstawowych punktów z bardzo dobrą średnią skutecznością na poziomie 95.62%. Wynik ten plasuje zaproponowany algorytm wśród najlepszych rozwiązań opublikowanych w latach 2005-2015.* ***Algorytm lokalizacji punktów charakterystycznych twarzy***

Keywords: analysis of face geometry, facial landmarks localization, face texture, color spaces.
Słowa kluczowe: analiza geometrii twarzy, identyfikacja punktów charakterystycznych twarzy, tekstura twarzy, przestrzenie barw.

Introduction

A face detection and identification are the basic skills in a human interaction. Since childhood, we verify the identity and recognize emotions on the basis of interpretation of the face position, their shape and appearance. For people, the analysis of the face, is a basic and natural ability. However, automatic recognition of facial features is a big challenge for digital image analysis. This is an important task in many systems concerning face image processing [1, 2]. It emphasizes the importance of the problem.

Detection of characteristic points (landmarks) in a face image is a position analysis of the elements of a very complex object with difficult surface geometry. Despite research conducted in the world for over 20 years, this task is still relevant and is still one of the most difficult problems of image analysis and recognition.

The aim of the research was to develop an efficient algorithm for the analysis of the geometry of the face, which would allow identifying a large number of landmarks. Additionally, due to the end-use application of the algorithm, which is the recognition of emotions, it is assumed that the developed algorithm will operate on a dynamic set of characteristic points (landmarks). This means that on the one hand, individual group of points will be treated locally, and on the other hand algorithm will allow selecting and adding next points.

Methods for landmarks localization

Known algorithms used in detection of landmarks can be divided into two groups:

- Local texture-based methods [2–5].
- Holistic shape-based methods [6–12].

Taking into consideration our assumptions we decided to select the first of these groups.

In order to isolate a region of interest (ROI) from the image, the researchers use methods based on the features extraction from the image and methods based on the data selection by machine learning. Methods, which are based on the features extraction, include applications of: edge detection [13], color segmentation [14], an analysis of the features distribution or deformable meshes [15].

In his paper [3], Arca considered the location of the 24 landmarks of face using morphological image processing and methods of deformable meshes. Sohail [2] used the morphological operations, Gaussian and Laplace filtration and also thresholding operations. It allows locating 18 facial landmarks. Zhao [5] estimates the global position of 9 landmarks by the progressive elimination of the objects with incorrect shapes from the image. In his algorithm, Zhao uses the cascade classifier based on local features, working in the area around landmark. In first step, landmarks are extracted from the image by applying the Haar wavelets. Zhao's classifier at the learning stage uses AdaBoost and SVM. One of the latest reports is the work published by Zhang [4]. The author uses the binary features of the image, and machine learning during the classification stage. To classify the 11 landmarks, Zhang uses a cascade algorithm based on random forest method, which is a kind of the decision tree ensemble. The paper [15] is a very good survey of the methods related to face recognition and landmarks detection. But, of course, does not include the latest publications.

Selection of the ROI

The proposed algorithm for landmarks detection is based on the analysis of texture in previously separated regions of interest. We consider four separate ROI (Fig.1.): the facial region (as a whole – for the analysis of the oval), the eye region, the region of the nose and the mouth region.

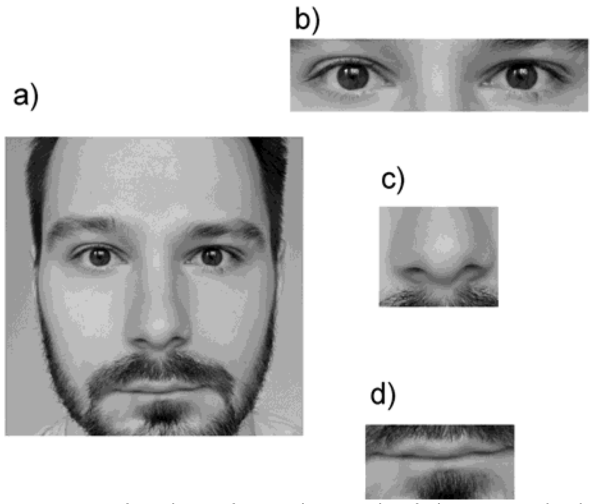

Fig.1. Image of author's face. The result of the ROI selection. Selected regions: a) the facial region (oval), b) the eye region, c) the region of the nose, d) the mouth region

Earlier determination of ROI has the task to decrease the size of the analyzed image. It leads to a reduction of calculations and allows reducing the possibility of incorrect location of the landmark. In the image processing we used the Viola-Jones face detector [16]. This detector ensures fast image processing with the high level of detection accuracy.

