A SURVEY ON 3D FACE MODELING

Serena Mejia and Marcelo Romero
Universidad Autónoma del Estado de México
sere379@yahoo.com.mx
mromeroh@uaemex.mx

Abstract

In the advance of science, different disciplines are interested to study the human body, among of them we have the facial anthropometry, which studies the human face's dimension and proportions by detecting what is called anthropometric landmarks over the facial surface. Such facial landmarks can be detected by using either direct or indirect techniques. Within a 3D face image it is generally know that the most distinctive facial landmarks are the inner-corners and the tip of the nose, although their appearance varies among human races. Hence, most face processing application generally relay on a face model for comparison and analysis based on such anthropometric landmarks. This paper presents a survey on 3D face modeling, which in the best of our knowledge is missed in relate literature and it is necessary to propose a new 3D facial model, and such model is a primary step in any face processing application. In particular, we are aim to produce a useful 3D face model for research in key areas based on Mexican anthropometric analysis.

1. Introduction

The study of human morphologic pattern is vital in different scientific disciplines as in medicine and social science to identify differences existing among races or demographic groups. Anthropometry consists in a series of technical measures that express quantitatively the dimensions of the human body and it has been increasingly used in the biomedical sciences (orthodontics, plastic surgery and maxillofacial surgery, etc.).

Hernández [1] mentioned that particularly craniofacial anthropometry refers to the study of the dimensions and proportions of the parts of the human face, in order to understand the physical changes of human kind and the differences between their races. The craniofacial anthropometric assessment is based on the determination of characteristic points of the face and skull, defined in terms of visible or palpable complex facial features. The measures of a population should show normal ranges that allow craniofacial measurements locate the subject within a range, establishing normal ranges of proportions of a human group.

In order to improve accuracy in such measures, since the early 80's has been systematizing this process under direct and indirect techniques these being the means of obtaining accurate data. Direct technical measures obtained directly from the surface of the head and face; while indirect techniques are based on *photogrammetry bidimensional* (2D) and *three-dimensional (3D) anthropometry* to measure the human face. Photogrammetry is based on obtaining measurements from 2D images of the subjects; while 3D anthropometry creates virtual models of the human face from 3D data.

One way to develop a face model is based on anthropometric points, so as mentioned Phothisane et al. [4] a facial anthropometric landmark describes a specific anatomic point over the human face, which may be the tip of the nose, corners of the eyes or the corners of the mouth. These anthropometric points exist in all the faces; however, its specific location differs depending on the subject. Then, by detecting those facial landmarks, researchers are interested in building facial models. For example, Golovinskiy et al. [11], postulate that a facial model such be universal and easy to be computed and edited. On the other hand, Cootes et al. [13], believes that a useful facial model such be specific. Another approach is found in the called generative 3D facial modelling, proposed by Paysan et al. [7], which are able to produce new faces from a sample.

This paper presents a literature review of the Facial models that have been created over the years, which in the best of our knowledge is missed in relate documents and it is a primary step in our research. We are aim to produce a 3D face model based on Mexican anthropometric analysis for face processing applications. The rest of this paper is as follows. Section 2 documents background information used into 3D face modelling. Finally, Section 3 discusses the paper presented in this survey, concludes about 3D face modelling and outlines our future work.

978-1-4673-6756-1/14 $31.00 © 2014 IEEE
DOI 10.1109/ICMEAE.2014.31

2. Face modelling in 3D

In this section we analysis most of the 3D face models available in the literature.

Conde [9] explains that today is possible to acquire two types of facial data, intensity images (2D) and three-dimensional data (3D). The first type of data represents the texture of the face while the second type collects facial geometric structure. Both types of data are complementary because they provide two different types of information. However the main limitation of face-based texture representation systems (2D) is its dependence on the lighting conditions and the position of the face. Meanwhile 3D rendering, by definition does not depend on the illumination and allows the normalization of the position of the face.

As for its characteristics and purpose, facial models ideally, according to Golovinskiy et al. [11], must be universal, easy to acquire and compute, its goal is to capture all the aspects of the face of a person into a digital model. Cootes and Taylor [13], also indicate that a model is useful if it is specific, this means that the model is able to represent the modeled object.

