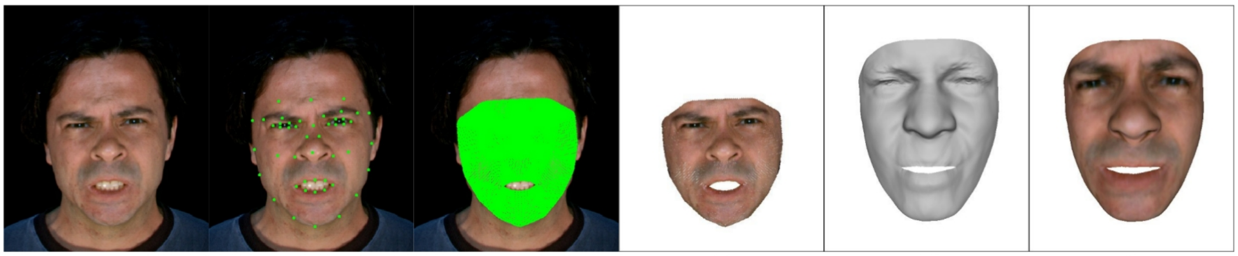
3D human face reconstruction using principal components spaces

Jesús P. Mena-Chalco[1], Luiz Velho[2], Roberto M. Cesar Junior[1]

[1] Institute of Mathematics and Statistics, University of São Paulo, São Paulo - Brazil
[2] National Institute of Pure and Applied Mathematics, Rio de Janeiro - Brazil
jmena@vision.ime.usp.br, lvelho@impa.br, cesar@ime.usp.br



(a) (b) (c) (d) (e) (f)

Figure 1. Overview of our method: from a 2D photo and their corresponding facial landmarks (a)-(b), the facial texture data is extracted by successive 2D triangular subdivisions (c)-(d), producing a new 3D face model (e)-(f).

***Abstract*—In this work we propose a new method of *3D face computational photography* based on a facial expressions training dataset composed of both facial range images (3D geometry) and facial texture (2D photo). The method allows to obtain a 3D representation of facial geometry given only a 2D photo and a set of facial landmarks, which undergoes a series of transformations through the estimated texture and geometry spaces. In the training stage, principal component analysis is used to represent the face dataset, thus defining an orthonormal basis of texture and another of geometry. In the reconstruction stage, an input is given by a 2D face image and their corresponding landmarks. This data is fed to the PCA basis transform, and a 3D version of the 2D input is built. Several tests using a 3D faces dataset, together with the adoption of a metric, show good results in the 3D facial reconstruction. Additionally, we explored two applications related to the facial expressions transferring and caricaturization. The results of these applications show a rapid and simple synthesis of new 3D models with new expressions and exaggerated facial proportions, useful for 3D facial animation. Results and demos are available at www.vision.ime.usp.br/~jmena/projects/3Dface**

***Keywords*-3D face reconstruction; principal components analysis; computer graphics**

I. Introduction

3D face models have an important role in various applications of *Computer Vision*, *Computer Graphics* and *Pattern Recognition*. Generally, we can divide the applications related to 3D face models into two types: (i) face analysis, which uses computational methods that allow to characterize models of real-world faces and (ii) face synthesis, which use computational methods to build or to generate face models.

(*) This manuscript is a summary of the PhD project *"Reconstrução de faces 3D através de espaços de componentes principais"* defended at December 14, 2010.

The 3D face reconstruction, as presented here, belongs to the type of face synthesis, in which, given a 2D input image, its corresponding 3D version is generated. The reconstruction process of 3D facial models is a relevant topic that has received special attention within the research community. This is an example of the so-called *3D facial computational photography*, where methods of Computer Vision and Computer Graphics are used to solve a problem related to synthesis of faces.

For the face synthesis, many methods of reconstruction and 3D facial animation have been proposed (e.g. [1]–[5]). The approaches considered in this category (i) are based on the location of facial landmarks, which are used to create correspondence between the 2D input face and the face model learned in the training stage, and (ii) suppose that the input face belongs to only one person with the neutral facial expression.

