Head Modeling from Pictures and Morphing in 3D with Image Metamorphosis based on triangulation

WON-SOOK LEE, NADIA MAGNENAT THALMANN

MIRALab, CUI, University of Geneva
24, rue General-Dufour, CH-1211, Geneva, Switzerland
Tel: +41-22-705-7763 Fax: +41-22-705-7780
E-mail: {wslee, lee, thalmann}@cui.unige.ch

Abstract. This paper describes a combined method of facial reconstruction and morphing between two heads, showing the extensive usage of feature points detected from pictures. We first present an efficient method to generate a 3D head for animation from picture data and then a simple method to do 3D-shape interpolation and 2D morphing based on triangulation. The basic idea is to generate an individualized head modified from a generic model using orthogonal picture input, then process automatic texture mapping with texture image generation by combining orthogonal pictures and coordinate generation by projection from a resulted head in front, right and left views, which results a nice triangulation on texture image. Then an intermediate shape can be obtained from interpolation between two different persons. The morphing between 2D images is processed by generating an intermediate image and new texture coordinate. Texture coordinates are interpolated linearly, and the texture image is created using Barycentric coordinates for each pixel in each triangle given from a 3D head. Various experiments, with different ratio between shape, images and various expressions, are illustrated.

Keywords: clone, a generic model, automatic texture mapping, interpolation, Barycentric coordinate, image morphing, and expression.

1. INTRODUCTION

Ask any animator what is the most difficult character to model and animate, and nine out of ten may respond "humans". There is a simple reason for that: we all know what humans are supposed to look like; we are all experts in recognizing realistic person.

In this paper, we describe a method for individualized face modeling and morphing them in 3D with texture metamorphosis. There are various approaches to reconstruct a realistic person using a Laser scanner [7], a stereoscopic camera [2], or an active light stripper [10]. There is also an approach to reconstruct a person from picture data [6][12]. However most of them have limitation when compared

practically to a commercial product (such as a camera) for the input of data for reconstruction and finally animation.

Other techniques for metamorphosis, or "morphing", involve the transformation between 2D images [14][17] and one between 3D models [15][16] including facial expression interpolation. Most methods for image metamorphosis are complicated or computationally expensive, including energy minimization and free-form deformation.

We present a method not only for reconstruction of a person in 3D for animation but also for morphing them in 3D. Our reconstruction method makes morphing between two people possible through 3D-shape interpolation based on the same topology and 2D morphing for texture images to get an intermediate head in 3D.

We first introduce, in Section 2, a fast head modeling method to reconstruct an individualized head modified from a generic head. In Section 3, texture mapping is described in detail how to compose two images and obtain automatic coordinate mapping. Then Section 4 is devoted to image morphing based on triangulation. In following Section 5, several results are illustrated to show 3D realistic morphing among several people. It includes shape interpolation, image morphing and expression interpolation. Finally conclusion is given.

2. Face modeling for individualized head

In this section, we present a way to reconstruct a head for animation from orthogonal pictures, which looks photo-realistic. First, we prepare a generic model with animation structure and two orthogonal pictures of the front and side views. The generic model has efficient triangulation, with finer triangles over the highly curved and/or highly articulated regions of the face and larger triangles elsewhere, that includes eyes and teeth.

The main idea to get an individualized head is to detect feature points on the two images and obtain 3D position of the feature points to modify a generic model using a geometrical deformation. The feature detection is processed in a semiautomatic way using the structured snake method with some anchor functionality as described in paper [12].

Figure 1 (a) shows an orthogonal pair of images. Detected features on a normalized image pair are also shown in **Figure 1** (b). Feature points are overlaid on image even though the space has different origin and different scaling.