According to the literature, common face processing applications are facial recognition and analysis of the position of the head. Table 1 summarizes the facial models indicating information such as author, year of publication, number of vertices, number of polygons and the number of subjects used for face modeling construction. The works are listed chronologically by year of publication.

TABLE 1. 3D FACIAL MODELS

Model/author/year	Purpose	Vertices	Polygons	Subjects
Park [3], 1974	Speech animation	356	334	---
Candide-3 [5,6], 2001	Facial expressions and animation	75	100	---
Morphable Model, Blanz y Vetter [2] 1999	Generation of human faces	70,000	---	200
Bronstein et al [10], 2003	Facial recognition	2,000 a 2,500	---	157
El-Hussuna [14], 2003	Facial model	---	---	8
Golovinskiy et al [11], 2006	Facial details modeling	---	500,000	149
Basel Face Model, Paysan et al [6] 2009	Generation of human faces	53,490	160,470	200
Ramirez et al [12], 2009	Facial recognition	2,777	2,676	---
Phothisane et al [4], 2011	Head tracking	3,000	---	200

Parke [3] presents a three-dimensional, symmetric and parametric facial model producing facial images. The face is constructed by polygonal surfaces and manipulated through parameters to control interpolation, translation, rotation and scaling of various facial features with the purpose of allow representing a facial expression. The model presented by Parke [3] requires little input information to specify and generate a specific face in a specific expression. The model of Park [3] is based on the algorithm of Watkins which to solve the problem of visible surface to polygonal objects and Gouraud technique for smooth shading of these objects. Figure 1 shows the Facial model developed by Park [3]. Parke [3] collected data from ten different faces and with them made an animated film to show the transitions from face to face.

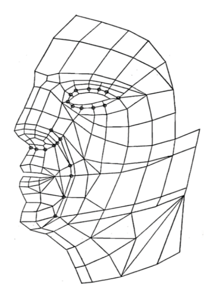
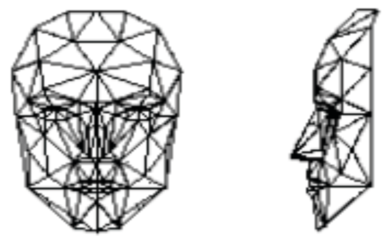


Figure 1. Facial model proposed by Parke [3] in 1974; the small circles indicate vertices found in the skin folds.

The facial model CANDIDE [5, 6] is a parameterized mask specifically designed to model human faces. This model handle a small number of polygons, about 100, that allow a fast reconstruction with a moderated computing power, and is controlled by local Actions Units (AUs). This model [5, 6] was developed by Mikael Rydfalk in the Group of image coding in Linköping University in Sweden en 1987. The original model has 75 vertices and 100 triangulations and it has being modified 3 times. The Figure 2 presents the CANDIDE model mask.



(a) (b)

Figure 2. Frontal (a) and lateral (b) view of CANDIDE model [5 y 6].

Blanz y Vetter [2] proposed a parameterized technique of facial modeling for arbitrary human faces. Studying statistics of facial data in 3D (shape and texture) the authors [2] built a *Morphable Face Model* and obtained knowledge about the facial variations using pattern classification. Computing the average face and the principal variations in the data group, they obtained a probability distribution to avoid "no human faces". The fundamental part of the work of Blanz y Vetter [2] is a general model of human face. They limited the range of acceptable faces based in prototypes of human faces, they used thousands of vertices per face leading to a direct triangulation of the surface which no require any interpolation techniques for the surface's variations and also add a texture variations model. Their goal is represent any face as a linear combination of a base group of face prototypes. The authors [2] used statistical measurements to maintain a credible deformation. The database used by Blanz and Vetter [2] consist of 200 scanned heads of young adults (100 males and 100 females), without facial hair, make up or accessories. The models obtained have 70,000 vertices for position and the same number for color values. Figure 3 shows the mean face that is the base to generate new faces.

Bronstein et al. [10] present a geometric method for facial recognition using data in 3D. Their method is based in geometric invariants of the human face and makes a comparison of non-rigid surfaces allowing typical deformations of the face due to facial expressions. This research explains that the human face cannot be considered as rigid object and based on empirical observations the authors conclude that facial

Figure 3. Mean face of *Morphable Face Model [2].*

expressions can be modeled as isometric transformations. These authors employed the bending invariant canonical forms as the center of their work in facial recognition. Regarding image acquisition the authors [10] used the coded light technique to obtain the geometry of the faces based on the pattern deformation. The images in 3D is transform in a triangulated surface by using splines, the facial surface is minimize to 2,500 vertices and then the canonical form is computed, the texture image can be mapped on the canonical surface. Interpolating the texture and form into a cartesian mesh they get flat representation for texture and shape as shown in Figure 4.

