Bosphorus Database for 3D Face Analysis

Arman Savran[1], Neşe Alyüz[2], Hamdi Dibeklioğlu[2], Oya Çeliktutan[1], Berk Gökberk[3], Bülent Sankur[1], and Lale Akarun[2]

[1] Boğaziçi University, Electrical and Electronics Engineering Department
{arman.savran,oya.celiktutan,bulent.sankur}@boun.edu.tr
[2] Boğaziçi University, Computer Engineering Department
{nese.alyuz,hamdi.dibeklioglu,akarun}@boun.edu.tr
[3] Philips Research, Eindhoven, The Netherlands
berk.gokberk@philips.com

Abstract. A new 3D face database that includes a rich set of expressions, systematic variation of poses and different types of occlusions is presented in this paper. This database is unique from three aspects: i) the facial expressions are composed of judiciously selected subset of Action Units as well as the six basic emotions, and many actors/actresses are incorporated to obtain more realistic expression data; ii) a rich set of head pose variations are available; and iii) different types of face occlusions are included. Hence, this new database can be a very valuable resource for development and evaluation of algorithms on face recognition under adverse conditions and facial expression analysis as well as for facial expression synthesis.

1 Introduction

In recent years face recognizers using 3D facial data have gained popularity due to their relative advantages over 2D schemes, notably lighting and viewpoint independence. This trend has also been enabled by the wider availability of 3D range scanners. The 3D face processing can be envisioned in two roles, either as a single modality biometric approach in lieu of the 2D version or in a complementary mode in a multi-biometric scheme. Another prime application of 3D facial processing is the understanding of facial expressions for affective human-computer interfaces.

Most of the existing methods for facial feature detection and person recognition assume frontal and neutral views only, and hence biometry systems have been designed accordingly. However, this forced posing can be uncomfortable for the subjects and limit the application domains. Therefore, the pressing need in this field is to develop algorithms working with natural and uncontrolled behaviour of subjects. A robust identification system can also cope with the subjects who try to eschew being recognized by posing awkwardly and worse still, by resorting to occlusions via dangling hair, eyeglasses, facial hair and other accessories.

On the other hand, understanding of facial expressions has wide implications ranging from psychological analysis to affective man-machine interfaces. Once the expression is recognized, this information can also be used to help the person identifier and/or adapt the interface.

B. Schouten et al. (Eds.): BIOID 2008, LNCS 5372, pp. 47–56, 2008.

The desiderata of a 3D face database enabling a range of facial analysis tasks ranging from expression understanding to 3D recognition are the following: i) Action units from Facial Action Coding System (FACS) [1], both single and compound; ii) Emotional expressions; iii) Ground-truthed poses; iv) Occlusions originating from hair tassel, eyeglasses and a gesticulating hand. Motivated by these exigencies, we set out to construct a multi-attribute 3D face database.

Table 1. List of some well known 3D face databases. Sub.: subjects Samp.: samples per subject, Occl.: occlusions, NA: not available.

Database	Sub.	Samp.	Total	Expression	Pose	Occl.
Bosphorus	105	31-54	4652	34 expressions (action units & six emotions)	13 yaw, pitch & cross rotations	4 occlusions (hand, hair, eyeglasses)
FRGC v.2 [2]	466	1-22	4007	Anger, happiness, sadness, surprise, disgust, puffy	NA	NA
BU-3DFE [3]	100	25	2500	Anger, happiness, sadness, surprise, disgust, fear (in 4 levels)	NA	NA
ND2006 [4]	888	1-63	13450	Happiness, sadness, surprise, disgust, other	NA	NA
York [5]	350	15	5250	Happiness, anger, eyes closed, eye-brows raised	Uncontrolled up & down	NA
CASIA [6]	123	15	1845	Smile, laugh, anger, surprise, closed eyes	NA	NA
GavabDB [7]	61	9	549	Smile, frontal accentuated laugh, frontal random gesture	Left, right, up, down	NA
3DRMA [8]	120	6	720	NA	Slight left/right & up/down	NA

1.1 Comparisons with Major Open 3D Face Databases

Various databases for 3D face recognition and occasionally 3D expression analysis are available. Most of them are focused on recognition; hence contain a limited range of expressions and head poses. Also, none of them contain face occlusions. One of the most popular 3D database FRGC v.2 [2], though the biggest one in the number of subjects has only a few mild expressions. The database richest in the spectrum of emotional expressions is BU-3DFE [3]. Every subject displays four intensity levels of the six emotions. Table I lists publicly available databases of relevance and compares with our database.

