SOLVING THE TASK OF FACE RECOGNITION IN CASES OF INSUFFICIENT TRAINING SET

Olga Krutikova and Aleksandrs Glazs
Riga Technical University
Setas street 1, Riga, Latvia, LV-1084

ABSTRACT

This paper describes methods that are aimed to solve face recognition tasks with an insufficient training set. These methods include: creation of a 3D model of a head that is based on a basic training set - three images of faces (profile, half turn, full face), placing and analyzing control points on a model, calculating distances between points (the ones not used in the creation of a model), which is followed by face recognition. The created 3D model allows acquiring additional images of faces (at different angles), which significantly increases the results of recognition of unknown faces, as compared to only using the basic training set. The proposed methods were tested on various images of faces. The results have shown that these recognition methods can be used in cases, when the initial information about the shape of the face is insufficient, for example, in forensics.

KEYWORDS

Face recognition, 3D model, insufficient training set, control points.

1. INTRODUCTION

The problem of face recognition in images is an important aspect of computer vision. Faces are similar in their anatomical structure, but they differ significantly in shape. Solving this problem is important for various fields, especially in cases, where initial information is limited.

A large amount of face recognition 2D algorithms have been proposed. For example, in (Turk and Pentland, 1991) it was proposed to use eigenfaces to find and recognize faces in images. Face recognition techniques using linear/Fisher discriminant analysis (LDA) (Zhao, 1998) or support vector machines (Heisele, 2001) for the classification were also developed. The main disadvantage of the proposed methods is that they are very dependent on the turn angle of the head, lighting, and they also require a large training set.

Using a 3D model of a face/head allows to solve the main problem of 2D recognition: information about the shape of the face at different turn angles of the head. In (Zhao, 2000) a method was proposed that preprocessed 2D images to create a 3D model. This shape-from-shading (SFS)-based method used a depth map to generate synthetic frontal images. The Linear Discriminant Analysis (LDA) was applied to the synthetic images instead of the original images. A disadvantage of this method is the difficulty of calculating a mathematical model of lighting, it is also not always possible to accurately recreate the shape of the face, and the model might contain distortions and holes.

In (Huang, 2001), (Blanz, 2002), (Blanz and Romdhani, 2003) it was proposed to use a 3D morphable model. The main disadvantage of these methods is that they use a 3D scanner to create models, which is impossible if the initial data are simple 2D images.

In (Bronstein, A. M., Bronstein, M. M. and Kimmel, 2003) it was proposed to use photometric or stereo light to acquire information about the 3D structure. The range image and the texture of the face are acquired. Next, the range image is preprocessed by removing certain parts such as hair, which can complicate the recognition process. A canonical form of the facial surface is computed. The recognition itself is performed on the canonical surfaces. The disadvantage of this method is that several images of the face are taken in different lighting conditions and the 3D geometry is extracted by assuming a Lambertian reflection model. Also, the images are not always available in different lighting conditions and there are problems with calculating a reflection model.

Another method (Popatheodorou, 2004) uses Principal Component Analysis (PCA), which is widely used for 2D recognition and was modified for 3D recognition. They propose to extend the capabilities of the algorithm and combine color, depth map and color map (Tsalakanidou, 2003). However, the main disadvantages of the PCA method are strict requirements for the images of faces (distinct background, size, rotation of the head up to 30°). This approach also cannot be used if the training sample contains only several images.

In (Godil, Ressler, Grother, 2004), a scanned 3D models of a head are used to increase the training set. Stickers or anthropometry landmarks are placed on the person, before scanning. The seventy-three Anthropometry Landmarks were extracted from the scans. The performance of the system based on the 3D information was evaluated by comparing it to the one based on color map information using a PCA based method. The disadvantages of this method is that it uses scanners for creating 3D models and the model topology was rough in the face region. Also, the recognition algorithm requires a color map, which is unacceptable, if the images that are used for creating the models and are synthesized with the help of models are halftone.

In (Howland and Wang, 2006), it was proposed to use the PCA and then LDA methods, but their algorithm was tested on the images of faces, where the turn angle of the head varied only in the range of -30° to 30°.

In this paper, it is proposed to use a semi-automatic method for the creation of the 3D model of the head from 3 images of faces (Krutikova, 2013, 2014), in which an average 3D model of a head is adapted to a specific face by using control point that are placed on the images of the face, which allows to avoid using any additional equipment (3D scanner, stereo cameras) in creation of the models.