Then two 2D position coordinates in front and side views, which are the (x, y) and the (z, y) planes, are combined to be a 3D point. First, we use a global transformation to move the 3D feature points to the space for a generic head. Then Dirichlet Free Form Deformations (DFFD) [8] are used to get new geometrical coordinates of a generic model adapting to the detected feature points. Then the shapes of the eyes and teeth are recovered to the original shape with translation and scaling adapted to a new head. The control points for the DFFD are feature points detected from the images. As shown in **Figure 2**, it is a rough matching method that does not get the exact points for every point except feature points. However, it is useful to reduce the

data size of a head to accelerate animation speed. To get it to be realistic looking, we use automatic texture mapping, which is described in below section.
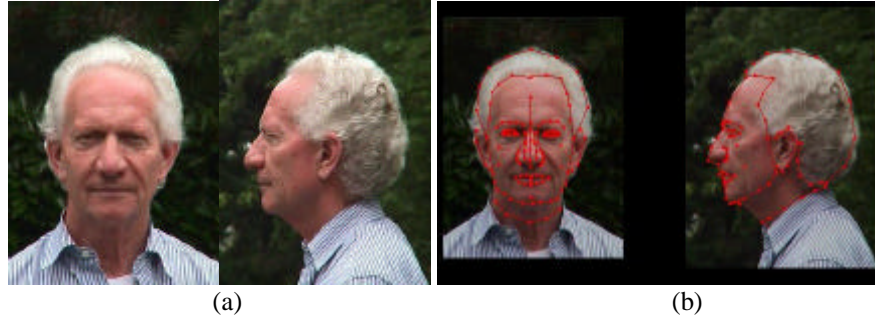

(a)   (b)

Figure 1: (a) The front and side views of a Caucasian man. (b) Scaling and translation of given images after normalization and detected features.
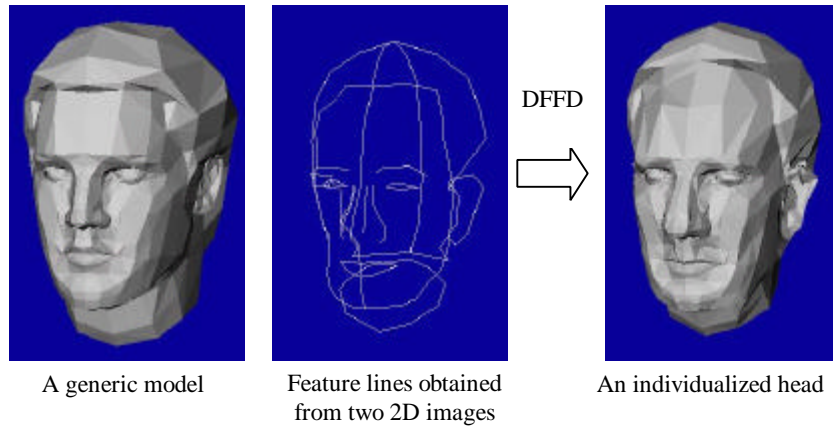

A generic model   Feature lines obtained   An individualized head
                       from two 2D images

Figure 2: Modification of a generic head with detected feature points

3. Texture mapping

Texture mapping is useful not only to cover the rough matched shape, here the shape is obtained only by feature point matching, but also to get a more realistic colorful face.

The information of detected feature points is used for automatic texture generation combining two views. The main idea of texture mapping is to get an image by combining two orthogonal pictures in a proper way to get highest resolution for most detailed parts. We first connect two pictures on predefined feature lines using a geometrical deformation and a multiresolution technique for smoothness without boundary effect, and then give proper texture coordinates of every point on a head following same transformation with image transformation.

3.1. Texture generation

Image deformation

A front view is kept as it is to keep high resolution and side view is deformed to be connected to front view in certain defined feature points lines. We deform the side view face to attach to the front view face in right and left direction. In the front image, we can draw feature lines as we can see two red lines on front face in **Figure 4**. There is a corresponding feature line on a side image. We deform the side image to transform the feature line, the same as the one on the front view. Image pixels in right side of feature lines are transformed with the same transform as the line transform as shown in **Figure 3**. To get the right image, we utilize side image as it is and deform it with the right the red feature line on the front image. For a left image, we flip a side image vertically and deform it with the left-hand red feature line on the front image. The resulted three images are shown in **Figure 4**.
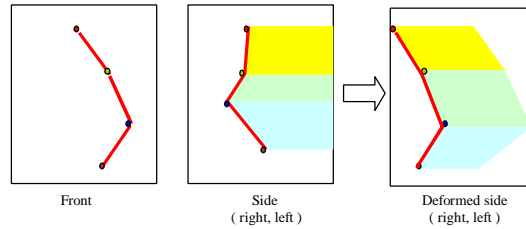

Front   Side (right, left)   Deformed side (right, left)

Figure 3: The side views are deformed to transform certain lines in side view to ones in front view.