The Bosphorus database represents a new comprehensive multi-expression, multi-pose 3D face database enriched with realistic occlusions. The database has the following merits: i) Many action units from the FACS [1] in addition to the basic six emotional expressions; ii) Various ground-truthed head poses are available; iii) A number of facial occlusion types are captured from the subjects. Finally in order to achieve more natural looking expressions, we have employed actors and actresses from professional theatres, opera and the conservatory school.

The content of the database is given in Section 2, and data acquisition is explained in Section 3. In Section 4 the acquired data are evaluated. Finally conclusion is given in Section 5.

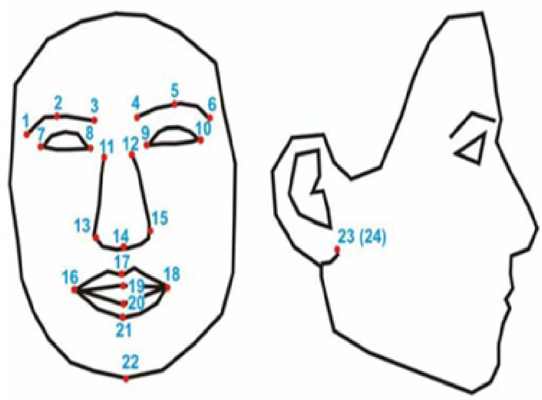
2 Database Content

The database consists of 105 subjects in various poses, expressions and occlusion conditions. 18 men had beard/moustache and 15 other subjects had short facial hair. The majority of the subjects are aged between 25 and 35. There are 60 men and 45 women in total, and most of the subjects are Caucasian. Also, 27 professional actors/actresses are incorporated in the database. Up to 54 face scans are available per subject, but 34 of these subjects have 31 scans. Thus, the number of total face scans is 4652. Each scan has been manually labelled for 24 facial landmark points such as nose tip, inner eye corners, etc, provided that they are visible in the given scan. These feature points are given in Table II.

In the following subsections, the collected facial expressions, head poses and occlusions are explained in detail.

Table 2. Manually labeled 24 facial landmark points

1-6. Inner/middle/outer eyebrow
7-8. Outer left/right eye corners
9-10. Inner left/right eye corners
11-12. Nose saddle left/right
13. Left nose peak
14. Nose tip
15. Right nose peak
16. Left mouth corner
17. Upper lip outer middle
18. Right mouth corner
19. Upper lip inner middle
20. Lower lip inner middle
21. Lower lip outer middle
22. Chin middle
23-24. Ear lobe left/right



2.1 Facial Expressions

Two types of expressions have been considered in the Bosphorus database. In the first set, the expressions are based on action units (AUs) of the FACS [1]. AUs are assumed to be building blocks of facial expressions, and thus they can constitute a flexible basis for them. Since each action unit is related with the activation of a distinct set of muscles, their veracity can be assessed quite objectively. Out of 44 AUs in FACS, we have collected a subset which was easier to enact. The selected action units were grouped into 20 lower face AUs, five upper face AUs and three AU combinations.

Fig. 1. Some samples from happiness expression captured from actors/actresses. Texture mapping and synthetic lighting is applied for rendering.

