Face recognition system using 2D and 3D information fusion

Svetlana V. Korobkova, Archil Tsiskaridze
Vocord Company, Moscow, Russia
svetlana.korobkova@vocord.ru

Abstract

The system using both 3D and 2D face information has been developed and tested. We use 3D face surface reconstructed from 4 images. These face images are obtained from 4 calibrated high definition cameras without any additional structured illumination. 2D face texture information is obtained using frontal face 3D view with face texture. Neural networks are used both for 2D texture recognition and for 2D and 3D fusion. The developed face recognition system achieves the accuracy of 98,5%.

***Keywords**: 3D object reconstruction, 3D face recognition, 2D face recognition, texture face recognition, face recognition methods fusion.*

1. INTRODUCTION

2D face recognition systems are the most prevalent recognition systems when it is required that the person can be scanned without any collaboration. All face recognition systems using 2D face images have a row of shortcomings. Such systems are sensitive to illumination conditions, direction and type of light source. The illumination on the image depends on the type of the light source, face geometry and reflectance and camera parameters. This makes it very difficult to design 2D face model which could correctly deal with all these parameters. 3D face recognition systems are free of these drawbacks. 3D face surface remains the same in any lighting conditions [5].

Another drawback of 2D face recognition systems is pose variance. 2D face recognition algorithms, which are capable of processing and recognizing faces with different poses have to include special processing steps, such as applying special transformation function to faces with pose variations [6]. The accuracy of such methods depends highly on the accuracy of landmark points detection on the face with pose variations and shows bad results when a part of face is poorly illuminated. Moreover this task cannot be strictly solved in 2D case in theory because of perspective projection property of 2D images. To overcome this problem it has been proposed to record and keep faces with pose variations in the database [9]. Nevertheless this approach requires storing of large amount of data and reduces face recognition system performance. Several more approaches have been proposed to deal with faces with pose variations. They are eigen light-fields method, flexible appearance models, statistical face models for interpolating to unseen views of the face [4, 7], and aggregate face models [12]. The use of 3D information makes the task of constructing frontal face view much easier to solve and provides rather accurate projected frontal view of the probe face [1].

According to face recognition methods overview [11], among state-of-the-art face recognition methods it is worth to mention global methods (with correct recognition rate 90-96%) and statistical methods (93-100%). Face recognition methods based on parametric face model construction have the accuracy about 88-96%.

After the analysis of the state of the art technology in face recognition we made a decision to use both 3D and 2D face information for recognition which led to achieving high accuracy rates of the system.

The developed face recognition system implements the method which uses the fusion of 3D face surface information and 2D face texture information. Due to this face recognition system uses the advantages of both approaches and makes it possible to overcome the drawbacks of each of them when they are used as 2 separate systems. Using this approach made it possible to achieve 98.5% correct recognition rate.

The results of 2D and 3D face similarity rate are obtained separately and used as the inputs of a neural network classifier. The output value of this classifier is the similarity measure for the gallery and the probe face image.

The input data for the face recognition system are the 4 images containing subject face, obtained from the high definition calibrated cameras. These images are processed (face search), and after that 3D face surface is reconstructed. Using the 3D face model and combination of face textures from the input images the 2D face texture is formed. Then both face 3D surface form and 3D face texture are used to find the corresponding gallery face in case it exists.

The developed face recognition system consists of the following processing steps:

- Face and landmarks search
- 3D face surface reconstruction and construction of frontal face texture;
- 3D face recognition;
- 2D texture face recognition;
- 2D and 3D recognition results fusion.

2. 3D FACE SURFACE RECONSTRUCTION

3D face surface reconstruction is done using images from 4 time-synchronized cameras as input. Cameras have 1280 x 1024 resolution and frame frequency is 10 pictures per second. Cameras are grouped in 2 vertical pairs, which makes it possible to place them on both sides of the checkpoint system. The distance between the cameras in a pair is 20 cm, the distance between 2 vertical camera pairs is about 1 m.