Despite more than three decades of research [6], [7] there are still some important 3D face photography open problems. In the literature, the 3D face reconstructions from single photos with different facial expression have been little explored (see an example in Fig. 1a). This limitation is especially due to the difficulty of calculating correspondences between the 3D geometry data and the facial expressions, and the lack of 3D face datasets with different facial expressions. Accordingly, in this work, we address the problem of realistic 3D face reconstruction from a single 2D photo, without considering the restriction on photos with different facial expressions.

Our approach employs principal components analysis (PCA) to represent the face model (texture and geometry separately) of 3D faces with different facial expressions. The PCA face model is composed by two separate orthonormal basis which represent texture and geometry, respectively. The reconstruction is possible because we assume that a 3D human face, with any facial expression, can be modeled/reconstructed using a combination of other 3D faces belonging to the training set. Our approach works with few training samples.

It is important to note that the proposed method considers the correct location of 48 facial landmarks in the 2D photo given as input (Fig. 1b). Certainly, this task can be solved in different ways. In particular, we could automate this task using a method based on Active Shape Models [8] or Active Appearance Models [9]. The proposal was developed and implemented in a prototype system, using few samples in the training stage and simple algorithms of Computer Vision and Computer Graphics. Thus, opportunities for future improvements can be defined and implemented in order to obtain more realistic face reconstructions.

Research objective and contributions

The main research objective of our study was to develop a method for reconstruction of 3D faces given a single 2D photo. The proposed method is based on spaces generated by principal component analysis using a dataset of real 3D faces, so that the reconstruction is performed quickly and the result is simple and realistic.

The contributions were the following:

- Study of the main methods of 3D face synthesis that constitute the state-of-the-art of face reconstruction based on real data.
- Proposition and evaluation of a semi-automated method for reconstruction of faces given a single 2D photo (main contribution of our work).
- Creation of a data acquisition protocol and a normalized 3D facial expressions dataset.
- Study and extension of a method for caricaturization of faces, from the neutral expression, exploring the proportions of facial elements.

II. IMPA-FACE3D: A DATASET OF FACIAL EXPRESSIONS

Current 3D face datasets usually include samples with pose variation and few acquisitions of facial expressions that allow to investigate spontaneous changes of human expressions. For this reason, in 2008, we created a dataset of 3D faces, called IMPA-FACE3D [10] in order to assist the research related to analysis and synthesis of neutral faces and six universal expressions among the human races proposed by Ekman [11]: joy, sadness, surprise, anger, disgust and fear. The data is hence composed by registered texture (color) and geometry data. Therefore, a depth value is associated with each pixel of the texture images. Devices that provided controlled lighting conditions were used. The digital camera had the white balance properly adjusted[1].

This dataset includes acquisition of 38 subjects. During acquisition, subjects were in frontal position. They were not allowed to use eyeglasses or other objects that may modify the face appearance, but there were no restrictions on clothing, hairstyle of haircut. The dataset also includes another 5 strongly differentiated facial expressions that may be explored in future applications: mouth and eyes open, mouth and eyes close, kiss expression, and wink expression. Altogether, the dataset is composed of 22 men and 16 women. 14 samples were acquired for all individuals, summarizing 532 samples in total.

Once the data has been acquired, procedures were used to eliminate noise originating by specular reflections of illumination in the face. Additionally, the problems of self-occlusion and discontinuities (holes) were treated with techniques of image processing. Finally, facial landmarks were manually placed over the facial geometry. We adopted an initial face model with 48 landmarks and a triangulation that contains 74 elements (triangles).

III. 3D FACE RECONSTRUCTION USING PRINCIPAL COMPONENTS SPACES

Starting from the work of Vlasic *et al.* [12] and Macêdo *et al.* [13] a method for 3D face reconstruction from 3D color images using a small training dataset has been created. It is worth noting that these previous works [12], [13] do not explore 3D data. The training set is composed by a small set of range images (registered texture and geometry data) corresponding to seven different facial expression.

Our approach uses PCA to represent 3D faces with texture and space geometry (both correlated but analyzed separately). In the reconstruction stage, it makes use of a parameterized 3D face model composed of two separated orthonormal basis that represent the texture and geometric data, respectively.

Given as input a 2D color image of a human face to be 3D reconstructed, and its corresponding facial landmarks (initial model), texture information is extracted following a triangular subdivision strategy using the initial face model. For convenience, we used the same subdivision strategy adopted in the training stage (See Fig. 1c-d), while the facial geometry is obtained by projecting the texture information onto the space of geometry obtained in the training stage. This projection is performed using a vectorial basis (estimated by PCA) and a linear optimization function to relate the 2D texture information with its corresponding 3D geometry.