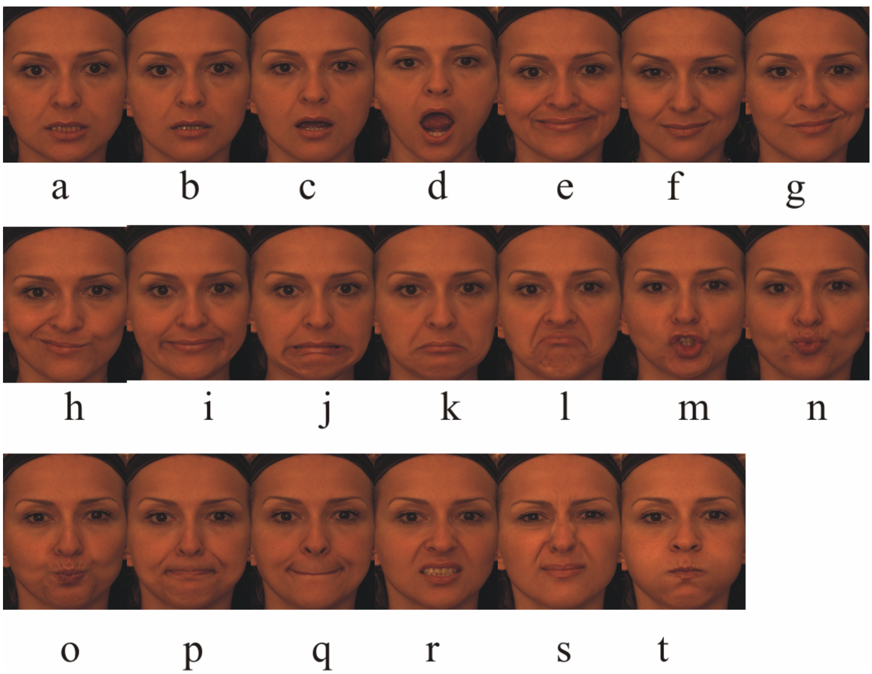


Fig. 2. Lower face action units: lower lip depressor (a), lips part (b), jaw drop (c), mouth stretch (d), lip corner puller (e), left lip corner puller (f), right lip corner puller (g), low intensity lip corner puller (h), dimpler (i), lip stretcher (j), lip corner depressor (k), chin raiser (l), lip funneler (m), lip puckerer (n), lip tightener (o), lip presser (p), lip suck (q), upper lip raiser (r), nose wrinkle (s), cheek puff (t).

In the second set, facial expressions corresponding to certain emotional expressions were collected. These are: happiness, surprise, fear, sadness, anger and disgust. These expressions were found to be universal among human races [9].

For the acquisition of action units, subjects were given explications about the intended action as well as negative feedback if they did not enact correctly. Also to facilitate the instructions, a video clip showing the correct facial motion for the corresponding action unit was displayed on the monitor [10]. However, in the case of emotional expressions, there were no video or photo guidelines so that subjects had to improvise. Only if they were unable to enact, they were told to mimic the expression in a recorded video or photograph. Moreover, a mirror was placed in front of the subjects for immediate visual feedback.

In Fig. 2, Fig. 3, Fig. 4 and Fig. 5, the 34 expressions in the database are given. Also, Fig. 1 shows some 3D faces displaying the happiness emotions of actors/actresses. These facial images are rendered with texture mapping and synthetic lighting.

It is important to note that not all subjects could properly produce all AUs, some of them were not able to activate related muscles or they could not control them. Therefore, in the database few expressions are not available for some of the subjects. Also, the captured AUs need to be validated by trained AU experts. Second, since no video acquisition was possible for this database, the AUs were captured at their peak intensity levels, which were judged subjectively. Notice that there was no explicit control for the valence of pronounced expressions. As in any other database, acted expressions are not spontaneous and thoroughly natural. All these factors constitute the limitations of this database for expression studies.

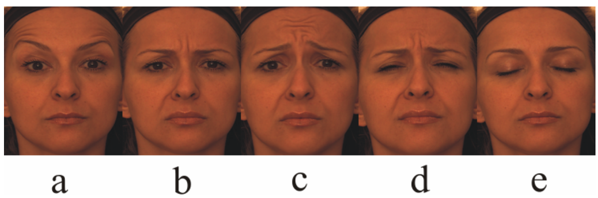


Fig. 3. Upper face action units: outer brow raiser (a), brow lowerer (b), inner brow raiser (c), squint (d), eyes closed (e)

Fig. 4. Action unit combinations: jaw drop + low intensity lip corner puller (a), lip funneler + lips part (b), lip corner puller + lip corner depressor (c)

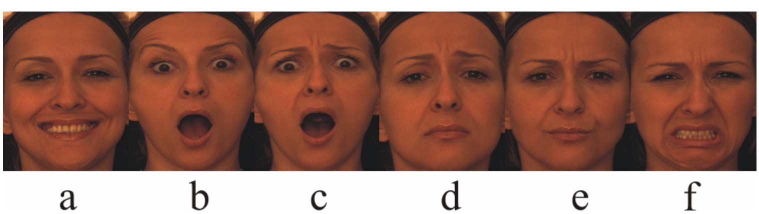


Fig. 5. Emotional expressions: happiness (a), surprise (b), fear (c), sadness (d), anger (e), disgust (f)