The proposed algorithm for matching points search on the images obtained from 2 cameras forming a stereo pair is implemented with subpixel accuracy of ¼ pixel value (due to image interpolation). This makes possible to reach overall effective accuracy about 0.2 mm, which comes from camera sensors discreteness. Such accuracy is enough for further high quality face recognition.

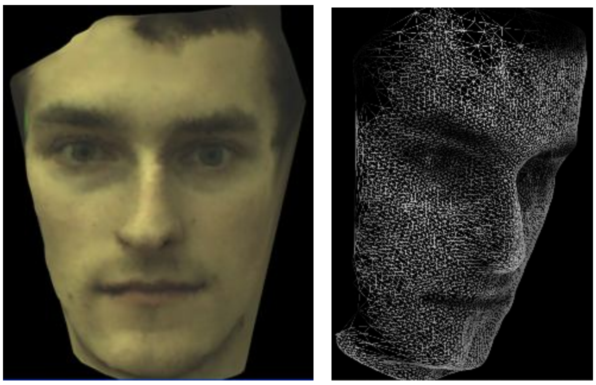
After the reconstruction using 2 camera stereo pairs is done the resulting 3D point sets are combined into one point set representing the face. The procedure of point sets combination is based on preliminary intrinsic and extrinsic camera calibration. Intrinsic camera parameters describe camera and lens properties and extrinsic describes camera positions and orientation. So the resulting point sets reconstructed from 2 stereo pairs on both sides of the checkpoint lay in the same coordinate system.

3D face surface reconstruction consists of the following steps:

- Building 3D point sets using each stereo pair and combining 2 point sets;
- Preliminary point set filtration;
- Building triangulated 3D face model (figure 1);
- Combining several face surfaces, obtained at different points of time into an aggregate 3D face surface.

To make 3D point set triangulation we use 2.5D triangulation method. The main idea of this method is projection 3D points to a 2D plane and then using 2D triangulation to produce 3D triangulated face surface. This method overcomes 3D triangulation methods both in reliability and in performance if applied to face 3D surfaces. Triangulation takes several milliseconds for the whole face surface.

The search for matching points on the images can be solved for each point independently. This feature of the algorithm makes possible significant speed up of the calculations due to performing them using parallel computation on a graphical processing unit. The reconstruction of a 3D object the image of which is 500x500 pixels (the usual size of the face in the image) with a 4 pixel step for 2 images from a stereo pair takes 5.1 sec. on Intel Core2 Q6600 2.4 GHz processor and only 0.17 sec. on NVIDIA GTX260 graphical processing unit. This provides 30 times speed up of the 3D surface reconstruction calculation.



(a) (б)

Figure 1: 3D face reconstruction. (a) – 3D surface with texture, (б) – triangulated face surface.

3D surface reconstruction can be made even under the condition of poor illumination of the face image (about 400-600 lux) with the distance from cameras to the reconstructed object up to 2.5 m. When using cameras with higher resolution, which become cheaper and easier to find, will make it possible to achieve good reconstruction results for even larger distances between cameras and the object of interest.

The 3D face model is based on parametric description, The construction of the face model is based on searching face parameters which fit closely to the surface obtained from cameras. If 3D face obtained from cameras includes artifacts (i.e. glasses, mustaches), they are eliminated, and thus do not contribute to recognition score of 3D recognition algorithm, see figure 2).

Figure 2: Face with glasses, glasses do not have volume and appear only as texture trails.

3. 3D FACE RECOGNITION

3D face recognition is done in two steps:

1. Fast database sample comparison using feature set of small length;
2. Precise comparison of several most similar database face samples.

Figure 3: Reconstructed 3D face surface - side view (while cameras are placed in front of the face).

4. 2D TEXTURE FACE RECOGNITION

To perform 2D face recognition first we form a synthetic face image, face texture is projected to the 3D face surface and the face is rotated to make frontal view of this face. The resulting image is then regarded as the input of 2D face recognition algorithm and called face texture. For this face texture we perform illumination correction procedure which yields face texture image with the specified mean illumination level.

For face recognition we use a set of classifiers, each of them analyzes different face parameters and outputs similarity rate for 2 face textures – probe face texture and database face texture. These classifiers are chosen to be neural network.

