Emotion Recognition from Micro-Expressions: Search for the Face and Eyes

Anna D. Sergeeva, Alexander V. Savin, Victoria A. Sablina, and Olga V. Melnik
Department of Electronic Computers and Department of Information-Measuring and Biomedical Engineering
RSREU
Ryazan, Russia
sablina.v.a@evm.rsreu.ru

Abstract— **In the paper the problem of the emotion recognition from the micro-expressions is considered. The psychological background of this problem from the modern viewpoint is described in short. There are seven universal emotions which are usually could be recognized. The application fields of the emotion recognition are indicated. The most known databases for the facial micro-expression research are highlighted. The state-of-the-art of the micro-expression spotting and recognition methods is given. One of the main preliminary steps before the emotion recognition is investigated more thoroughly. It is the task of searching for the areas of a face and eyes in the image. Two algorithms implemented to solve this task are compared: the algorithm based on the color intensity estimation and the Viola – Jones algorithm. The results for of these algorithms for images of the people with the different skin colors are represented. The advantages and the drawbacks of the considered algorithms are discussed. The survey and experimental research results are proposed to further use in the emotion recognition pipeline.**

Keywords-emotion recognition; micro-expressions; face detection; eyes detection; skin color; Viola – Jones algorithm; facial expression datasets

I. INTRODUCTION

It is important to know what emotions a person really feels at present moment. Such a feedback from the interlocutor can be get using facial micro-expressions.

The most known specialist in the field of psychology of emotions Paul Ekman discovered that from the point of mime people from any culture express their feelings and emotions the same way [1]. Many investigations carried out in this field showed that a human could not control his micro-expressions and is not able to suppress them. Therefore, the one who can recognize them always can get the trustworthy information about the interlocutor state at present moment.

Usually seven universal emotions, which could be recognized using micro-expression analysis, are considered, i.e. anger, contempt, fear, disgust, happiness, sadness and surprise. Involuntary expressions are always coming out from the "mask" created on the face. In this case, they can be determined from micro-movements. Normally such expressions appear only for a small part of a second. That is why the task of micro-expression detection is complex also from the technical viewpoint since high-speed cameras are required for its solution.

To learn how to recognize the deception on the faces it is necessary to pay attention to the five aspects, i.e. the face morphology, the time characteristics of the emotion, the emotion manifestation area on the face, the micro-expressions, and the social context. People who control their face expression are paying the most attention to its lower part, i.e. a mouth, a nose, a chin and cheeks. After all, we perform the sound communication exactly through the mouth. However, the eyelids and the eyebrows more often reveal the true feeling.

Micro-expressions consist in the significant amount of information about the true human emotions. It can be used by employees of a customhouse, a police and a frontier station, by HR specialists, by journalists during the interview, by lecturers on the lessons, by businessmen in the meetings and by people who often have to search for the deception or verify the facts.

II. RELATED WORK

The facial micro-expression research started at the beginning of the twentieth century. From that moment many databases for the revelation and analysis of the facial micro-expressions are generated. The examples of the earliest datasets of the emotions which are not spontaneous are the Polikovsky's database, USF HD, and YorkDDT. But the emotions represented in these datasets differ significantly from the natural emotions. Because of that they are not very useful. The most known datasets with spontaneous emotions are CMIC, CASME, CASME II, CAS(ME)^2, SAMM.

Many algorithms are developed using emotion oriented databases. But as a result of the obstacles arising during the micro-expression analysis and the distribution of the micro-expressions into the determined classes the more attention is paying to the objective classes oriented to the movements of the facial muscles. SAMM is the only dataset oriented not to the emotional component but to the objective movements of the facial muscles [2]. Investigations carried out using the CMIC, CASME II and SAMM datasets show that the datasets oriented to the movements of the facial muscles give the better results than emotionally oriented.