In the application of the algorithm Viola-Jones, we used four basic solutions:
- new image presentation called integral image,
- feature extraction using modified Haar function,
- AdaBoost learning method used at the stage of constructing the classifier,
- cascading structure of the classifier.

In the developed algorithm, we proposed the division of regions of interest into independent groups. For each of them, we used the selected color space in the analysis of the binary characteristics of the image.

- The facial region (oval), the eye region.

On the stage of highlighting the shape and boundaries of the face, we used the *YIQ* color space [17]. *YIQ* space in the NTSC system is a model for the television signal. Components of *YIQ* space can be obtained by a linear transformation of the *RGB* color space. Component Y is defined as luminance, I and Q determine the chrominance.

In our analysis, the image component Q is subjected to binarization by thresholding. Threshold of binarization is fixed for the entire image, and is determined for each image individually using the Otsu method [18]. This method is based on the histogram analysis of the grayscale image and is used to separate foreground objects from the background effectively. The result of the operation is an image, where the shape and borders of the oval face in binary images can be easily extracted. In addition, the same image can emphasize the eye region.

- The mouth region.

In order to highlight the mouth area, we have used the properties of *CIE-Lab* space [17]. It is a standardized color space *Lab*, which illustrates the difference between the colors seen through the human eye in the best way. The color is defined by three components L, a and b. Component L describes the luminance, a determines the color from green to magenta, b color from blue to yellow. In the algorithm for the shape of the mouth highlighting, we used two components a and b. After the transformation from *RGB* to *CIE-Lab*, the image component b is subtracted from the image component a. As the result of such operation we obtain an image with a highlighted region of color, which is characteristic for the lips color.

- The region of the nose.

In order to highlight the areas of the nose, we applied the method based on the analysis of the image in *RGB* space. In this case, we use all three *RGB* components of the image: R - red, G - green, B - blue.

At the initial stage of the conversion, based on the matrix $\boldsymbol{H}$ (1), two factors are determined: $h_{max} = max(\boldsymbol{H})$ and $h_{min} = min(\boldsymbol{H})$. These factors are used for image normalization.

$$\boldsymbol{H}(i,j) = \frac{\boldsymbol{R}(i,j)}{\boldsymbol{G}(i,j) + \boldsymbol{B}(i,j)} \tag{1}$$

where $\boldsymbol{R}(i,j)$, $\boldsymbol{G}(i,j)$, $\boldsymbol{B}(i,j)$ – a red, green, blue components of the image for the pixel (i,j) .

I the next step, we calculate the matrix of factors $\boldsymbol{H}_{temp}$ (2). It determines the normalized matrix of the image $\boldsymbol{H}_{norm}$ (3).

$$H_{temp}(i,j) = \frac{\boldsymbol{5} \cdot R(i,j)}{G(i,j) + B(i,j)} \tag{2}$$

$$H_{norm}(i,j) = \frac{H_{temp}(i,j) - h_{\boldsymbol{min}}}{h_{\boldsymbol{max}} - h_{\boldsymbol{min}}} \tag{3}$$

Finally, we obtain a normalized image, where the region of nose is highlighted. Additionally, normalized matrix $\boldsymbol{H}_{norm}$ helps to highlight elements of the eyes (iris) and eyebrows in the binary image.

Facial landmark localization in the ROI

The proposed detection algorithm allows extracting basic 21 facial landmarks (Fig. 2). The selection of points is typical for such algorithms and is associated with the most important elements of the face that are recognized by a man. Tasks of the location were carried out independently in each separate area – ROI. The similar localization process is applied to facial region, the eyes and mouth. Different process is applied separately for the region of the nose. In the analysis we used the binary parts of face image, extracted earlier.