This relation is feasible because, intuitively, a human face can be modeled by a combination of faces from different individuals with different facial expressions. Finally, the texture information is directly mapped to the obtained 3D facial geometry (Fig. 1e-f). See in Figure 2 a flowchart of the proposed method. For the training stage, 2D photos, their corresponding 3D facial geometries and 3D facial landmarks are used. In the reconstruction stage only a 2D photo is considered, with its corresponding 2D facial landmarks.

Note that the facial landmarks considered into the training stage are related to the 3D facial geometry. However, the reconstruction stage consider 2D facial landmarks (2D points associated with the 2D photos). An initial version of this reconstruction approach was presented in the SIBGRAPI [14], and an expanded version was published in [15].

[1]See in Table 3.1 of the PhD thesis a description of current 3D face datasets.

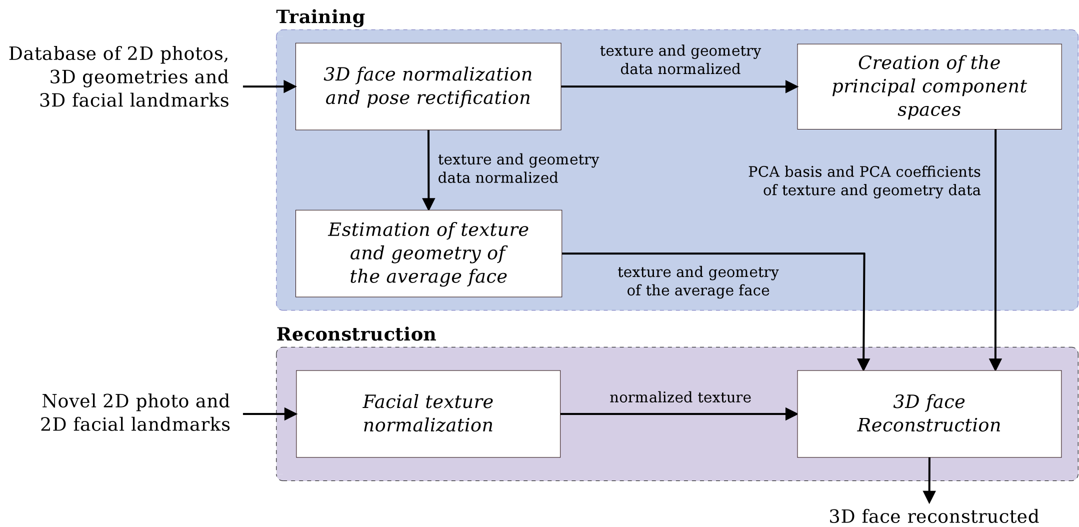


Figure 2. Schematic data flow diagram of the 3D facial reconstruction system (*dotted lines*). Each block (*solid lines*) represents a process, while each arrow represents the information flow between processes.

A. Mathematical model overview

The list below summarizes the symbols used in the work, being presented to help the reader:

- $\mathbf{X}_i = (\mathbf{x}_i^t, \mathbf{x}_i^g)$: the i-th 3D face composed by texture and geometry data;
- l_i: landmarks of the i-th 3D face;
- L_i^t, L_i^g: texture and geometry of the i-th 3D face, normalized, cropped and with rectified pose, respectively;
- E^t, E^g: texture and geometry PCA basis, respectively;
- t_0, g_0: texture and geometry of the average face;
- α^t, α^g: texture and geometry coefficients expressed in terms of E^t, E^g, respectively;
- s_x: weighting coefficients of $\mathbf{x}^t$ in terms of the training samples;
- $\mathbf{x}^t$, $\mathbf{x}^g$: normalized input texture face and the corresponding output reconstructed geometry.

The proposed approach is based on learning a 3D face model using texture and geometry of a training face for some different facial expressions. An input 2D face image (i.e. a color photo) is then reconstructed by projecting it on the trained 2D texture space, decomposing it as weights of the training samples. The obtained weighting coefficients are then used to build a 3D model from the 3D training geometry samples.