The development of the machine learning technology leaded to the emergence of many algorithms of the micro-expression detection and analysis. The most widespread approaches to the micro-expression detection and analysis are based on the segmentation of the face into the regions, the

978-1-7281-1740-9/19/$31.00 ©2019 IEEE

feature extraction and the further search of the correspondences of the extracted features using the existing classificators. The Moilanen method is one of the first methods to detect the spontaneous micro-expressions in the long videos. This method uses as the original data the CMIC and CASME II datasets. The texture descriptor is used as the spatio-temporal characteristic. At the micro-expression classification stage the method uses the Support Vector Machine (SVM), the Multiple Kernel Learning (MKL) and the Random Forest (RF). This method recognizes approximately 72 % of the micro-expressions [3]. The method in [4] is based on the spatio-temporal characteristics determines the facial micro-expressions in the long videos. This method uses both macro-expressions and micro-expressions. The deformation magnitude of the facial muscles gives us the possibility of tracking not only the macro-expressions but also the instantaneous micro-expressions. This method recognizes approximately 74 % of the micro-expressions. The method in [5] is also based on the physical deformation of the facial muscles in the process of the emotion expression. The method is stable towards the rotation of the head and the change in the illumination. It is used for the description of the human face geometry. The deep learning methods are not often applied in the direction of the micro-expression analysis for the present. The main limitations of the deep learning methods are the long training time and the need for the use of the big dataset to train the system. In general methods of the automatic analysis of the facial micro-expressions can be classified into two groups: the micro-expression spotting and recognition methods [6].

The micro-expression spotting consists of the following stages: the preliminary processing, the feature extraction and the detection of the facial micro-expressions. The preliminary processing includes the steps of detection and tracking the facial landmarks, face registration, masking, and face region retrieval. Facial landmark detection can be done manually or automatically using the following techniques: Active Shape Model (ASM), Discriminative Response Maps Fitting (DRMF), Subspace Constrained Mean-Shifts (SCMS), and Constrained Local Model (CLM). Face registration is based on one of the two approaches: area-based or feature-based. Area-based methods use template matching or correlation. Feature-based methods use region features, line features and point features. In masking step the methods for masking facial regions, i.e. eyes, nose, mouth, are used. Face region retrieval is the step for splitting the face into several regions. In this step the facial segmentation methods are applied: segmentation of the face into regions, blocks, Delaunay triangulation. The feature extraction is utilized further at the stage of the detection of the facial micro-expressions. The methods of feature extractions for micro-expression spotting are Histograms of Oriented 3D Gradients (3D-HOG), Local Binary Patterns (LBP), and Histograms of Oriented Optical Flow (HOOF). The detection of the facial micro-expressions includes the movements spotting and the apex spotting. The movement spotting can be based on the methods of the feature difference analysis, e.g. Chi-Square distance. The apex spotting can be based on the methods searching for the apex frame, e.g. the binary search method.

The micro-expression recognition consists if the following stages: the preliminary processing, the feature extraction, and the classification. The preliminary processing in general can be performed by the same methods as it is described above for the micro-expression spotting. The methods of feature extractions for micro-expression recognition are based on Local Binary Patterns (LBP) or Optical Flow (OF). The classification can be done using the following methods: K-Nearest Neighbor (k-NN), Support Vector Machine (SVM), Random Forest (RF), Sparse Representation Classifier (SRC), Relaxed K-Singular Value Decomposition (RK-SVD), Group Sparse Learning (GSL), and Extreme Learning Machine (ELM).

III. Description of the Used Methods

One of the main preliminary steps required before the emotion recognition is the stage of search for the areas of a face and eyes in the image for the subsequent analysis of these parts. In this work we consider and compare two methods of search for the areas of a face and eyes.

The first one is the method of Viola – Jones [7] proposed in 2001 by Paul Viola and Michael Jones. In this method the Haar features are used to search for the desired object. In our case these objects are a face and eyes. This method at present is considered as the most optimal from the viewpoints of the execution time and the required computational resources [8].

All features enter to the classifier input. The classifier gives the result “true” or “false”. The classifier is trained slowly but the results of the search can be obtained fast. Different versions of the Viola – Jones method are used to search for the areas of a face and eyes in numerous research works [9, 10]. In this work we used the implementation of this method included in the OpenCV 4.0 library. Besides, pretrained configurations for the classifier which are the prepared beforehand and provided with the given library are used.

The second method is based on the color characteristic of the skin in the color space YCrCb and RGB. This approach is used in different research works [11-15].

This method consists in the pipeline of the several image processing and filtering stages of the original image. In general the algorithm for the given method can be represented as in Fig. 1.

It can be noticed that this method can be applied only to the color images. It is its main drawback.