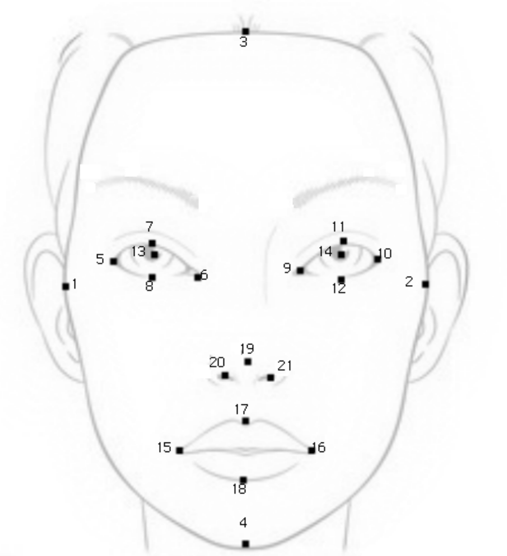

Fig.2. Location of the facial landmarks analyzed in this paper. The free face chart is used from [19]

- The facial region (oval), the eye and the mouth region.

For these regions, in the process of landmark detection, we just assumed that the shape of the face and eyes is oval. The process of localization is divided into following steps:
- A set of morphological operations, responsible for the isolation of appropriate set of pixels. We used here the analysis of neighborhood and cut off "protruding" pixels (artifacts). For the iris all the "holes" in the image are removed. This stage ends with a smoothing shape using a convex hull method.
- Shape analysis of the resulting set of points. To the resulting convex hull we try to fit the shape of ellipse in the best way. Based on this ellipse the landmarks points are selected.
- In addition, in the location of the iris we used the Hough transform, in order to detect objects whose shapes are closest to the circle. In this way, we assigned a circle to a set of pixels representing the iris.

- The region of the nose.

Landmarks location for the nose is based on the segmentation method, taking into account the brightness of the analyzed area. First, the input image of the nose is thresholded by Otsu method [18]. Based on the histogram, 8 levels of gray are determined. In the next step, we group pixels in to two areas: the darkest areas for the location of the nostrils of nose and lightest areas for location of the tip of nose. After such processing, the centroid of the collected pixels formed landmarks for the nose. These are: the center of the nose (among the brightest pixels) and nostrils (in the two groups of darkest pixels).

Tests of the proposed method of localization

The accuracy of the presented algorithm for facial landmarks location was verified using the database of facial images Stimuli [20]. This database was developed at the Center for Vital Longevity and is available free of charge for research conducted in academic centers. Stimuli base includes color photos of people's faces in frontal position of different race and age. All images in the database are prepared in jpg format and in resolution of 639 x 480 pixels.

Table 1. The results of the tests of algorithm for landmarks detection

Point number (Fig. 1)	Landmark	Correctness of the detection [%]	D_{eye}
6	The inner corner of the right eye.	96	0,0417
9	The inner corner of the left eye	95,2	0,0418
5	The outer corner of the right eye	95,2	0,0405
10	The outer corner of the left eye	96,8	0,0438
7	The upper point of the right eye	98,4	0,0448
11	The upper point of the left eye	98,4	0,0447
8	The lower point of the right eye	96,8	0,0467
12	The lower point of the left eye	96	0,0423
13	The center of the iris of the right eye	92	0,0464
14	The center of the iris of the left eye	95,2	0,0409
19	The center of the nose	91,2	0,0781
20	The right nostril	96,8	0,0411
21	The left nostril	96,8	0,0452
15	The right corner of the mouth	96,8	0,0547
16	The right corner of the mouth	96	0,0572
17	The upper point of the mouth	95,2	0,0537
18	The lower point of the mouth	92,8	0,0626
	The average value	**95,62**	**0,0486**

To evaluate the accuracy of landmarks location we have used a measure of detection error D_{eye} (4), defined as the normalized Euclidean distance.

$$D_{eye} = \frac{1}{N}\sum \frac{\sqrt{(x-\bar{x})^2 + (y-\bar{y})^2}}{d_{eye}} \tag{4}$$

where N – number of face images from the database, x , y – coordinates of landmark obtained by the automatic location (as a result of our algorithm), $\bar{x}$, $\bar{y}$ – real coordinates specified manually by authors, d_{eye} – scaling factor, which defines the distance between the pupils. This factor is used in most of the studies in order to standardize the size of faces in the pictures.