There are also many manual, semi-automatic and automatic face recognition algorithms (Scheenstra, 2005). Since the recognition process is complex and consists of many stages, the manual algorithms are too slow and are susceptible to operator error. In large data bases, the search for faces can take a considerable amount of time, which is not always acceptable. Automatic method requires minor interventions from the operator, but it is not always possible (lighting, turn of the head), since the control points may not be properly detected and would require additional configuration.

In this paper, a method of face recognition is proposed, which uses information that was acquired with the help of a 3D model of the head (Krutikova, 2013, 2014). Points that are placed on the base images, are transferred on to the model of the head, which allows to calculate the distances between the points at different turn angles of the head. Control points are also placed on the new image of the face and the face recognition is based on the calculated distances between the points.

2. PROPOSED METHODS

2.1 Method of Face Recognition Based on the Control Points

The proposed method allows recognizing a new image of a face based on the data that was acquired from the 3D model of the head, which is created for each class based on the images from the insufficient training set. To create the model, control points are placed on the images from the training set. The distances between the control points of each model of the head and the control points on the new image are then calculated. The sums of squared distances are calculated. The smallest sum of squared distances corresponds to the found face. The proposed method consists of several steps:

1. Creation of the 3D model of the head for each class using the training set (3 images – profile, half turn, full face,) (Krutikova, 2013, 2014). The recognition does not need textured models, it only uses the mesh, which contains all the necessary vertices of the model, which allows to reduce the consumption of resources and make loading and reading information from the model faster. For example, loading a model without textures takes (0.372 s), with textures (1.32 s).

The images from the insufficient training set with placed control points can be seen on Figure 1.1. The 3D model of the head (Figure 1.3) was created, using an 3D model of an average head (Figure 1.2), by adapting its vertices according to the placed control points.

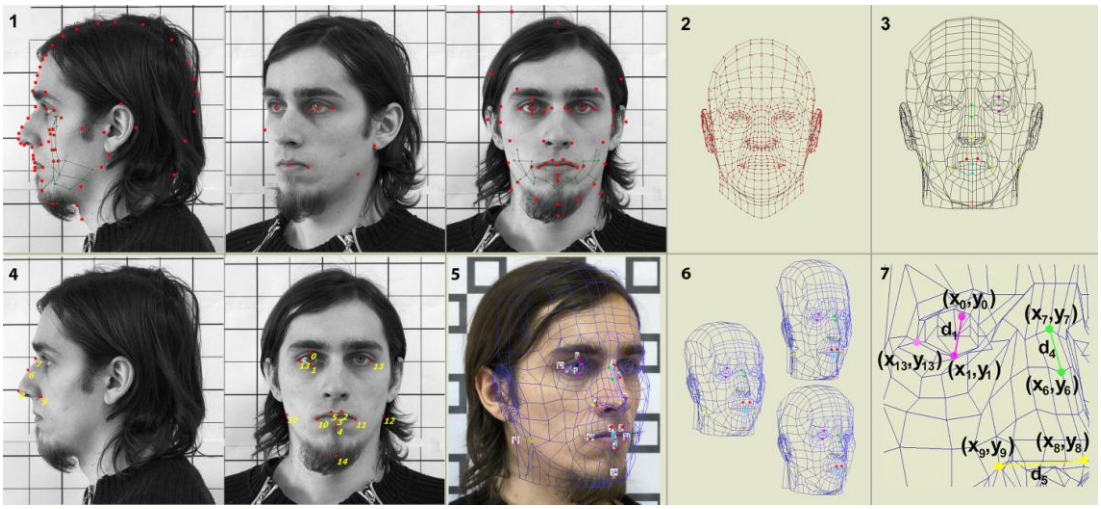

Figure 1. Creation of the 3D model of the head

2. The control points (Figure 1.4), which are used to calculate distances, are automatically transferred from the training set images on to the 3D model of the head and are marked in different colors (Figure 1.3, Figure 1.7).

3. Control points are placed on the new image of the face from the examination set (Figure 1.5).

4. Each 3D model *j* of the head is automatically positioned on the image of the face (if needed the model is scaled) and the model *j* is rotated to most resemble the rotation angle of the head on the image (or its closest angle) (see Figure 1.5, Figure 1.6).

5. The distances between the control points on all the head model images and the control points on the new face image are calculated.

5.1. To calculate distances between the control points on the 3D model, all control points are marked in pairs with different colors.