IV. Experimental Results

To carry out the experiments the following computer configuration is used:

– Intel Core i5 2.3 GHz;
– 8 Gb DDR3 RAM;
– MacOS 10.14.3;
– OpenCV 4.0;
– Intel Iris Plus Graphics 640 1536 MB.

As the original images to carry out the experiments the open dataset IMM frontal face DB and also several personal images are used. The examples of the original images are represented in Fig. 2.

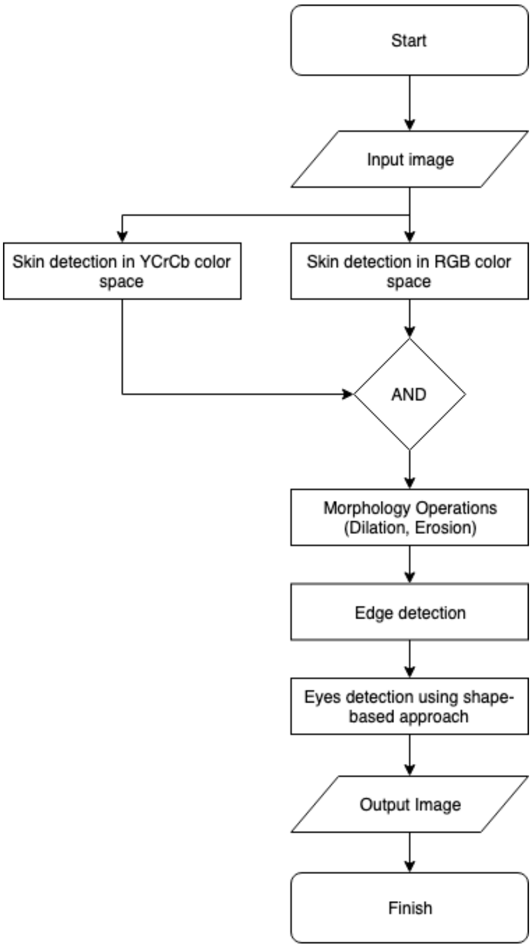


Figure 1. The general outline of the algorithm of the search for the area of the face and eyes.

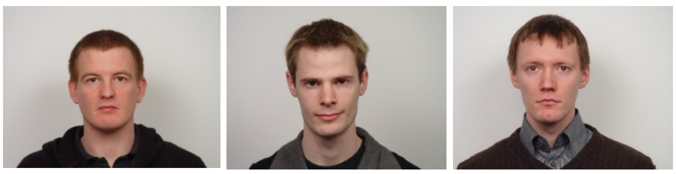


Figure 2. The examples of the original images.

The results of the considered algorithms are represented in Fig. 3 for face segmentation algorithm and in Fig. 4 for Viola-Jones algorithm.

Besides, these algorithms also can be used to process the images of the people with the dark skin color. The results of searching for the face and the eyes in such case are represented in Fig. 5.

As the main parameter the execution time of the algorithms is estimated. Besides, the mean-square deviation σ is also calculated. To do this 10 experiments are carried out for each image. The comparison of the investigated algorithms is represented in Table 1.

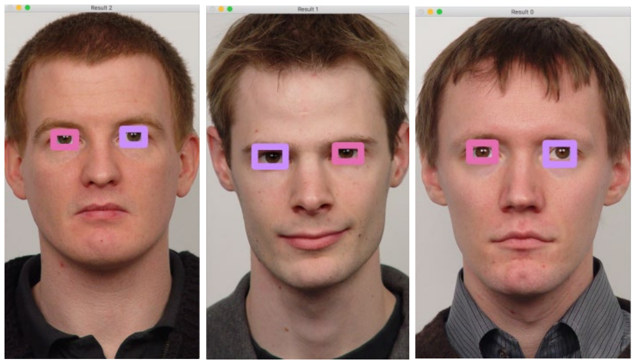


Figure 3. The results of the algorithm based on the color intensity estimation.