Figure 4: Three images after deformation ready for merging.

Multiresolution image mosaic

The two resulted images after deformation are merged using pyramid decomposition of image [4] using the Gaussian operator. We utilize REDUCE and EXPAND operators to obtain G_k (Gaussian image) and L_k (Laplacian image) and merge three L_k images on each level on any given curves, here they are feature lines for combination. Then the merged images P_k is augmented to get S_k, which is the resulted image for each level obtained from P_k and S_{k+1}. The final image is S_0.

Figure 5 shows the whole process of the multiresolution technique to merge three images and
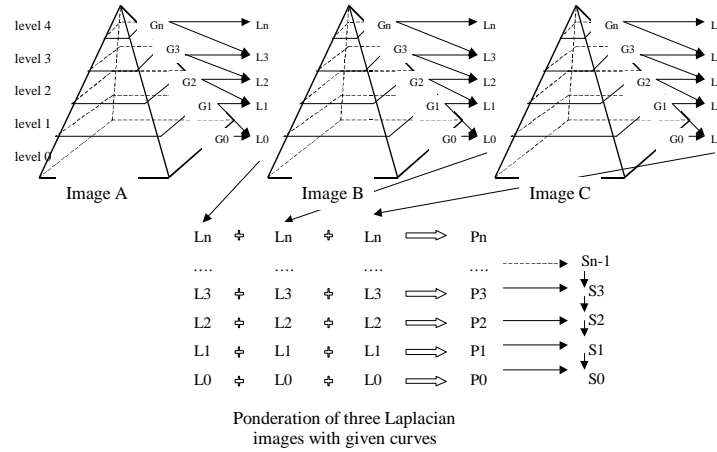**Figure 6** shows an example from level 3 to level 2.

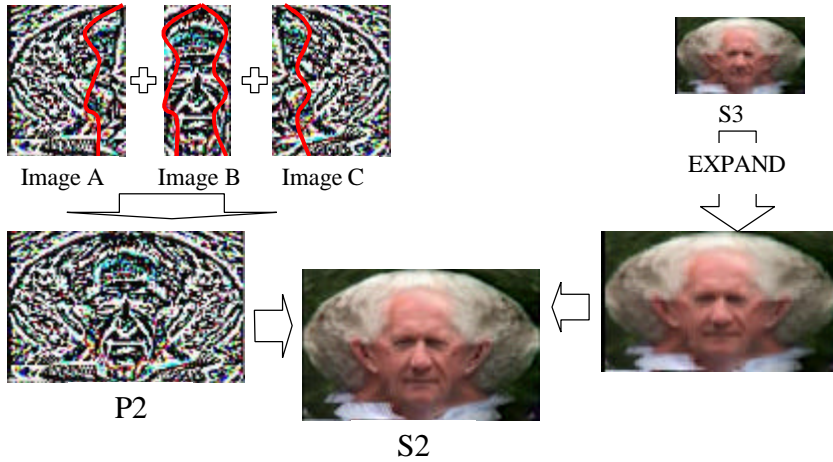**Figure 5**: Pyramid decomposition and merging of three images.

Figure 6: The process from level 3 to level 2.

This multiresolution technique is very useful to remove boundaries between the three images. Although we try as much as possible to take pictures in a perfect environment, boundaries are always visible in real life. As we can see in **Figure 7** (a) and (c), skin colors are quite different when we do not use the multiresolution technique. The images in (b) and (d) show the results after the multiresolution technique, which removes boundaries and makes a smooth connection between images.
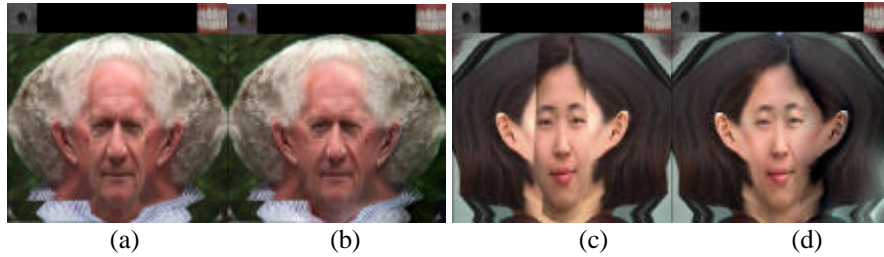
(a) (b) (c) (d)

Figure 7: The generated texture images combined from the three (front, right, and left) images without multiresolution techniques in (a) and (c) and with the technique in (b) and (d).