5.2 The image is scanned horizontally *w* and vertically *h*, until the intensity of red I_r, green I_g, and blue I_b channels is the same as the desired color.

5.3 The Euclidian distances *k*, $k \in [1:N]$ between the control points *p+1* and *p*, for the image of the model *j* (Figure 2.6) can be determined as follows (see Formula 1) :

$$d_{j_k} = \begin{cases} \frac{1}{h}\left(y(p+1)_{k,j} - y(p)_{k,j}\right), k = 1,2 \\ \frac{1}{h}\sqrt{\left(x(p+1)_{k,j} - x(p)_{k,j}\right)^2 + \left(y(p+1)_{k,j} - y(p)_{k,j}\right)^2}, k \in [3:N] \end{cases} \quad (1)$$

The Euclidian distances for a new image of the face (Figure 1.7) can be determined similarly (see Formula 2):

$$d_{e_k} = \begin{cases} \frac{1}{h}\left(y(p+1)_{k,e} - y(p)_{k,e}\right), k = 1,2 \\ \frac{1}{H}\sqrt{\left(x(p+1)_{k,e} - x(p)_{k,e}\right)^2 + \left(y(p+1)_{k,e} - y(p)_{k,e}\right)^2}, k \in [3:N] \end{cases} \quad (2)$$

where $j \in [1:m]$ – model number, *m* – amount of models;
$k \in [1:N]$ – distance number, *N* – amount of distances;
$p \in [1:Q]$ – control point index, *Q* – amount of control points;
d_{j_k} – distance (with number *k*) between the control points of the model (*j*), pixels;
d_{e_k} – distance (with number *k*) between the control points of the new image (*e*), pixels;
x_{j_p}, y_{j_p} ,– coordinates of the control point *(p)* of the model *(j)*, pixels;

x_{e_p} , y_{j_p} – coordinates of the control point *(p)* of the new image *(e)*, pixels;

h - height of the image of the model, pixels.
H - height of the new image of the face, pixels.

6. Sum of squares is calculated between certain distance numbers for the new image *e* and the image *j* that was acquired from the 3D model (see Formula 3).

$$d_j = \sum_{k=1}^{N} (d_{j_k} - d_{e_k})^2 \tag{3}$$

where *j*- class number, $j \in [1..m]$
d_j -sum of squares *(j)*

8. To recognize the face now it is necessary to solve the optimization task - find the minimal sum. The minimal sum of squares of distances is calculated and it is determined, whether the face belongs to any class *j**.

$$d_j \rightarrow \min_{j \in [1..m]} \Rightarrow j^*$$

3. EXPERIMENTS

The proposed method was tested on images from twelve classes (12 people) with 3 images (Figure 3. a) profile, b) half turn, c) profile) in each class that are used to create 3D models (Figure 3. d), e), f)). Images that were used to create the models and test the algorithm were taken using as -lens reflex digital camera Nikon D5200 24.1 MP CMOS Digital SLR with AF 70-300mm F/4-5.6 Tamron Tele-macro lens, with resolution 6000x4000 and stored using the .JPG format, and for model creation the image resolution was reduced to 1702x1200. For the examination set, the image resolution was further reduced to 762x861. The average time to create models (without placing the control points) -30 sec.

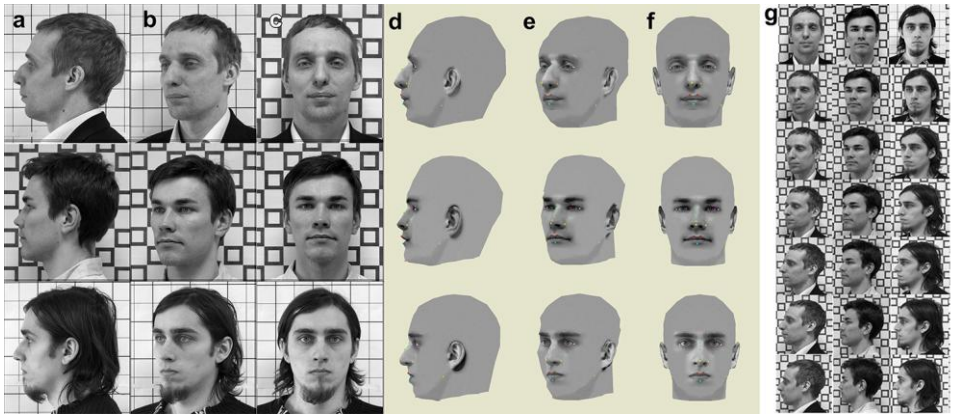

Figure 2. An example of initital data for 3 clases: a), b), c) training set, d), e), f) 3D models of each class, g) examination set

The proposed method was tested (Dell Inspiron 5710, Intel(R) Core(TM) i5-3210M CPU (2*2.5 GHz), Video card - NVIDIA GeForce GT 630M), using an examination set, which consisted of 84 images with various turn angles of the head that varied from -90 to 0 degrees with an interval of 15 degrees.