In order to verify the correctness of the algorithm we used the 17 landmarks. Measures of error D_{eye} for 4 landmarks of facial contour has not been taken into account. This is due to the large uncertainty in manual determination of the true coordinates for these points. Such a procedure is consistent with the evaluation methods described in the literature. The results of the tests are presented in Table 1. We determined the correctness of the detection for the landmarks and the average error detection D_{eye}. In straight majority of tests described in the literature, the error not greater than 0.1, is considered as the measure of the error for the correct location of the landmarks.

A comparison with other methods for detection of facial landmarks was performed. For comparison, we selected the most important methods described in the years 2005 - 2015. Summary data of this comparison is shown in Table 2.

Table 2. Set of basic parameters for selected methods for detection of facial landmarks

Author, year of publication	Method	Number of points	Correctness of the detection [%]	D_{eye}
Vukadinovic [6] 2005	shape-based	20	93	*
Arca [3] 2006	texture-based	24	85.8	*
Cristinacce [7] 2006	shape-based	17	92-95	**
Sohail [2] 2008	texture-based	18	90.44	*
Millborrow [8] 2008	shape-based	17	66	0.061
Kozakaya [9] 2010	shape-based	14	95.1	0.058
Valstar [10] 2010	shape-based	20	80-95	0.051
Amberg [11] 2011	shape-based	23	86-92	*
Dibeklioglu [12] 2012	shape-based	22	92.21	0.043
Zhao [5] 2013	texture-based	9	99	*
Hong [21] 2014	shape-based	17	88-98	0.049
Zhang [4] 2015	texture-based	11	90-96	0.052
Proposed algorithm	**texture-based**	**21**	**95.62 (91.2-98.4)**	**0.0486**

* other criterion
** no data available

In Table 2. the latest solutions of the problem of landmarks location are shown. These solutions represent both classes of algorithms – local texture-based and holistic shape-based. They have similar properties, and we cannot point out one group as definitely better or worse. However, differences in the applications, for which algorithms have been developed, are significant and cannot be neglected. Analyzing the data collected in Table 2. it can be stated that the proposed algorithm has very good properties. Correctness of the detection places it among the best solutions. It should also be emphasized the fact that the proposed solution allows detecting a large number of facial landmarks – 21 points. Among the solutions locating so many (or more) points, a new algorithm is one of the best.

Summary

In the paper new algorithm for the location of facial landmarks has been presented. The proposed algorithm can extract 21 points. This is a large number of points taking into account the methods published in the last few years. New algorithm has been tested for the correctness of operation and for the effectiveness. The accuracy of the presented algorithm of localization has been verified using the Stimuli base of face images. This is a base commonly used in testing for this type of algorithms. Tests were carried out on a set of 125 images from the database. Detection of each considered landmark has been analyzed independently. The proposed algorithm achieved an average correctness of the detection on the level of 95.62%, and the average location error on the level of 0.0486. The highest correctness (98.4%) algorithm reached in location of the upper point of the right eye. The lowest correctness (91.2%) algorithm reached in the center of the nose. Comparison with methods described in the literature was also carried out. The proposed algorithm has achieved a very good effectiveness that allows placing it among the best solutions published in the years 2005-2015. It is also worth to notice the simplicity of the operations used in the developed algorithm.

The initial separation of region of interest using different color spaces has brought very good results. It is worth noting that a good result which is locating 21 points practically represents a minimum for use in the emotions recognition. Work on the algorithm will be continued, primarily in order to increase the number of effectively recognized landmarks. However, effective locating further points could be more difficult.

