Algorithm of Preparation of Training Sample Using 3d-Modelling of Faces

Frolov I., Samal D.

Belarusian State University of Informatics and Radioelectronics, 6 P. Brovki Street, Minsk 220013 Belarus, samal@bsuir.by

Abstract**:** *The algorithm to preprocess the training sample of multi-class classifier* support vector machine *(SVM) is described in the paper. This provided approach is based on the modeling of possible changes of the face features of human recognition, which influence on training classifier and respectively on further results of recognition. Age, emotional face expressions, head rotation, different lighting conditions, noise, and some combinations of mentioned parameters were selected as the major parameters for modeling.*

1. INTRODUTION

Detection algorithms of face area on image focused on robustness of the low quality of input images – contrast, brightness, etc. However, process of valid classification is very problematic even in case of accurate face detection because of a difference of lighting (fig. 1) when carrying out photographing.

Fig.1 - Exapmles of images with different light conditions.

It should be noted, that nowadays there are no universal algorithms to improve images and to adjust the bright-contrast characteristics, used for different image processing tasks and, particularly, in face recognition. Therefore, there is an open question about selection of existing or needed to develop algorithms of digital image processing for face recognition.

In system of biometric face identification in Technest company (www.genextech.com) 3D-modeling technology was used to add missing foreshortenings with different poses, lighting and face expression to faces from database, however that site does not report any leveling of age difference.

The paper [1] overviews possibility of using filters, noise in order to extend training set for setting of biometrical system recognition. However, that work was focused on noise modeling more and did not consider questions of lighting modeling, gesture and age difference.

This investigation overviews complex approach to solve tasks of generation and extension of training sample from images with different brightness characteristics, contrast, lighting conditions, emotional face expressions, age difference. This result is reached by combining of 3D-modeling face with changeable parameters and methods of digital image processing for noise imitation, also adapting to single processing conditions.

SVM-classifiers, as well as neural networks, refers to so-called static classifiers, therefore, the more number of face images in different variations (lighting, gesture, age, rotation, angle) will be introduced to system during training, the more this classifier will have generalizing ability, and the higher percent of recognition will be at the step of face recognition (that statement will be checked at the step of algorithm testing).

2. PARAMETERS OF FACE 3D-MODELLING

To construct the 3D-model face the system requires at least one photo portrait of full frontal type with resolution, not lower than the required (not lower than 90 pixels between centers of pupils of the eyes) and without deflection, rotation and head inclines. 3D-model construction is performed with using third-party software FaceGenhttp://www.facegen.com.

High degree of the photo conformity of real people and their 3D-models were determined with the help of 3D-modeling technologies (fig. 2).

Fig.2 - Matching photo and 3D-model.

Construction of a virtual set of faces with different camera angles, emotions, lighting, noise, age allows to extend training sample, also to adjust influence of deforming factors. The most popular fact, influenced on recognition accuracy, is changing lighting sources, which cause appearing shadows on images. If training sample does not contain images, made by different lighting variations, level of recognition by system such faces will decrease rapidly. Fig. 3 shows variations of face model lighting (from left to right): above - diffused straight lighting, light source on the right, light source on theleft.

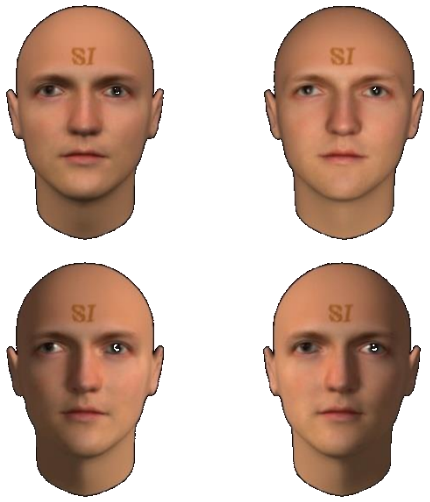

Fig.3 - Modeling different conditions of lighting.

Age plays main role in personal identification by digital photo portrait, while receiving photos and during identification process. In that case, human face is subjected by aging and, therefore, face features are changing (fig. 4).

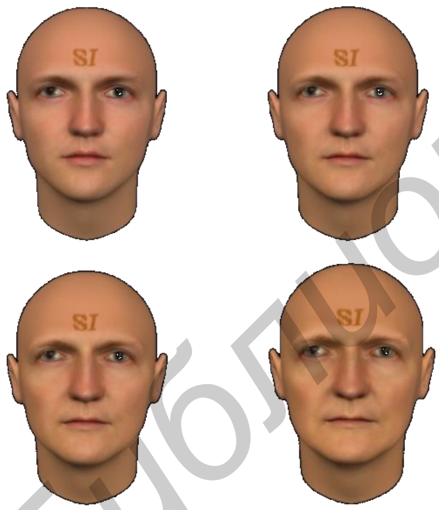

Fig.4 - Model of age changing from left to right: 30 years, 40 years, 50 years, 60 years.

Besides, angle of photographing is important aspect of recognition: head deflection from a vertical position to right or left side, head rotation around vertical axis to right or left side, head bend around vertical axis up-down. These changing head positions can be set in one plane (fig. 5), and at the same time in all mentioned planes (fig 5).