Figure 4. Flat texture in cartesian mesh, (A) texture mapped in the facial surface, (B) canonical form, (C) flat texture and (D) is the canonical image.

Bronstein et al. [10] used a database of faces consisting in 64 children and 93 adults, 115 males and 42 females; the database contained images of identical twins. The authors [10] analyzed four approaches: the eigendecomposition of images in 3D, the combination of texture and form image in eigenfaces, the decomposition of canonical images and the eigenforms algorithm. Bronstein et al. [10] concluded that their algorithm is able to extract the geometric features of facial surfaces, also that the use of canonical representations made robust their approach for recognition of facial expressions and that significant deformation of the face do not confuse their algorithm.

Golovinskiy et al. [11] presented an approach to model the geometry of small facial features like wrinkles and pores that appear on the skin surface. Golovinskiy et al. [11] proposed a statistical facial model that is able to extract, transfer and synthesize small facial features; their approach is based in the analysis of scans in high resolution which provided the facial geometry in 3D, also they used surface subdivision to separate the facial details. The next step was the parametric analysis to process the variations of the statistical model. Golovinskiy et al. indicated that in order to exploit the statistical information other techniques can be used like PCA. Basically their system consist of an analysis step and a synthesis step, the first begins with the face scan in high resolution 500 thousand polygons, in this step the statistical information is obtained; the synthesis step consist in adjust the image to form a new face. These authors mentioned that some applications of their method are the facial details analysis to study age effects and gender, as well the generation of new faces adding new details or age changes. Their experiments were on 149 subjects of different age, gender, race; the age range was 15 to 83 years, participate 114 males, 35 females and regarding races there were 81 Caucasians, 63 Asian and 5 African Americans. The mean face of the Golovinskiy et al. [11] is shown in Figure 5.

Figure 5. Mean face developed by Golovinskiy et al. [11] from which generate new faces with aging details.

In 2009 Paysan et al. [7] presented a generative model in 3D denominated Basel Face Model, this method can be used in facial recognition among other tasks. This model can adjust images in 2D and in 3D acquire under different situations and with different devices; the resulting parameters separate the lighting, position and identity parameters. According with Paysan et al. [7], Basel Face Model work independently texture and shape, building two linear models, the authors applied PCA to obtained a parametric facial model. The authors assumed that the data is normally distributed and the mean is computed correctly. For the correspondence step they modified ICP algorithm and then the faces were parameterized as triangular meshes with 53,490 vertices and 160,470 polygons, each vertex has associated a color value, then a face is represented by two vectors in 3D: shape and texture. The Figure 6 presents the mean face of this model.

Figure 6. BFM's Mean face [8].

Ramirez et al. [12] presented a model to generate multiple expressions from a neutral face, which is modeled with a 3 layers mesh to represent skin, fat tissue and the skull. To make the expressions in 3D, the model is complemented with the most important muscles that participate in facial expressions. The model proposed by Ramirez et al. [12] is divided in 3 principal steps: noise filtering, control points detection and register. The generic face model consists of 2,676 polygons and 2,777 vertices; the Figure 7 presents the generic face model of Ramirez et al. [12]. Ramirez et al. [12] mentioned that the facial muscles are defined as a vertices group, they work with 18 facial muscles and for the facial expressions simulation they used Facial Action Coding System (FACS) to describe and measure the facial behavior and sets a mechanism to classify facial expressions. After identify landmarks in generic face and in the input face the authors [12] proceed to transform the generic face model in the face of input using points for correspondence. To track changes on the generic model during the alignment process this phase is divided into three steps: global alignment, deformation in the xy-plane and deformation in the direction of z-axis. After the alignment process the generic facial model may be used to generate synthetic facial expressions. This synthetic scheme captures facial expressions, which are compared with training data to define how realistic they are.

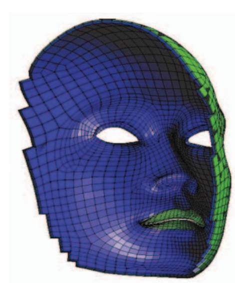


Figure 7. Ramirez's generic face model [12].