The training dataset is composed by pairs of texture and geometry data from subjects with seven different facial expressions (six universal facial expressions plus a neutral pose). A set of landmarks is placed on the texture image and used to represent the input texture information. Therefore, the i-th facial texture, $i = 1, \ldots, N$, is represented by L_i^t composed by the texture information obtained as result of the face normalization. In the same way, the facial geometry is represented by L_i^g, composed by the geometry information. Therefore, for the i-th sample consisting of M vertices:

$$L_i^t = [r_{i1}, \ldots, r_{iM}, g_{i1}, \ldots, g_{iM}, b_{i1}, \ldots, b_{iM}]$$

$$L_i^g = [x_{i1}, \ldots, x_{iM}, y_{i1}, \ldots, y_{iM}, z_{i1}, \ldots, z_{iM}]$$

where (x_{ij}, y_{ij}, z_{ij}) and (r_{ij}, g_{ij}, b_{ij}) represent the spacial position and the RGB value of the j-th landmark, $j = 1, \ldots, M$, belonging to the i-th sample. The training stage consists in defining good texture and geometry space representations based on a given set of facial expression samples from several subjects. Thus, texture and geometry are obtained from N samples of facial expression of different subjects, being denoted as $\{L_1^t, L_2^t \ldots, L_N^t\}$ and $\{L_1^g, L_2^g \ldots, L_N^g\}$, respectively. This data helps to define the initial texture and geometry spaces.

In order to have a more efficient and statistically optimized representation, both texture and geometry spaces are PCA-transformed [16]. Each training sample represents a vector expressed in these spaces.

The main goal of our approach is to obtain a geometry representation of a given face $\mathbf{x}$, provided as a texture image (Fig 1a). A set of facial landmarks is manually positioned, and a normalized input texture $\mathbf{x}^t$ is extracted from such input image (Fig 1d), and undergoes a series of transformations through the texture and geometry spaces. The final result is the reconstructed geometry of the input face $\mathbf{x}^g$ (Fig 1e), i.e. a point in the geometry space.

B. Training

The training stage is composed by three phases. Firstly, the training dataset is normalized, considering the initial facial landmarks composed by 48 elements. This procedure allows to establish a *dense correspondence* among the 3D faces of training dataset (the normalization allows to correlate the 3D faces in the training dataset). The facial landmarks positioned in all 3D faces allow to establish an homology (relationship) among the individual representations. This fact also provides a simple way to create an average 3D face using a triangulation, i.e., texture (t_0), and geometry (g_0) of the average 3D face. Note that the *dense correspondence*, as considered in this work, allows to use the same triangulation on all models of training dataset. Thus, equivalent meshes are obtained. For practical purposes, we consider a 3D triangulation of 18 944 elements, as product of successive triangular subdivisions of the initial face model: two linear subdivisions followed by two Loop subdivisions. The subdivision strategy allows to accomplish the resulting meshes into per-vertex correspondences yielding, in each iteration, face models with refined approximations of the raw mesh.

Two PCA procedures are carried out separately for the geometry L_i^g and for the texture L_i^t data. This analysis leads to:

- An orthonormal basis ($E^t = \{e_i^t\}$) for the facial texture space, and the coefficients ($\{\alpha_i^t\}$) for each texture image in the training dataset expressed w.r.t. $\{e_i^t\}$.
- An orthonormal basis ($E^g = \{e_i^g\}$) for the facial geometry spaces and the coefficients ($\{\alpha_i^g\}$) for each geometry data in the training dataset expresses w.r.t. $\{e_i^g\}$.

In order to work with the same number of principal components in the aforementioned spaces, we use the minimum amount of components representing a pre-defined amount of total variance kept. The results shown in this work were drawn from those in which principal components kept at least 95% of the total variance of both texture and geometry data.
The following algorithm summarizes the training procedure:

FACE3D-TRAINING$\big($**X**=$\{(\mathbf{x}_1^t,\mathbf{x}_1^g),\ldots,(\mathbf{x}_N^t,\mathbf{x}_N^g)\}$, *triangulation*, *landmarks*=$\{l_1,\ldots,l_N\}\big)$

$\{L^t,L^g\} \leftarrow$ NORMALIZATION(**X**, *landmarks*, *triangulation*)
$\{t_0,g_0\} \leftarrow$ MEAN(L^t, L^g)
$\{(E^t,\alpha^t), (E^g,\alpha^g)\} \leftarrow$ PCA(L^t, L^g)

The procedure MEAN calculates the average texture model, t_0, and the average geometry model, g_0, based on L^t and L^g, respectively. The procedure PCA allows one to perform an analysis based on principal components using the texture and geometry data, generating two basis (E^t, E^g) and their corresponding texture and geometry coefficients (α^t,α^g).