Eyes and teeth images are added automatically on top of an image, which are necessary for animation of eyes and mouth region.

3.2. Texture fitting

To give a proper coordinate on a combined image for every point on a head, we first project an individualized 3D head onto three planes. With the information of feature lines, which are used for image merging in above section, we decide which plane a point on a 3D head is projected. Then projected points on one of three planes are transferred to one of feature points spaces between the front and the side in 2D. Finally, one more transform on the image space is processed to obtain the texture coordinates. The origins of each space are shown in **Figure 8** (a) and the final mapping of points on a texture image is generated. 3D head space is the space for 3D head model, 2D feature points space is the one for feature points which are used for feature detection, 2D image space is the one for space for orthogonal images which are used for input, and 2D texture image space is for the generated image space. The 2D-feature point space is different from 2D-image space even though they are displayed together in **Figure 1**.
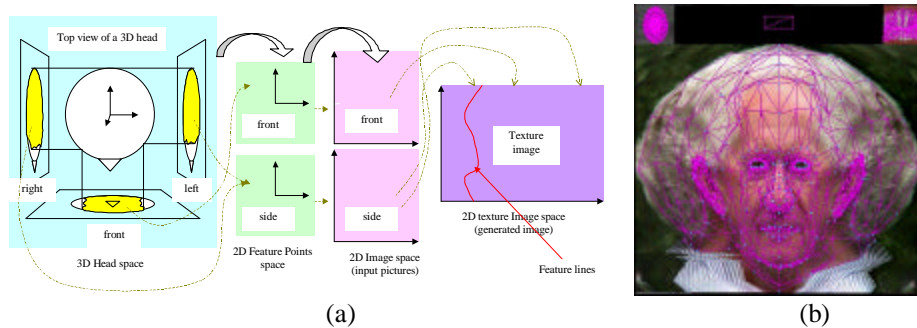

(a) (b)

Figure 8: (a) Texture fitting process to give a texture coordinate on an image for each point on a head. (b) Texture coordinates overlaid on a texture image.

The final texture fitting on a texture image is shown in **Figure 8** (b). The eyes and teeth fitting process are done with predefined coordinates and transformation related to the resulted texture image size, which is fully automatic after one process for a generic model.

The brighter points in **Figure 8** (b) are feature points while the others are non-feature points and the triangles are a projection of triangular faces on a 3D head. Since we utilize a triangular mesh for our generic model, the texture mapping is resulted on efficient triangulation of texture image showing finer triangles over the highly curved and/or highly articulated regions of the face and larger triangles elsewhere as the generic model does. This resulting triangulation is used for 3D-image morphing in Section 4.2.

The final texture mapping results in smoothly connected images inside triangles of texture coordinate points, which are given accurately. **Figure 9** shows several views of the reconstructed head out of two pictures in **Figure 1** (a).

Figure 9: snapshots of a reconstructed head of a Caucasian male in several views

4. 3D morphing between two persons

When we morph one person to another person in 3D, there are two things needed. One is the shape variation and the other is texture variation.

4.1. 3D interpolation in shape based on same topology

Every head generated from a generic model shares the same topology with a generic model and has similar characteristic for texture coordinates. Then resulted 3D several shapes are easily interpolated. An interpolated point P between P_L and P_R is found using a simple linear interpolation. **Figure 10** shows a head after interpolation where $a = b = 0.5$. The left head is slightly rotated and the middle has an interpolated shape and an interpolated position with some rotation.
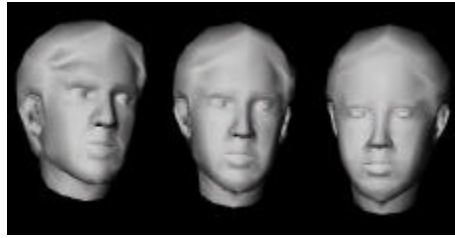
Figure 10: Shape interpolation

4.2. Image morphing

Beside shape interpolation, we need two items to obtain intermediate texture mapping. First texture coordinate interpolation is performed and image morphing follows.