2.2 Head Poses

Various poses of the head are acquired for each subject (Fig. 6). There are three types of head poses which correspond to seven yaw angles, four pitch angles, and two cross rotations which incorporate both yaw and pitch. For the yaw rotations, subjects align themselves by rotating the chair on which they sit to align with stripes placed on the floor corresponding to various angles. For pitch and cross rotations, the subjects are required to look at marks placed on the walls by turning their heads only (i.e., no eye rotation).

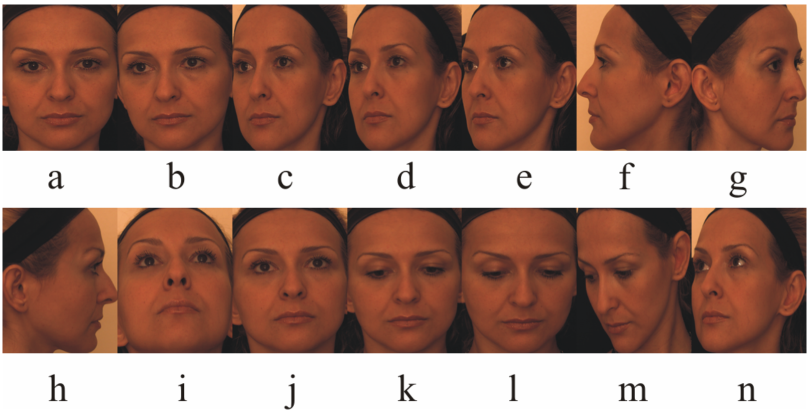


Fig. 6. Head poses: neutral (a); yaw rotations: +10° (b), +20° (c), +30° (d), +45° (e), +90° (f), -45° (g) and -90° (h) pitch rotaions: upwards (i), slight upwards (j), slight downwards (k), downwards (l); right-downwards (m) and right-upwards (n)

Notice that pose angles are prone to slight errors. Especially poses including pitch rotations can be subjective, since subjects were requested to turn their head toward target marks. This introduced slight angular errors due to differences of rotation centres which depended on subjects. Whenever subjects were tempted to rotate their eyes in lieu of their heads toward the targets, they were warned.

2.3 Occlusions

For the occlusion of eyes and mouth, subjects were allowed to choose a natural pose for themselves; for example, as if they were rubbing their eyes or as if they were surprised by putting their hands over their mouth. For occlusion with eyeglasses, we had a pool of different eyeglasses so that each subject could select at random one of them. Finally, if subjects' hair was long enough, their faces were also scanned with hair partly occluding the face (Fig. 7).

The subject to subject variation of occlusions is more pronounced as compared to expression variations. For instance, while one subject occludes his mouth with the whole hand, another one may occlude it with one finger only; or hair occlusion on the forehead may vary a lot in tassel size and location.

Fig. 7. Occlusions: eye occlusion with hand (a), mouth occlusion with hand (b), eyeglasses (c) and hair (d)

3 Data Acquisition

Facial data are acquired using Inspeck Mega Capturor II 3D, which is a commercial structured-light based 3D digitizer device [11]. The sensor resolution in x, y & z (depth) dimensions are 0.3mm, 0.3mm and 0.4mm respectively, and colour texture images are high resolution (1600x1200 pixels). It is able to capture a face in less than a second. Subjects were made to sit at a distance of about 1.5 meters away from the 3D digitizer. A 1000W halogen lamp was used in a dark room to obtain homogeneous lighting. However, due to the strong lighting of this lamp and the device's projector, usually specular reflections occur on the face. This does not only affect the texture image of the face but can also cause noise in the 3D data. To prevent it, a special powder which does not change the skin colour is applied to the subject's face. Moreover, during acquisition, each subject wore a band to keep his/her hair above the forehead to prevent hair occlusion, and also to simplify the face segmentation task.