***Authors**: Michał Ochocki Ms Eng., Dariusz Sawicki PhD DSc Eng., Warsaw University of Technology, Institute of Theory of Electrical Engineering, Measurements and Information Systems, Koszykowa 75, 00-661 Warsaw, Poland, e-mail: michal-ochocki@wp.pl, dariusz.sawicki@ee.pw.edu.pl*

REFERENCES

[1] Efraty B., Huang C., Shah S.K., Kakadiaris L.A., Facial Landmark Detection in Uncontrolled Conditions, *Proc. of Int. Joint Conf. on Biometrics*. (2011), 1-8

[2] Sohail A.S.Md., Bhattacharya P., Detection of Facial Feature Points Using Anthropometric Face Model. *Signal Processing for Image Enhancement and Multimedia Processing*, 31 (2008), 189-200

[3] Arca S., Campadelli P., Lanzarotti R., A face recognition system based on automatically determined facial fiducial points, *Pattern Recognition*., 39 (2006), n. 3, 432-443

[4] Zhang Z., Zhang W., Ding H., Liu J., Tang X., Hierarchical facial landmark localization via cascaded random binary patterns, *Pattern Recognition,* 48 (2015), n. 4, 1277-1288

[5] Zhao X., Shan S., Chai X., Chen X., Cascaded Shape Space Pruning for Robust Facial Landmark Detection, *Proc. of IEEE International Conference on Computer Vision (ICCV 2013),* (2013), 1033-1040

[6] Vukadinovic D., Pantic M., Fully automatic facial feature point detection using Gabor feature based boosted classifiers. *Proc of IEEE International Conference on Systems, Man and Cybernetics (CSMC)*, (2005), 1692-1698

[7] Cristinacce D., Cootes T., Feature detection and tracking with constrained local models, *Proc. of British Machine Vision Conference* (*BMVC 2006),* (2006), 95.1-95.10

[8] Milborrow S., Nicolls F., Locating Facial Features with an Extended Active Shape Model, *Proc. of 10th European Conference on Computer Vision*, (2008), 504-513

[9] Kozakaya T., Shibata T., Yuasa M., Yamaguchi O., Facial feature localization using weighted vector concentration approach, *Image and Vision Computing*, 28 (2010), n.5, 772-780

[10] Valstar M, Martinez B., Binefa X., Pantic M., Facial point detection using boosted regression and graph models, *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (2010), 2729-2736

[11] Amberg B., Vetter T., Optimal landmark detection using shape models and branch and bound, *Proc. of IEEE International Conference on Computer Vision (ICCV 2011)*, (2011), 455-462

[12] Dibeklioglu H., Salah A., Gevers T., A Statistical Method for 2-D Facial Landmarking, *IEEE Trans. Image Processing*, 21 (2012), n.2, 844-858

[13] Segundo M., Silva L., Bellon O., Queirolo C., Automatic face segmentation and facial landmark detection in range images, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 40 (2010), n.5, 1319-1330

[14] Wang J.G., Sung E., Frontal view face detection and facial feature extraction using color and morphological operations, *Pattern Recognition Letter,* 20 (1999), 1053-1068

[15] Hjelmas E., Low B.K., Face detection: A Survey, *Computer Vision and Image Understanding*, 83 (2001), 236-274

[16] Viola P., Jones M., Rapid Object Detection using a Boosted Cascade of Simple Features, *Computer Vision and Pattern Recognition (CVPR)*, 1 (2001), 511–518

[17] Ford A., Roberts A., Colour Space Conversions. Westminster University., London (1998)

[18] Otsu N., A threshold selection method from gray-level histogram, *IEEE Transactions on Systems, Man, and Cybernetics*, 9 (1979), n.1, 62-66

[19] Want a free mua face chart? http://smashinbeauty.com/want-a-free-mua-face-chart/ Retrieved 25 January 2013

[20] Minear M., Park D.C., A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, and Computers,* 36 (2004), 630–633

[21] Hong S., Khim S., Efficient facial landmark localization using spatial–contextual AdaBoost algorithm, *Journal of Visual Communication and Image Representation*, 25 (2014), n.6, 1366-1377