Fig. 5 - Head inclines and rotation: from above – in three planes, from down – in one plane in uniform diffuse lighting.

The next factor, which changes facial features, is expression of emotions (fig. 6).

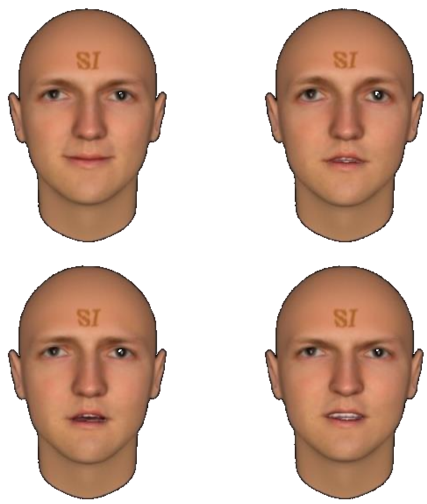

Fig.6 - Emotions: closed smile, surprise, fear, anger.

Besides, in order to construct complete training sample of numerous lighting variants, head inclines/rotations, emotional face expression, combinations of mentioned variants were implemented for each photo portrait. Paying attention to factor of different conditions of receiving, storage and transferring images, conjunction of faces 3D-models and several popular kinds of noise was made for leveling different possible noise on images (fig. 7).

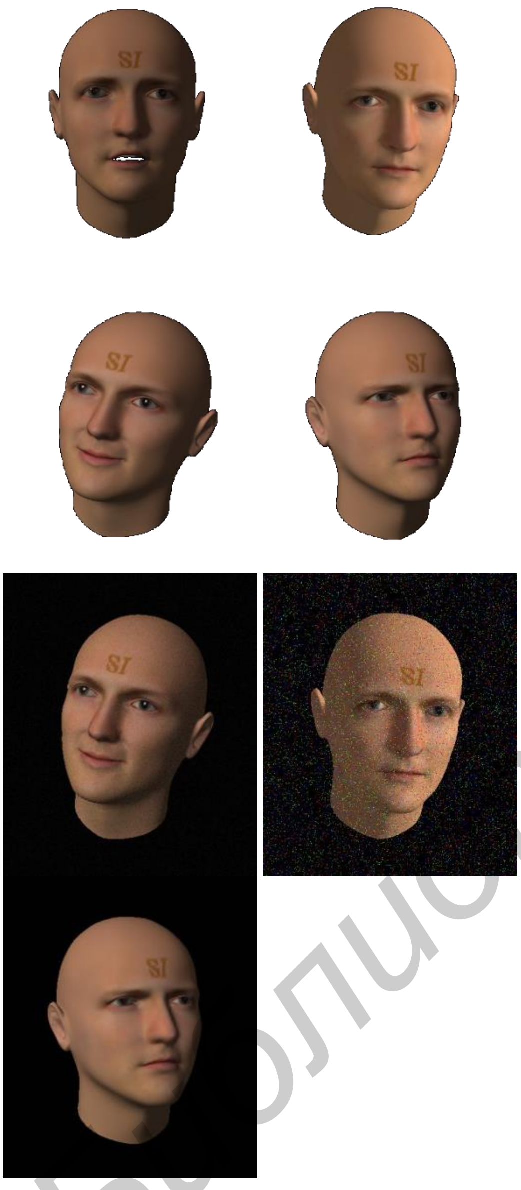

Fig.7 - Modeling variations of lighting, emotions, age, noise.

An applying of those approaches allows extending set of photo portraits in process of forming and processing training sample, receiving more numbers of possible variations of photo conditions. Therefore, number of characteristic vectors describing images increases, and respectively, considering their difference, number of support vector increases while training classifiers and it increases the generalizing ability of classifiers and, therefore, tend to increase recognition coefficient.

Learning limited set of changing parameters can help to calculate common number of additional images, which are built for extending training sample according to a formula (1).

$$images = \frac{n*(n-1)}{2} \qquad (1)$$

where images – common number of addition images for each initial photo portrait, n – number of parameters variation.

17 parameters were used in that research for extending training sample – minimum set of variations, allowing to capture several important factors simultaneously: rotate angle and incline (7 positions), age (2 positions), emotional face expression (4 positions), lighting (4 positions), that allowed to increase training sample in 136 times (because for each image was generated exactly 136 additional images). In order not to make a significant part of the noisy images, noise effects and filters were applied selectively only for face images with normal dispelled lighting without additional modified parameters, which allows to form noise effect for 40 images of each(randomly selected) class (fig. 7). The following modifications were used in work: gauss noise, impulse noise (like «salt and pepper»). Therefore, common number of images, generated for each initial image, was 176 examples.

That sequence of operations in forming of training sample can be described as following algorithm.

Step 1. Forming initial sample of face images for training SVM-classifier.

Step 2.3D-model building of face processing with the help of third-party software FaceGen (www.facegen.com) and generating additional 176 images (quantity and kinds of parameters – lighting, age, emotional face expression, mustache, glasses, inclines/rotations – are set depending on using system conditions and user requirement) for each initial photo portrait in according to selected modified parameters of photography.