The propriety software of the scanner is used for acquisition and 3D model reconstruction. We reconstructed faces right after the image acquisition, which although laborious, gave us an opportunity to quickly check the quality of the scanning, and repeat it if necessary. In this phase data is also segmented manually by selecting a polygonal face region. In order to remove noise, several basic filtering operations (like Gaussian and Median filtering) are applied. Finally, each scan is down-sampled and saved in two separate files that store colour photograph and 3D coordinates. A segmented 3D face approximately consists of 35K points.

4 Discussion of Data Quality

Despite precautions and adjustments for maximum quality data, some errors and noise persisted due to the 3D digitizing system and setup conditions. The remaining bugs are explained below.

Movements: Though images are captured within one second, motion of the subjects' faces can be source of severe data corruption. A comfortable seat with a headrest was used to diminish the subject movements during long acquisition sessions. However, this problem can also happen for instance due to breathing or muscle contractions during expressions. Therefore, faces that were deemed to be seriously faulty were recaptured. In the database, movement noise emerges especially in case of expressions, but depends on the subject and occurs sporadically. An example is shown in the middle row of Fig. 8.

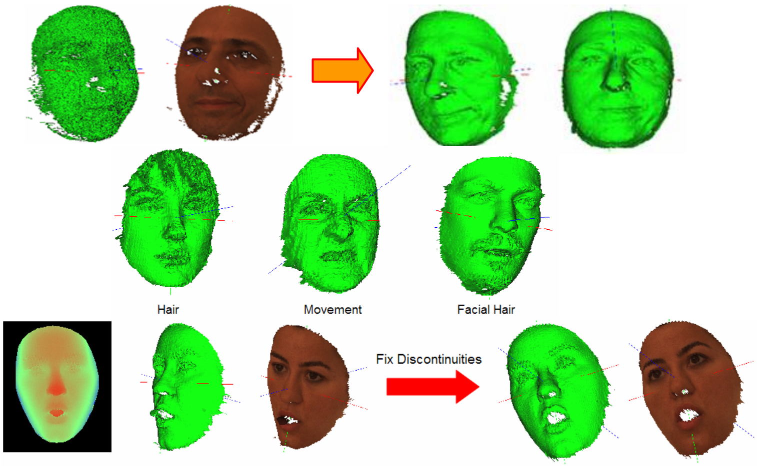


Fig. 8. Commonly occurring problems during image acquisition and face reconstruction. Top row shows basic filtering and self-occlusion problem. In the middle row, noise due to hair, movement, and facial hair is seen. At the bottom left, an error in the depth level of the tongue, and at the right, its correction is displayed.

Hairs and Eyes: Data on hair and facial hair, such as beard and eyebrows, generally causes spiky noise. Spiky surfaces arise also over the eyes. Basic smoothing filtering reduces these types of noises (Fig. 8).

Self-occlusions: Since data are captured from single views with this system, self-occlusions occur. The consequences are holes in the facial data, and uncompleted and distorted facial contours. Holes are formed due to missing data, mostly at the sides of the nose. Even slight head rotations generate high amount of self-occlusions. In Fig. 2 some of these problems are illustrated. No processing was performed for these problems.

Discontinuity: Discontinuity problems develop either inside the mouth when mouth is open, or in occluded face scans. The reconstruction of depth values at these discontinuous regions can sometimes be faulty. These errors are corrected by manual intervention using the system's software (Fig. 8).

5 Conclusion and Future Work

We have described the components, merits and limitations of a 3D face database, rich in Action Units, emotional expressions, head poses and types of occlusions. The involvement of actors/actresses, especially in the case of expressions, is considered to be an advantage.

Several projects in the area of biometry and man-machine interfaces are being conducted on this database. Face recognition experiments have already been carried out on this database. These experiments consider the effect of face registration on the identification performance when the reference face model is obtained from neutral faces while test faces contain a variety of expressions. This research is presented in a companion paper [12]. Another research path is that of automatic facial landmarking. Automatically located landmarks can be used as initial steps for better registration of faces, for expression analysis and for animation. Various algorithms ranging from active appearance models to bunch graphs and statistical matched filter are studied [13].