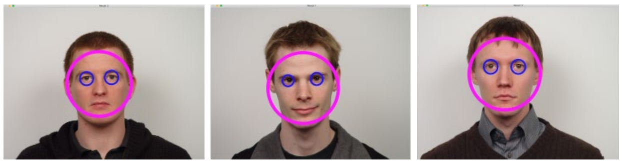
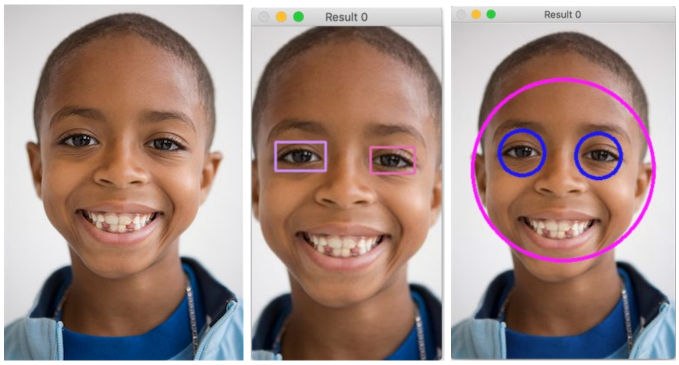


Figure 4. The results of the Viola – Jones algorithm.



a b c

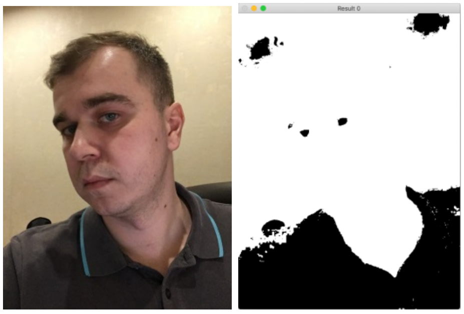
Figure 5. The results of searching for the face and the eyes for the people with the dark skin color: a – the original image; b – for the algorithm based on the color intensity estimation; c – for the Viola – Jones algorithm.

TABLE I. COMPARISON CHARACTERISTICS OF THE ALGORITHMS

Resolution	**Color-based approach**		**Viola-Jones**	
	$\bar{t}$, *ms*	σ, *ms*	$\bar{t}$, *ms*	σ, *ms*
664×1000	63.57	3.08	190.41	11.06
810×1071	88.48	7.15	227.81	33.36
2560×1920	466.78	2.78	437.19	20.33
2560×1920	469.28	4.34	735.16	15.88
2560×1920	469.19	2.02	461.19	34.02
2560×1920	464.33	5.04	356.91	19.27
2560×1920	485.89	9.31	420.29	6.82

We can see from the obtained experimental results that the method based on the color intensity estimation wins on the execution time at smaller image sizes. In turn, Viola – Jones method works faster in some runs when image sizes reach 2560×1920 pixels. Besides, we should highlight the limitations of the methods. As it is noticed above, the method based on the color intensity estimation works only with color images while the Viola – Jones method works both with the color and non-color images.

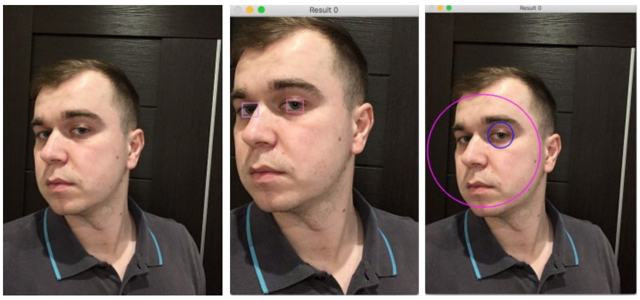
Also from the experiments we determined that the image background plays a major role in the method based on the color intensity estimation. If the background will be of the same color as the face color then this method gives erroneous results. The original image with the light background similar to the skin color and the mask of the skin detection are represented in Fig. 6.



a b

Figure 6. An error in face skin detection obtained using the algorithm based on the color intensity estimation: a – the original image; b – the mask of the face.

Also the experiments with the images in which the person doesn't look directly to the camera are carried out. In this case the method based on the color intensity estimation determines the face and both eyes while the Viola – Jones method determines the face and only one eye (Fig. 7).



a b c

Figure 7. An error of the Viola – Jones algorithm: a – the original image; b – the result of the algorithm based on the color intensity estimation; c – the result of the Viola – Jones algorithm.