Step 3. Estimating of camera angle for each image (initial and generated): preliminary face search on image, search of pupils, calculating rotation angle of image and rotating to the eyes layout on the same horizontal line, calculating rotating angle α_1^{new} and head incline α_2^{new}.

Step 4. Localization of face area on the borders of the eyebrows with the help of localization algorithm of Viola-Jones face.

Step 5. Preliminary processing and image normalization: extension of range of the brightness of pixels and histogram equalization for alignment of bright-contrast characteristics of image processing.

Step 6. Scaling of localized and processed face area to size 200×200 pixels (that extension complies with minimum distance in 90 pixels between pupils of eyes centers in face localization on the borders of the eyebrows).

Step 7. Adding localized and processed face area to training set of photo portraits and appropriate recognition system of face database.

Step 8. Forming characteristic vectors of scaled and processed face images in according to principal component analysis(NIPALS algorithm), using 100 main components for introduction of each vector. Saving of characteristic vectors in training set and appropriate base system.

Step 9. Multi-class training of SVM-classifier (4). End of the algorithm.

While developing of system researches were made and experimentally expediency of implementation was founded on the first work operation step of selecting interest area – face – on the initial image, and on following step – implementation of preliminary procedure of digital image. Efficiency of performing those procedures in specified order showed on fig. 8. Images on the right side have more contrast and deeper view, it means, they contain more informative characters than monotone and low contrast image on the left side. Those results are explained by background influence and «extra» parts of photo portrait (background, clothes) while applying equalization algorithm and histogram normalization and smoothing filter [3] of image, because while using them results are calculated with considering of pixels of whole image, background from that image sometimes is darker, or lighter than face area.

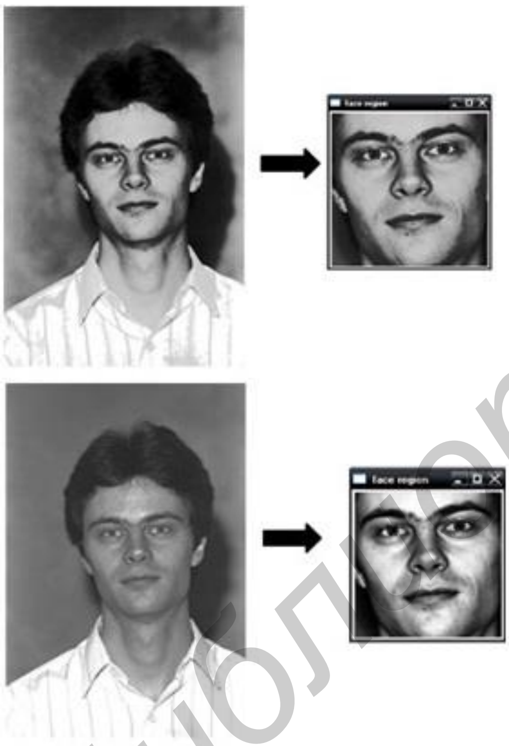

Fig.8 - Procedure example of image preprocessing

Performing of improving image procedure only on specified face area allows getting more contrast images, more suitable for following procedure of initial data reducing and example recognition.

Experiment was made to assess the efficiency of applying suggested approach in forming and extending training sample for training SVM-classifier. For recognition SVM-classifier was trained with training parameters C = 8, γ = 0,003125. The results of experiment are showed in table 1.

Table 1. Influence of extending volume of training sample on recognition coefficient

Following sample	Number of classes	Common number of images for training	Number of each class images	Recognition coefficient in identification, %
2 images of each class	500	1000	2	84,8
2 images of each class + additional virtual faces set	500	89000	178	90,6

3. CONCLUSION

Findings confirm expediency of applying developed forming algorithm of training sample for training classifier for face recognition. Applying of face 3D-modeling allows to compensate lack of real photo portrait number with different photography conditions, age difference and emotional face expressions, noises. Developed algorithm differs stability not only to lighting changes, camera angle, emotions, but also to age difference and noise, from famous approaches in systems of biometrical identification.

4. REFERENCES

[1]Starovoytov, V.V. Extansion of training set for setting biometrical recognition systems / V.V. Starovoytov, Nikodimov D.Y. // Data processing and management in extraordinary and extreme situations: materials of the 4-th international conference. – Minsk, 2004. – P. 204–209.

[2]Yanushkevich, S. N. Biometric-Based Decision Support Assistance in Physical Access Control Systems / S. Chague, B. Droit, O. Boulanov, S. N. Yanushkevich, V. P. Shmerko, A. Stoica // Bio-inspired Learning and Intelligent Systems for Security : Proceedings of ECSIS Symposium, Aug. 4–6, 2008, Edinburgh, UK. – Edinburg, 2008. – P. 11–16.

[3]Gonzalez, R. Digital image processing / R. Gonzalez, R. Woods. – Moscow: Technosphere. – 2005. – 1072 p.

[4]Haykin, S. Neural networks: full course / S. Haykin. – Moscow: Williams. – 2006. – 1104 p.