For facial analysis and synthesis applications, non-rigid registration of faces is a very important intermediate step. Although variations due to expressions can be analyzed by rigid registration or landmark-based non-rigid registration methods, more faithful analysis can only be obtained with detailed non-rigid registration. Improved registration with non-rigid methods facilitates automatic expression understanding, face recognition under expressions and realistic face synthesis studies. The ill-posed problem of non-rigid registration has been addressed in [14].

Finally, this database has been used for 3D face detection purpose. A recently introduced generic transformation invariant 3D feature detector [15] is being experimented to locate scanned faces in 3D space.

Acknowledgements. We would like to thank to subjects who voluntarily let their faces to be scanned. We are very thankful to Aydın Akyol from Istanbul Technical University; Jana Trojanova from West Bohemia University; Semih Esenlik and Cem Demirkır from Boğaziçi University; Nesli Bozkurt, İlkay Ulusoy, Erdem Akagündüz and Kerem Çalışkan from Middle East Technical University; and T. Metin Sezgin from University of Cambridge. They worked and helped for the collection and processing of large part of this database during the Enterface'07 workshop project (http://www.enterface.net/enterface07/). This work is supported by TÜBITAK 104E080 and 103E038 grants.

References

1. Ekman, P., Friesen, W.V.: Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto (1978)
2. Phillips, P., Flynn, P., Scruggs, T., Bowyer, K., Chang, J., Hoffman, K., Marques, J., Min, J., Worek, W.: Overview of the face recognition grand challenge. In: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference, vol. 1, pp. 947–954 (2005)
3. Yin, L., Wei, X., Sun, Y., Wang, J., Rosato, M.J.: A 3D Facial Expression Database For Facial Behavior Research. 7th Int. In: Conference on Automatic Face and Gesture Recognition (FGR 2006), April 10-12, 2006, pp. 211–216 (2006)
4. Faltemier, T.C., Bowyer, K.W., Flynn, P.J.: Using a multi-instance enrollment representation to improve 3d face recognition. In: First IEEE Int. Conf. on Biometrics: Theory, Applications, and Systems, BTAS 2007, pp. 1–6 (2007)
5. Heseltine, T., Pears, N., Austin, J.: Three-dimensional face recognition using combinations of surface feature map subspace components. Image and Vision Computing 26, 382–396 (2008)
6. Zhong, C., Sun, Z., Tan, T.: Robust 3d face recognition using learned visual codebook. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2007, pp. 1–6 (2007)

7. Moreno, A., Sanchez, A.: Gavabdb: A 3d face database. In: Proc. 2nd COST275 Workshop on Biometrics on the Internet (2004)
8. Beumier, C., Acheroy, M.: Face verification from 3d and grey level cues. Pattern Recognition Letters 22, 1321–1329 (2001)
9. Ekman, P., Friesen, W.V.: Constants Across Cultures in the Face and Emotion. Journal of Personality and Social Psychology 17(2), 124–129 (1971)
10. Wallraven, C., Cunningham, D.W., Breidt, M., Bülthoff, H.H.: View dependence of complex versus simple facial motions. In: Bülthoff, H.H., Rushmeier, H. (eds.) Proceedings of the First Symposium on Applied Perception in Graphics and Visualization, vol. 181, ACM SIGGRAPH (2004)
11. Inspeck Mega Capturor II Digitizer, `http://www.inspeck.com/`
12. Alyüz, N., Gökberk, B., Dibeklioğlu, H., Savran, A., Salah, A.A., Akarun, L., Sankur, B.: 3D Face Recognition Benchmarks on the Bosphorus Database with Focus on Facial Expressions. In: First European Workshop on Biometrics and Identity Management Workshop (BioID 2008) (2008)
13. Çeliktutan, O., Sankur, B., Akakın, H.Ç.: Multi-Attribute Robust Facial Feature Localization. FG 2008: 8th IEEE on Automatic Face and Gesture Recognition, Amsterdam (September 2008)
14. Savran, A., Sankur, B.: Non-Rigid Registration of 3D Surfaces by Deformable 2D Triangular Meshes. In: CVPR 2008: Workshop on Non-Rigid Shape Analysis and Deformable Image Alignment (NORDIA 2008), Alaska (2008)
15. Akagündüz, E., Ulusoy, İ.: 3D Object Representation Using Transform and Scale Invariant 3D Features. In: Workshop on 3D Representation for Recognition (3dRR 2007), ICCV 2007 (2007)