2D interpolation of texture coordinate

It is straightforward as 3D-shape interpolation. An interpolated texture coordinate C between C_L and C_R is found using a simple linear interpolation in 2D.

2D-image metamorphosis based on triangulation

We morph two images with a given ratio using texture coordinates and the triangular face information of the texture mapping; we first interpolate every 3D vertex on the two heads. Then to generate new intermediate texture image, we morph triangle by triangle for every face on a 3D head. Parts of image, which are used for the texture mapping, are triangulated by projection of triangular faces of 3D heads since the generic head is a triangular mesh. See **Figure 8** (b). With this information for triangles, Barycentric coordinate interpolation is employed for image morphing. **Figure 11** shows that each pixel of a triangle of an intermediate image has a color value decided by mixing color values of two corresponding pixels on two images. Three vertexes of each triangle are interpolated and pixel values inside triangles are obtained from interpolation between two pixels in two triangles with the same Barycentric coordinate. To obtain smooth image pixels, bilinear interpolation among four neighboring pixels is processed.
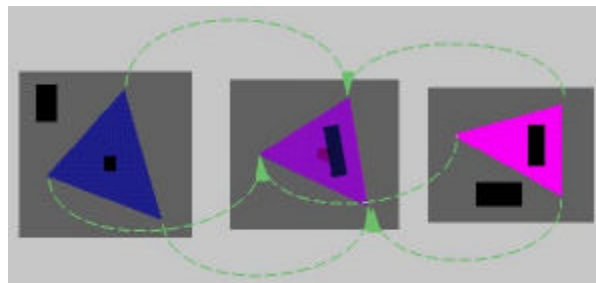

Figure 11: Image transformation using Barycentric coordinates.

We process the triangles on the image one by one. The new problem is how to fill the triangle in an efficient way. For the three vertexes *P1, P2,* and *P3* of a triangle, we rename them according to the *x* coordinate in descending order like $P1_x \leq P2_x \leq P3_x$. Then a line between *P1* and *P2* and one between *P1* and *P3* are compared to fill the triangle to find the range of *y* for a given *x* where $P1_x \leq x \leq P2_x$. After similar checking is performed to find *(x, y)* where $P2_x \leq x \leq P3_x$.

For each pixel *(x, y)* in an intermediate image, the Barycentric coordinate *(u, v, w)* is found inside a triangle with three points *P1, P2,* and *P3*. There are corresponding points in two input images, say *P1L, P2L,* and *P3L* for the first image and *P1R, P2R,* and *P3R* for the second one. Then *u*P1L + v*P2L + w*P3L* is the corresponding pixel in the first image and vice versa for the second. The color value *M(x, y)* of a given pixel *(x, y)* is found from linear interpolation between the color value *ML(x, y)* of the first image and *MR(x, y)* of the second as the formula below.

$$M(x, y) = aM_L(uP1_{Lx} + vP2_{Lx} + wP3_{Lx}, uP1_{Ly} + vP2_{Ly} + wP3_{Ly})$$
$$+$$
$$bM_R(uP1_{Rx} + vP2_{Rx} + wP3_{Rx}, uP1_{Ry} + vP2_{Ry} + wP3_{Ry})$$