Control points were placed on each new image of the face and the distances were calculated, the sum of squares and minimal values were also calculated. The algorithm was tested using the extended training set (the model was used), and the base training set (without the model). When the algorithm was tested using the extended training set there were no mistakes. Table 1. shows the test result for the 1st experiment, where all calculated minimal values belong to the correct class.

Table 1. An example of recognition results for first image, when the extended training set was used

New image (e)	Model (j)	Rotation angle, degree						
		-90°	-75°	-60°	-45°	-30°	-15°	0°
1	1	0.57	1.125	2.171	0.693	0.025	0.237	0.145
	2	3.046	3.444	5.912	2.67	0.772	0.59	0.691
	3	3.449	4.233	7.687	4.146	0.578	0.749	0.559
	4	1.861	1.795	2.476	1.81	1.305	1.002	1.163
	5	1.576	2.952	5.271	1.379	0.746	0.478	0.518
	6	3.16	4.897	7.976	3.311	0.873	1.482	1.804
	7	3.88	3.416	2.424	1.466	1.49	1.332	1.383
	8	1.378	2.802	2.746	0.78	0.68	0.819	0.739
	9	2.066	2.499	4.644	2.275	0.596	3.164	2.762
	10	5.894	5.444	6.532	3.876	2.285	1.778	2.278
	11	3.551	4.648	8.748	4.105	0.765	0.794	0.714
	12	2.369	4.796	10.4	2.355	0.65	1.714	1.8

In Table 1. every image from the examination set has a corresponding rotation angle (from -90 to 0 degrees). The table shows the 1st block (experiment) that containins 12 rows, where each row describes a different model. The 1st block contains the distances of the examination face from the 1st class for all 12 models and the row with the minimal distances is highlighted. As can be seen in the table, for instant the 1st experiment consisted of recognizing faces that belong to the 1st class. In the 1st block the model, which corresponds to this class, was rotated so that it would most resemble the turn angle of the face on the image and then the distances were calculated using formulas 1,2,3. As can be seen in the Table 1, the 1st line of the 1st block contains the smallest sums of distances. The graphical representation of the 1st block of the table can be seen in Figure 3, where 12 classes are shown in different colors, at different rotation angles of the head (from -90 to 0 degrees), and the smallest values of the sum of squared distances correspond to the 1st class. The second block had similar results (recognizing faces that belong to the second class) etc.

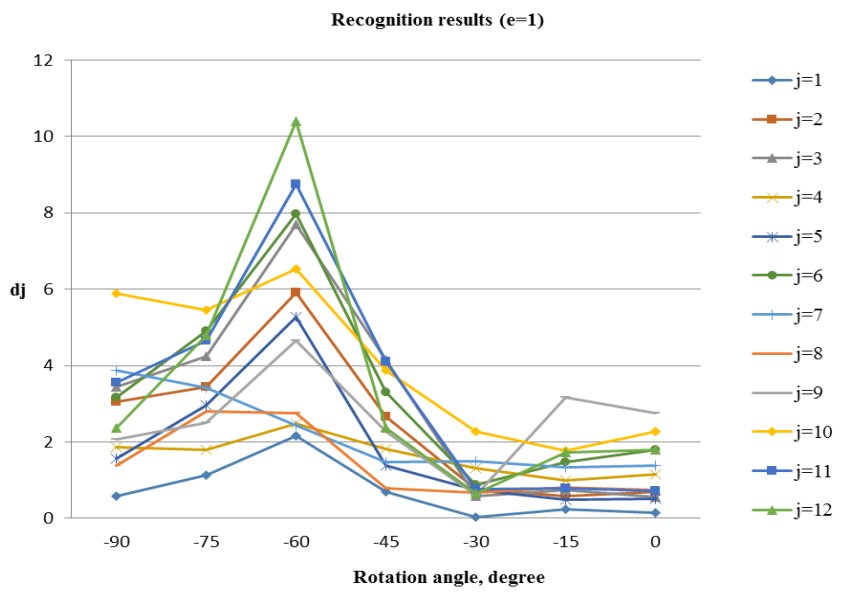

Figure 3. Recognition results for images e from the expanded training set that correspond to 1st class

In the case when the base training set (3 initial images) was used, the amount of errors for 49 images (7 classes) was 6, for (12 classes) 18.