Each neural network classifier takes a vector of difference features as input. For each component of the difference feature vector we take its absolute value, so that classification is symmetric. The output of each classifier is the estimation of probability (the value from 0 to 1) that 2 faces are of the same person by this classifier.

The outputs of these classifiers are used as the input values for the overall classifier, which determines the probability of 2 face textures being of the same person, which is considered as the output of 2D face recognition system part.

The training on separate face texture neural network classifiers and of the overall classifier has been conducted on a training set of about 700000 difference features. We have collected and sorted face database and know the person name for each texture sample.

5. RESULTS

The developed face recognition system using both 3D and 2D information for recognition was tested on big enough database containing both 3D information and 2D texture for each face instance and known classes (person names) for these samples. We have reached 98,5 recognition accuracy rate at FAR = FRR point. The ROC plot for the proposed face recognition algorithm using 2D and 3D information fusion is shown in figure 4.

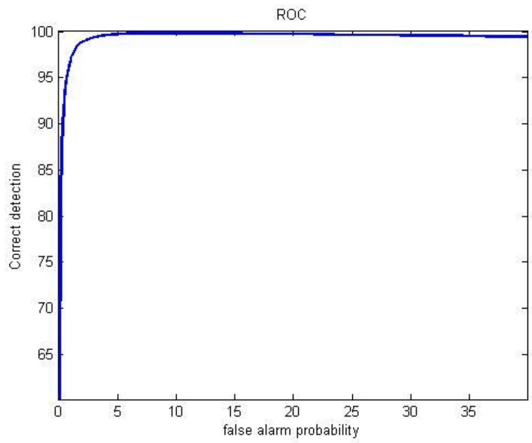


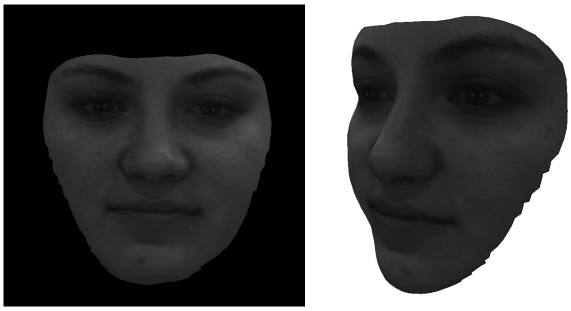
Figure 4: ROC of the face recognition system.

Our face recognition system proved to be stable when probe faces had different poses (such as rotation (+/- 30 degrees), leaning down and looking up faces (50/30 degrees up and down respectively)), moreover the system still can recognize a person if some parts of face are occluded (or there are no key points on this region, i.e. eyes, mouth).

The proposed face recognition algorithm is stable when input face has slight face expression variations: smiles, open mouth, squint faces.

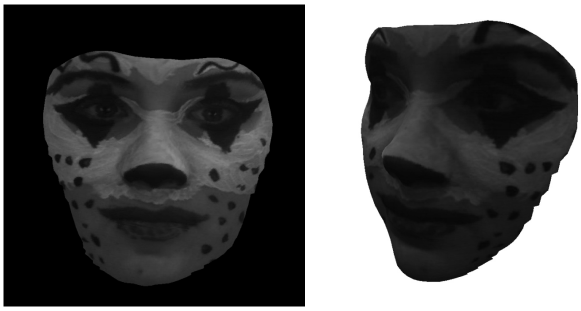
We have conducted experiments with painted faces and found out that face recognition system still gives correct answer for heavily painted faces due to the usage of 3D face information which is unchanged in case of painted faces.

Figure 5 shows the face sample (3D form and 2D texture) without painting. This sample has been used as a database sample. Figure 6 shows a painted face of the same person. This face was used as the probe face. Our system still could recognize the person whom this face belongs to, in spite of great difference of the textures.



a) б)

Figure 5: Original face: a) – texture, б) – 3D surface



a) б)

Figure 6: Painted face: a) – texture б) – 3D surface

Our system is able to work in 2D-only recognition mode, This mode is important if no preliminary 3D information about the face is available. In this case recognition is based only on 2D information (i.e. single personal photograph) stored in the database. However, the benefit of our full approach compared to other 2D face recognition methods is pose correction before image comparison with database. This correction is carried out based on our 3D reconstruction with face rotation and severely improves recognition rate.