Phothisane et al. [4] present an alternative method for pose tracking using an exact 3D model extract from real human faces, they decimated the original model but keeping the landmarks. The model proposed by the authors [4] is based in Basel Face Model (BFM), the texture distribution was not used because the albedo of BFM does not support sensors classic cameras. To build a texture model, authors [4] compared the models with manually low-resolution face images of multiple databases (FacePix, CMU PIE, Yale face database). Texture is extracted by computing PCA from 474 different textures (captured on controlled lighting conditions) to initialize the position of the head after face detection. The authors [4] tested their system on the Boston database. Figure 8 shows a texture model with 3,000 vertices, the red points are the landmarks, which were manually indicated for illustration purpose.

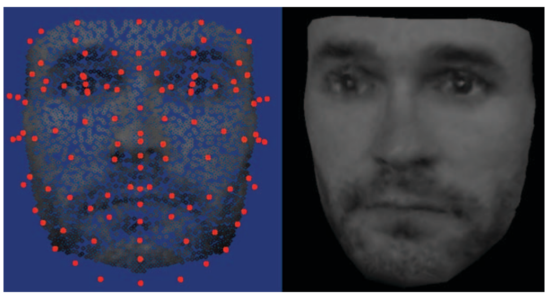


Figure 8. Phothisane et al [4] present an example of a texture model with 3000 vertices.

El-Hussuna [14] conducted a study that used facial anthropometric measurements and statistics to build a 3D model of facial Caucasian face. He mentioned that his study can be applied in plastic, aesthetic and reconstructive surgery. El-Hussuna [14] used scans of 8 Caucasian men whose ages are between 25 and 35, no mustache, beard, glasses and with neutral expression. He obtained two images of each subject, both images are aligned and the holes are filled. He used 64 landmarks for a simple display face through TCL/VTK software. After processing the scanned images El-Hussuna [14] proceeded to choose one of the scans to be the base mesh, the alignment vectors were used to create the point of finding dense correspondence points corresponding to those on the base mesh. This was achieved through the least Euclidean distance. The vertices were treated as pseudo-landmarks, the author [14] used Procrustes analysis technique followed by Principal Component Analysis to identify variations in his sample.

3. Discussion

During the development of this literary review it has been found that main applications of the 3D facial modeling is to test algorithms that perform tasks such as new faces generation, facial deformations analysis, face detection, face recognition, pose tracking, facial animation, expression generation and analysis, aging and de-aging faces to allow study how the face changes over time. Every application mentioned above is observed on disciplines as biometrics, forensics, plastic surgery, etc.

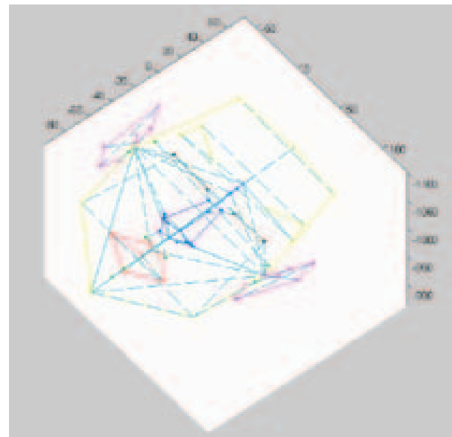


Figure 9. Mean face obtained by El-Hussuna [18], cyan lines represent the anthropometric distances and angles of interest in medical applications.

Although 2D/3D face modeling is not discussed in this paper. One should note that some facial models have evolved from 2D. For example, Chung Kuo et al. [15] presented a method to synthesize lateral face from only one 2D frontal face image. The depth information are estimated from this image based on the prior knowledge of anthropometry of the head and face. Abiantun et al. [16] create the model from 2D images generate the depth map (with the USF database HUMAN) and then employ 3D GEM (generic elastic model) to put the face in front position (pose correction), then apply face recognition and use PCA for reconstruction of the face.

Analyzing the characteristics of the facial models presented in this review, (Table 1) we can observe the evolution of the models reflected in the amount of vertices, polygons and the model's purpose, it is notable that the first models did not have a realistic representation of the face (Park [3], Candide [5,6]) fact that Blanz and Better [2] and Paysan et al [7] improve in a significant way.