C. Reconstruction

The input is a 2D photo (frontal face image) and their corresponding facial landmarks. The texture normalization of input is done through an iterative process similar to one adopted in the normalization of training dataset. The process begins with 48 landmarks and a 2D representation of the 3D triangulation used in training stage. Afterwards, four triangular subdivisions are made: two linear subdivisions followed by two Loop subdivisions.

The 2D positions of the new landmarks are used to extract the texture information corresponding to the facial region (see an example in Figs.1c-d). Note that we use the same strategy of triangular subdivision to maintain a direct relationship between the 2D normalized texture input vector and the normalized data calculated in the training stage.

Let $\mathbf{x}^t$ be the normalized texture of the input image $\mathbf{x}$, and t_0 the average texture obtained in training stage. The texture coefficients α_x^t, are calculated by projecting $(\mathbf{x}^t - t_0)$ onto the respective orthonormal basis ($\{e_i^t\}$):

$$\alpha_x^t = E^t.(\mathbf{x}^t - t_0) \tag{1}$$

where E^t is a transformation matrix defined by the orthonomal basis for the texture space learned in the training stage.

Once the texture coefficients α_x^t are obtained, the texture coefficients α^t of all images considered in the training stage are used to calculate the coefficients s_x, defined as

$$\alpha^t.s_x = \alpha_x^t \tag{2}$$

where α^t is the matrix defined by the coefficients for each texture image in the training dataset. The coefficients s_x can be calculated through a regression procedure or by least square estimation, such that $||\alpha_x^t - \alpha^t.s_x||$ is minimized.

Intuitively, s_x represents weighting coefficients obtained by projection α_x^t onto α^t. It is important to recall that each sample represented in α^t is associated to a subject of the considered training data. Therefore, s_x represents the decomposition of $\mathbf{x}^t$ in terms of different facial expressions belonging to several subjects of training data. So, we could say that a human face can be modeled by a linear combination of different facial expressions of several subjects.

The geometry coefficients α_x^g of $\mathbf{x}^t$ are then calculated using the geometry coefficients of all training geometry samples α^g:

$$\alpha_x^g = \alpha^g.s_x \tag{3}$$

Finally, the normalized geometry $\mathbf{x}^g$ of the input image $\mathbf{x}$ is then reconstructed by

$$\mathbf{x}^g = (E^g.\alpha_x^g) + g_0 \tag{4}$$

where E^g is a transformation matrix defined by the orthonormal basis for the geometry space learnt in the training stage.

As mentioned above, the color values (r, g, b), for each element of the texture vector have been extracted from the positions (x, y) estimated by successive subdivision of the initial 2D face model. In doing so, we use the positions (x, y) of the 2D face model to correct/rectify the (x, y) positions of the obtained vector $\mathbf{x}^g$, in order to keep a more realistic 3D face reconstruction. Currently, the correction is carried out by a simple linear weighting between the (x, y) values of the initial 2D face model and the (x, y) values of the reconstructed face.

In order to reduce the noise on the reconstructed facial geometry (surface), a Laplacian filter for smoothing polygonal surface meshes has been applied. Finally, the input texture $\mathbf{x}^t$ is directly mapped onto the resulting 3D smooth geometry $\mathbf{x}^g$.
The following algorithm summarizes the proposed approach:

FACE3D-RECONSTRUCTION$\big($ $\mathbf{x}$, $l_\mathbf{x}$ $\big)$

$\mathbf{x}^t \leftarrow$ TRIANGULAR-SUBDIVISION($\mathbf{x}$, $l_\mathbf{x}$)
$\alpha_x^t \leftarrow E^t.(\mathbf{x}^t - t_0)$
$s_x \leftarrow$ LSQR(α_x^t, α^t)
$\alpha_x^g \leftarrow \alpha^g.s_x$
$\mathbf{x}^g \leftarrow (E^g.\alpha_x^g) + g_0$
$\mathbf{x}^g \leftarrow$ GEOMETRIC-CORRECTION($\mathbf{x}^g$)
$\mathbf{x}^g \leftarrow$ GEOMETRIC-SMOOTHING($\mathbf{x}^g$)
return ($\mathbf{x}^t$, $\mathbf{x}^g$)