4. CONCLUSION

In this paper, a method of face recognition is proposed, which uses information that was acquired with the help of a 3D model of the head and is aimed to solve face recognition tasks with an insufficient training set.

The created 3D model allows acquiring additional images of faces (at different angles), which significantly increases the results of recognition of unknown faces, as compared to only using the basic training set. The proposed methods were tested on various images of faces. As can be seen from the results, when the extended training set was used there were no mistakes.

The results have shown that these recognition methods can be used in cases, when the initial information about the shape of the face is insufficient, for example, in forensics.

REFERENCES

Blanz, V., Vetter, T. 2003. Face Recognition Based on Fitting a 3D Morphable Model. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 9, pp. 1063-1074.

Blanz, V., Romdhani, S., Vetter, T. 2002. Face Identification across different poses and illuminations with a 3D morphable model. *Proceedings of the IEEE International Automatic Face and Gesture Recognition.* Freiburg, Germany, pp. 192-197.

Bronstein, A. M., Bronstein. M. M., Kimmer, R. and Spira, A. 2003. 2003. 3D face recognition without facial surface reconstruction. *Center for Intelligent Systems Report CIS#2003-05, Technion-Israel Institute of Technology*.

Godil, A., Ressler, S., Grother, P. 2004. Face Recognition using 3D Face Shape and Color Map Information: Comparison and Combination, *Proceedings of SPIE, Vol 5404 Biometric Technology for Human Identification*, Orlando, FL, pp. 351-361.

Heisele, B., Ho, P., Poggio, T., 2001. Face recognition with support vector machines: Global versus component-based approach. Proceedings Eighth IEEE International Conference on Computer Vision. Vancouver, Canada, pp. 688–694.

Howland, P.; Wang, J.; Park, H. 2006. Solving the small sample size problem in face recognition using generalized discriminant analysis; Pattern Recognition, Volume 39, Issue 2, pp. 277-287

Huang, J., Heisele, B., and Blanz, V. 2001. Face recognition with support vector machines: global versus component-based approach. *Proceedings Eighth IEEE International Conference on Computer Vision.* Vancouver, Canada, pp. 688-694.

Krutikova, O., Glazs, A. 2013. Development of a New Method for Adapting a 3D Model from a Minimum Number of 2D Images. In *Technologies of Computer Control*, Vol.14, pp. 12-17.

Krutikova, O., Glazs, A. 2014. Increasing the Training Set in Face Recognition Tasks by Using a 3D Model of a Face. In *Technologies of Computer Control*. Vol.15, pp.14-19.

Popatheodorou, T., Rueckert, D. 2004. Evaluation of Automatic 3D Face Recognition Using Surface and Texture Registration, *Proceedings Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, Seul, Korea, pp. 321-326.

Romdhani, S., Vetter, T. 2003. Efficient, Robust and Accurate Fitting of a 3D Morphable Model. *Proceedings Ninth IEEE International Conference on Computer Vision.* Nice, France, pp. 59-66.

Scheenstra, A., Ruifrok, A. and Velktkamp, R. C., 2005. *Proceedings of* 5th International Conference, AVBPA 2005, NY, USA, pp. 891-899.

Tsalakanidou, F., Tzocaras, D. and Strintzis, M. 2003. Use of Depth and Colour Eigenfaces for Face Recognition. *In Pattern Recognition Letters*, Vol. 24, pp. 1427-1435.

Turk, M.A., Pentland, A.P. 1991. Face Recognition Using Eigenfaces. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Maui, Hawaii, USA, pp. 586-591.

Zhao, W.Y, Chellappa, R., Krishnaswamy, A., 1998. Discriminant analysis of principal components for face recognition. Proceedings of the 3rd IEEE Internatational Conference on Automatic Face and Gesture Recognition. Nara, Japan, pp. 336–341.

Zhao, W.Y., Chellappa, R. 2000. SFS Based View Synthesis for Robust Face Recognition. Proceedings of the IEEE International Automatic Face and Gesture Recognition. Grenoble, France, pp. 285-292.