Two factors of interest are the model resolution and the amount of polygons that will represent the face, here is mandatory the detail level required by of researchers in order to represent faces that may be perceived as real.

Another fact that is noted is the amount of subjects, noting that the human face is a flexible surface and each image even from the same subject is different.

For face modeling, some researchers used by Basel Face Model and others decrease resolution for optimizing.

We also found that most of face models are computed from images of adult subjects, hence, it is open a research using subjects from a wider range of ages. By doing this, it would be relevant to consider users from different races to compute a most generic human face model.

The anthropometric measures based approach (El-Hussuna [14]) results important to consider in the construction of a facial model especially when the applications are medical or biometrics.

We are aim to produce a useful 3D face model for research in key areas based on Mexican anthropometric analysis as part of our future work, so we conclude that in order to development a new facial model there are some characteristics to consider. For example, the face model's purpose, facial landmarks to consider, model resolution, characteristics of the subjects of study, specify if the subject may wear accessories, number of subjects that provide the facial information and images per subject.

Acknowledgments

This research is sponsored by the Research Department of the Autonomous University of the State of Mexico and CONACyT, grants 3720/2014/CID and 553366, respectively.

References

[1] Hernández E. O. "Valoración directa de las características antropométricas cráneo-faciales de la población del estado de Michoacán, México". Universidad Michoacana de San Nicolás de Hidalgo, México. Octubre 2011.

[2] V. Blanz, T. Vetter. "A morphable model for the synthesis of 3d faces". SIGGRAPH '99 Proceedings of the 26th annual conference on Computer graphics and interactive techniques Pages 187-194.

[3] Park, F. I. "A parametric model for human faces". 1974, Universidad de Utah, USA.

[4] Phothisane, P., Bigorgne, E., Collot, L., Prévost, L. "a robust composite metric for head pose tracking using an accurate face model". 2011, Automatic Face & Gesture Recognition and Workshops (FG 2011), IEEE International Conference.

[5] "Candide: a parameterized face". Url: http://www.icg.isy.liu.se/candide/

[6] Jörgen Ahlberg, "Candide 3 a updated parameterized face". Image Coding Group, Dept. of Electrical Engineering, Linköping University, Linköping Sweden.

[7] Pascal Paysan, Reinhard Knothe, Brian Amberg, Sami Romdhani and Thomas Vetter (2009). "A 3d model for pose and illumination invariant face recognition". In proceedings of the 6th IEEE International Conference, Advanced Video and Signal Based Surveillance, Italy.

[8] Basel Face Model, url: http://faces.cs.unibas.ch/bfm/main.php?nav=1-1-0&id=details.

[9] Conde, C. "verificación facial multimodal 2D y 3D". 2006. Tesis Doctoral. Departamento de Arquitectura y tecnología de computadores, ciencia de la computación e inteligencia artificial. Universidad Rey Juan Carlos.

[10] Bronstein, A., Bronstein, M., Kimmel, R., "Expression invariant 3d face recognition". (2003). Springer-Verlag Berlin Heidelberg 2003.

[11] Golovinskiy, A., Wojciech Matusik, Hanspeter Pfister, Szymon Rusinkiewicz, Thomas Funkhouser. "A statistical model for synthesis of detailed facial geometry", (2006) ACM Transactions on Graphics (Proc. SIGGRAPH), July 2006.

[12] Ramírez Valdez Leonel, Hasimoto Beltran Rogelio. "3D-facial expression synthesis and its application to face recognition systems". Journal of Applied Research and Technology 2009, Redalyc.

[13] Cootes, T.F., Taylor, C.J. "Statistical models of appearance for computer vision". Imaging Science and Biomedical Engineering, University of Manchester. 2004.

[14] El-Hussuna, A. "Statistical variation of three dimensional face models". Master Thesis. IT-University of Copenhagen, Multimedia Technologies. March 2003.

[15] Chung J. Kuo, Ruey S. Huang, and Tsang G. Lin. "Synthesizing lateral face from frontal facial image using anthropometric estimation". 1997. Image Processing, 1997. Proceedings, International Conference on Vol.1, pp 133 - 136.

[16] Abiantun, R.; Prabhu, U.; Savvides, M., "Sparse feature extraction for pose-tolerant face recognition". Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol. 99.