The procedure TRIANGULAR-SUBDIVISION allows to normalize the input texture $\mathbf{x}$ using the corresponding 2D facial landmarks $l_{\mathbf{x}}$. The procedure LSQR allows one to represent α_x^t in terms of α^t, as indicated in Eq. (2).

Measuring accuracy of 3D facial reconstructions

The quality of 3D face reconstructions based on real samples has been little explored by researchers. Evaluations of reconstructed models are usually conducted in a subjective way and are based only in visual inspection between the reconstructed 3D face and the real 3D facial geometry.

Our proposed measure corresponds to the estimation of similarity between the real 3D face model and the reconstructed 3D face model. The defined procedure consists of two steps: (i) alignment of both facial models in order to register the two points' cloud that make up the 3D faces, and (ii) estimation of the distances between the real 3D face model and the reconstructed 3D face model. For the second step, we considered the projection algorithm of normal vectors of the reconstructed model onto the real 3D face model. This measure was called *projection distances map* and allows to numerically compare 3D facial models.

IV. TRANSFER OF FACIAL EXPRESSIONS AND CARICATURIZATION

The face normalization performed in the training stage was based on the processes of 3D triangular subdivision and projection of the normal vectors of each vertex of the face model. This normalization allows to maintain a *dense correspondence* (correspondence vertex-to-vertex) among all the 3D face model of the considered dataset. In this sense, the i-th vertex of any two face models corresponds to the same facial region.

In our work, given the *dense correspondence* between the 3D faces, we can easily manipulate the facial expressions of any 3D face. The proposed approach consists of two steps: In the first step, we use the neutral face as well as the six universal facial expressions of individuals in order to calculate the displacement vectors (the expression transferring for a 3D face is made considering as displacement source the average faces of each facial expressions obtained from the dataset). In the second step, these displacement vectors are used to apply a new expression (or combination of expressions) in the 3D facial geometry.

For automatic creation of new caricatures, we presented an adaptation of the work of exaggeration of facial relations proposed by Chen *et al.* [17]. The method, based on the exaggeration of dissimilarities with respect to an average face, considers the calculation of a proportions vector of facial elements (e.g. distance between eyes, mouth). This obtained proportions vector is compared with the proportions vector of an average face. Thus, the facial elements of the input face with 'different' proportions to the average face are the main candidates to be changed, and thereby shown in evidence.

V. OBTAINED RESULTS

Different experiments were conducted to evaluate the quality of 3D facial reconstruction generated by our proposed method. Figure 3 shows some of the results obtained considering subjects present in the training dataset (column 1), subjects not present in the training dataset and with controlled lighting (columns 2 and 3), and subjects who are not present in the training and without controlled lighting (columns 4-6). As can it be observed, the method was able to satisfactorily reconstruct the 3D facial geometry of all subjects, keeping for each expression, a spatial coherence[2]. Note that the projection distances map is represented with a color scale that indicates the similarity degree among facial regions. Cold colors indicate low values, while hot colors indicate high values of projection distances. In column 3 we show an example of reconstruction with non-coherent distortion in the region near the mouth. This phenomenon is due to the use of a photo belonging to a person with mustache (this feature alters texture regions used in the reconstruction method).

Finally, in columns 4-6 of Figure 3c only the 3D face geometry obtained by our method is shown. Here, actually, there is no form to numerically compare 3D face geometries, except for subjective visual comparison (we do not have the 3D face models of these famous people).

Figure 4 shows an example of caricature and transfer of 3D facial expression. By applying the adopted approach, the more distinctive features, with respect to average face, are modified. For example, if a person's nose is thin, then its caricaturization generates a face deformation, such that the nose is even thinner. On the other hand, if the length of a person's chin is large, when compared to an average face, then his caricature will correspond to a face with a greater chin.