To evaluate 2D recognizer itself we tested it on FERET data base. Consider situation when 2D photograph available is also not perfect: i.e. it is taken from another camera under poor illumination conditions, with variable head pose, glasses etc, In such situation it is important for 2D recognizer to show stable results within some image variation range, even if test image (acquired by the system) is purely frontal after 3D correction.

The basic steps of 2D recognizer are as follows and include preprocessing of the input image consisting of the following steps:

- face detection;
- landmark points search;
- deleting background;
- illumination correction.

After we get the input face image of specific size and pose we use it as the input for 2D face neural network classifiers similar to texture classifiers, but trained on the difference features of real face images.

This 2D face recognizer has been tested on FERET database, consisting of more than 2000 face images of about 700 different people. The images included pose variations with 10 degrees. Each person has more than 1 face instance in the database. Face pictures in FERET database were taken at different points of time, with different poses and in different illumination conditions.

Also the faces in FERET database have such variation as presence or absence of glasses, beard and moustache.

Our 2D only recognition system reached 88,5% correct recognition rate at FAR=FRR point.

6. CONCLUSION

We have designed and implemented face recognition system (figure 7) which uses both 3D and 2D information for recognition. 3D face information is obtained using only high definition cameras without any help of additional structured light sources. The system can deal faces with different poses, rotation, occluded parts of faces, faces with emotions and painted faces. 2D and 3D face data fusion make possible to achieve 98,5% correct recognition rate.

Figure 7: 2D+3D face recognition system.

7. REFERENCES

[1] Blanz, V., Grother, P., Phillips, J., and Vetter, T. (2005). *Face recognition based on frontal views generated from non-frontal images. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 454–461.*

[2] Chua, C. and Jarvis, R. (1997). *Point signatures - a new representation for 3D object recognition. International Journal of Computer Vision, 25(1): pp. 63–85.*

[3] Gordon, G. (1992). *Face recognition based on depth and curvature features. IEEE Computer Vision and Pattern Recognition (CVPR), pp. 808–810.*

[4] Gross, R., Matthews, I., and Baker, S. (2002). *Eigen light-fields and face recognition across pose. International Conference on Automatic Face and Gesture Recognition.*

[5] Hesher, C., Srivastava, A., and Erlebacher, G. (2003). *A novel technique for face recognition using range imaging. International Symposium on Signal Processing and Its Applications, pp. 201–204.*

[6] Kim, T. and Kittler, J. (2005). *Locally linear discriminant analysis for multimodally distributed classes for face recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(3): pp. 318–327.*

[7] Lanitis, A., Taylor, C., and Cootes, T. (1995). *Automatic face identification system using flexible appearance models. Image and Vision Computing, 13(5): pp. 393–401.*

[8] Lee, Y., Song, H., Yang, U., Shin, H., and Sohn, K. (2005). *Local feature based 3D face recognition. International Conference on Audio- and Video-based Biometric Person Authentication, pp. 909–918.*

[9] Li, Y., Gong, S., and Lidell,H. (2000). *Support vector regression and classification based multiview face detection and recognition. International Conference on Face and Gesture Recognition, pp. 300–305.*

[10] Moreno, A., Sanchez, A., Velez, J., and Diaz, F. (2003). *Face recognition using 3D surface extracted descriptors. In Irish Machine Vision and Image Processing Conference.*

[11] Papatheodorou, T. and Rueckert, D. (2007). *3D face recognition, book Face recognition, Vienna, Austria, 2007, pp. 417-423.*

[12] Prince, S. and Elder, J. (2006). *Tied factor analysis for face recognition across large pose changes. British Machine Vision Conference.*

About the author

Svetlana Korobkova is a senior researcher at Vocord. Her contact e-mail is svetlana.korobkova@vocord.ru.

Archil Tsiskaridze is a researcher at Vocord. His contact e-mail is svetlana.korobkova@vocord.ru.