V. CONCLUSION

The obtained results show that both considered algorithms could be useful to perform the preliminary processing in the pipeline of the micro-expression spotting or recognition. The algorithm based on the color intensity estimation is suitable for color images and faster for the small images. The Viola – Jones algorithm can be also applied to grayscale images and faster for images of a higher resolution. Experiments approved that considered algorithms provide sufficient results for the tasks of determining the face and eyes areas in the image. In the further research we intend to investigate the whole pipeline to detect and recognize the human micro-expressions, thereby to develop the more precise emotion recognition methods.

REFERENCES

[1] Paul Ekman, Emotion in the Human Face, 2nd Edition, Malor Books, 2013, 456 p.

[2] Adrian K. Davison, Cliff Lansley, Nicholas Costen, Kevin Tan, Moi Hoon Yap, "SAMM: A Spontaneous Micro-Facial Movement Dataset," in IEEE Transactions on Affective Computing, Volume 9, No. 1, 2018, pp. 116-129.

[3] Antti Moilanen, Guoying Zhao, and Matti Pietikäinen, "Spotting Rapid Facial Movements from Videos Using Appearance-Based Feature Difference Analysis," Proceedings of the 2014 22nd International Conference on Pattern Recognition (ICPR), Washington, DC, USA, 2014, pp. 1722-1727.

[4] Matthew Shreve, Sridhar Godavarthy, Dmitry B. Goldgof, and Sudeep Sarkar, "Macro- and Micro-Expression Spotting in Long Videos Using Spatio-Temporal Strain," Face and Gesture 2011, Santa Barbara, CA, 2011, pp. 51-56.

[5] Zhaoqiang Xia, Xiaoyi Feng, Jinye Peng, Xianlin Peng, and Guoying Zhao, "Spontaneous Micro-Expression Spotting via Geometric Deformation Modeling," Computer Vision and Image Understanding Journal, Volume 147, 2016, pp. 87-94.

[6] Yee-Hui Oh, John See, Anh Cat Le Ngo, Raphael C.-W. Phan, and Vishnu M. Baskaran, "A Survey of Automatic Facial Micro-Expression Analysis: Databases, Methods, and Challenges," Frontiers in Psychology Journal, Volume 9, Article 1128, 2018, 21 p.

[7] Paul Viola and Michael J. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Kauai, HI, 2001, pp. 511-518.

[8] Elena G. Kuznetsova, Evgeny A. Shvets, Dmitry P. Nikolaev, "Viola-Jones Based Hybrid Framework for Real-Time Object Detection in Multispectral Images," Proceedings SPIE, Eighth International Conference on Machine Vision (ICMV 2015), Volume 9875, 98750N, 2015, 6 p.

[9] Kunhui Lin, Jiyong Huang, Jiawei Chen, and Changle Zhou, "Real-time Eye Detection in Video Streams," Proceedings of the Fourth International Conference on Natural Computation, Jinan, 2008, pp. 193-197.

[10] Aleksey V. Alekseev, Yuliya A. Orlova, Vladimir L. Rozaliev, "Automatic Search of Human Face and Distinguish the Contours of the Mouth, Eyes on the Portrait Images," Izvestia VSTU Journal, No. 2 (157), 2015, pp. 49-53 (in Russian).

[11] Jalal A. Nasiri, Sara Khanchi, and Hamid R. Pourreza, "Eye Detection Algorithm on Facial Color Images," 2008 Second Asia International Conference on Modelling & Simulation (AMS), Kuala Lumpur, 2008, pp. 344-349.

[12] Alireza R. Azar and Farhad Khalilzadeh, "Real Time Eye Detection Using Edge Detection and Euclidean Distance," 2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI), Tehran, 2015, pp. 43-48.

[13] Douglas Chai and King N. Ngan, "Face Segmentation Using Skin-Color Map in Videophone Applications," in IEEE Transactions on Circuits and Systems for Video Technology, Volume 9, No. 4, 1999, pp. 551-564.

[14] Mohammad R.R. Fini, Mohammad A.A. Kashani, and Mohammad Rahmati, "Eye Detection and Tracking in Image with Complex Background," 2011 3rd International Conference on Electronics Computer Technology, Kanyakumari, 2011, pp. 57-61.

[15] Nusirwan Anwar bin Abdul Rahman, Kit Chong, and John See. "RGB-H-CbCr Skin Colour Model for Human Face Detection," MMU International Symposium on Information & Communications Technologies (M2USIC), 2006, 6 p.