VI. CONCLUDING REMARKS

In this work, we have described a method for reconstructing the 3D geometry of human faces from 2D monocular color photos. This task is accomplished through a "resynthesis by analysis" approach, in which models for human facial texture and geometry spaces are built from acquired real-world data, being subsequently used to perform 3D face reconstruction.

Additionally, using the dense correspondence among the 3D facial geometry of the dataset, we present two applications concerning expressions facial transfer and 3D face caricaturization. The obtained results showed good synthesis of facial geometry, useful for future facial animations and characterizations.

Additional information: Our collaborative work has resulted in a journal paper [15], three refereed conference papers [14], [18], [19], a technical report [10] and a conference poster [20]. The PhD thesis manuscript, dataset, demos and several results are available at www.vision.ime.usp.br/~jmena/projects/3Dface

[2]It is understood as spatial coherence the correct visual appearance of facial elements (e.g. eyes, nose, mouth), when confronted with the photo used as input data.

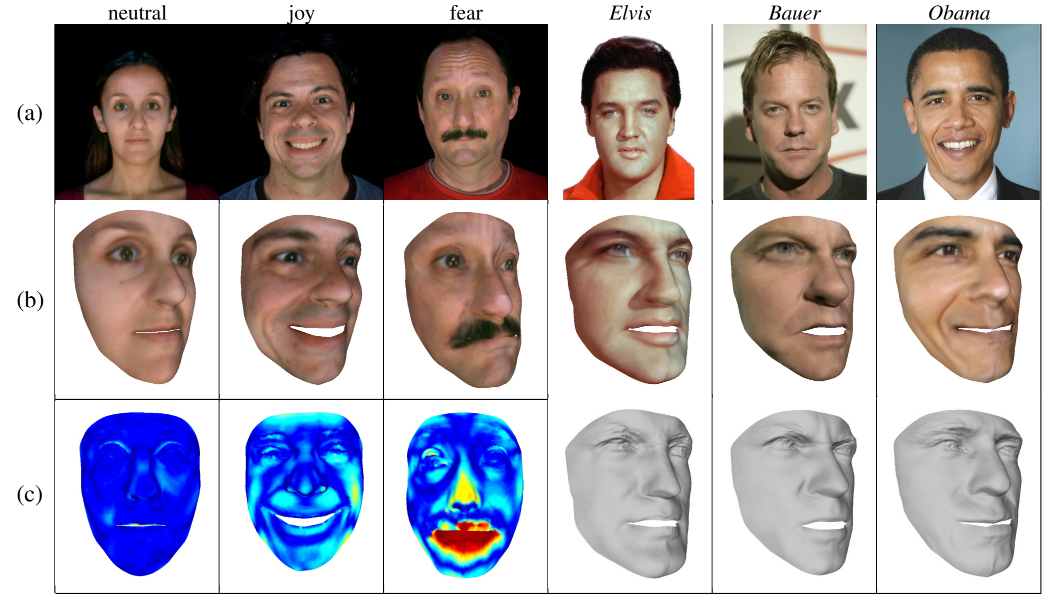


Figure 3. 3D reconstruction of different subjects: (a) 2D photos, (b) 3D geometry with the mapped texture, and (c) projection distances map.

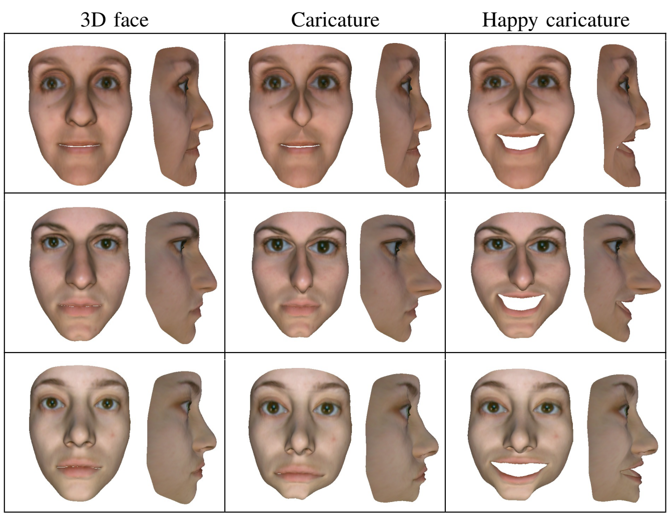


Figure 4. Caricature and transfer of facial expression. For each 3D face has been calculated a proportions vector of facial elements. This vector can be used to decide which facial proportion could be modified.

ACKNOWLEDGMENT

The authors express their gratitude to Carlos Hitoshi Morimoto, Ives Macêdo, Marcio Lobo Netto and Maria Cristina Ferreira de Oliveira for their valuable discussions and comments. This research was supported in part by CAPES, CNPq, FAPESP and FINEP.

REFERENCES

[1] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces," in *Computer graphics and interactive techniques*, 1999, pp. 187–194.

[2] D. L. Jiang, Y. X. Hu, S. C. Yan, L. Zhang, H. J. Zhang, and W. Gao, "Efficient 3D reconstruction for face recognition," *Pattern Recognition*, vol. 38, no. 6, pp. 787–798, June 2005.

[3] S. Park, J. Heo, and M. Savvides, "3D face reconstruction from a single 2D face image," in *Conference on Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08*, 2008, pp. 1–8.

[4] X. Gong, G. Wang, and L. Xiong, "Single 2D image-based 3D face reconstruction and its application in pose estimation," *Fundamenta Informaticae*, vol. 94, no. 2, pp. 179–195, 2009.

[5] W. Widanagamaachchi and A. Dharmaratne, "3D face reconstruction from 2D images," *Digital Image Computing: Techniques and Applications*, pp. 365–371, 2008.

[6] T. Kanade, "Picture processing system by computer complex and recognition of human faces," in *Doctoral dissertation, Kyoto University*, 1973.

[7] J. V. Kittler, A. Hilton, M. Hamouz, and J. Illingworth, "3D assisted face recognition: A survey of 3D imaging, modelling and recognition approaches," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 3, 2005, pp. 114–114.

[8] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models: Their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, January 1995.

[9] G. Dedeoglu, S. Baker, and T. Kanade, "Resolution-aware fitting of active appearance models to low resolution images," in *European Conference on Computer Vision*, 2006, pp. 83–97.

[10] J. P. Mena-Chalco, R. Cesar-Jr, and L. Velho, "Banco de dados de faces 3D: IMPA-FACE3D," IMPA - RJ, Tech. Rep., 2008.

[11] P. Ekman and W. Friesen, "Constants across cultures in the face and emotion," *Journal of Personality and Social Psychology*, vol. 17, no. 2, pp. 124–129, 1971.

[12] D. Vlasic, M. Brand, H. Pfister, and J. Popović, "Face transfer with multilinear models," *ACM Transactions on Graphics*, vol. 24, no. 3, pp. 426–433, 2005.

[13] I. Macêdo, E. V. Brazil, and L. Velho, "Expression transfer between photographs through multilinear AAM's," in *SIBGRAPI*, 2006, pp. 239–246.

[14] J. P. Mena-Chalco, I. Macêdo, L. Velho, and R. M. Cesar-Jr, "PCA-based 3D face photography," in *SIBGRAPI*, 2008, pp. 313–320.

[15] J. P. Mena-Chalco, I. Macêdo, L. Velho, and R. M. Cesar-Jr, "3D face computational photography using PCA spaces," *The Visual Computer*, vol. 25, no. 10, pp. 899–909, 2009.

[16] I. T. Jolliffe, *Principal Component Analysis*, ser. Series in Statistics. Springer Verlag, 2002.

[17] W. Chen, H. Yu, and J. Zhang, "Example based caricature synthesis," in *CASA 2009*, 2009.

[18] A. Mattos, J. P. Mena-Chalco, R. M. Cesar-Jr, and L. Velho, "3D linear facial animation based on real data," in *SIBGRAPI*, 2010.

[19] S. C. D. Pinto, J. P. Mena-Chalco, F. M. Lopes, L. Velho, and R. M. Cesar-Jr., "3d facial expression analysis by using 2d and 3d wavelet transforms," in *ICIP*, Belgium, 2011, pp. 1293–1296.

[20] S. C. D. Pinto, J. P. Mena-Chalco, F. M. Lopes, and R. M. Cesar-Jr, "Preliminary results on 3D facial expression analysis using 2D and 3D wavelet transforms," WIP - SIBGRAPI, 2010.