To show 2D-image metamorphosis, we show another example of a face as seen in **Figure 12**. It shows very similar structure to the Caucasian male example.
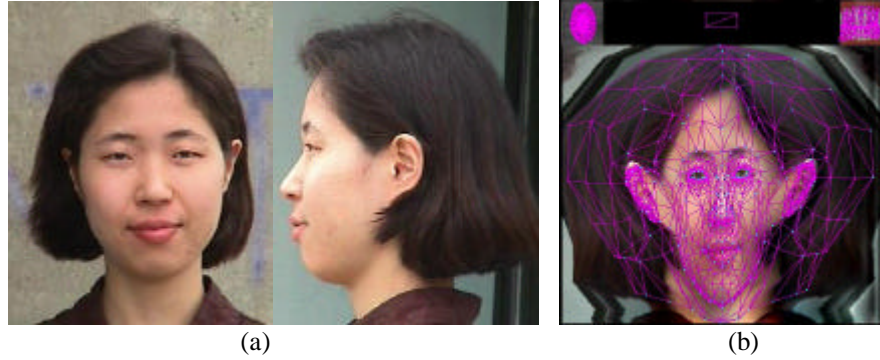

| (a) | (b) |

Figure 12: (a) The input front and side view images of an Asian female. (b) A generated texture image with visual texture coordinate for triangular faces of a corresponding 3D head.

We vary the morphing ratio *a* and *b* to show dynamic morphing. **Figure 13** illustrates intermediate texture images between one in **Figure 8** (b) and the one in **Figure 12** (b). Pixels like eyes, mouth and some part of hair regions which are not covered by triangles are not calculated to get intermediate values, but are not used for texture mapping anyway. If different eye colors are used, it interpolates eye colors too.

Figure 13: Intermediate images. Starting from top left going in a clockwise direction, it is mixed for (Caucasian male, Asian female).

5. Various morphing experiments

Figure 14: Morphing steps in clockwise direction starting from left-up for (Caucasian male, Asian female).

The resulting head shows very smooth images without any hole in the textured parts. **Figure 14** shows the final result of 3D morphing. It shows how to interpolates the shapes and skin and hair colors between old Caucasian male and young Asian female and **Figure 15** shows another variation of an Asian female mixing with a Caucasian female. The middle ones have mixed characteristics.

Figure 15: 50% morphing between Caucasian and Asian females.

Since it is a morphing in 3D, changing the view is an easy task to enable the visualization of various profiles. **Figure 16** shows two examples. Left-hand shows a Caucasian male with the Asian female and right-hand shows the Caucasian male with another male.

Figure 16: Different views of morphing between three people. They show the change of shape and image morphing. Each middle head is mixed with a ratio of 50%.

A different ratio for shape and image variation

It is possible to set different ratios for the interpolation of the 3D shapes and for morphing images. **Figure 17** shows different ratios of morphing for shape and texture images. The Asian female looks older with a lot of wrinkles and white hair, which are borrowed from a Caucasian male on the left side.

Figure 17: 90% of shape is from the Asian female and the image is mixed 50% for each of Caucasian male and Asian female.

Interpolation between several expressions

Since a reconstructed head can be animated in active way, which keeps the same topology and texture mapping, it shows an easy solution for the experimentation of mixing people with different expressions. **Figure 18** and **Figure 19** show other example for reconstruction. Input pictures in **Figure 19** are rotated interactively to get orthogonal pair.

Figure 18: A reconstructed Indian man from orthogonal pictures

Figure 19: Input images in gray for J.F.K and reconstructed J.F.K.

The (generic) face model is an irregular structure defined as a polygonal mesh. The face is decomposed into regions where muscular activity is simulated using rational free form deformation [13]. As model fitting transforms the generic face without changing the underlying structure, the resulting new face can be animated.

Animation can be controlled on several levels. On the lowest level, we use a set of 65 minimal perceptible actions (MPAs) related to muscle movements.

The two heads in **Figure 20** have different expressions and head positions. The middle head with in-between shape and in-between texture shows in-between expression.

Figure 20: Morphing in 3D with various expressions.

6. Conclusion

We introduced methods from for realistic looking face reconstruction from two pictures and to for 3D morphing between given two faces, which is not only 3D-shape interpolation or 2D-image metamorphosis, but both of them.

We have first shown an efficient face modeling for animation with new texture techniques. It shows the processes modifying a generic model for shape acquirement and producing texture images combining orthogonal pair input deforming geometrically and then using multiresolution technique.

The resulting heads have wide range covering man and woman, and also different kinds of race from only one generic model. These models have the same topological structure for the 3D shape and similar characteristic for texture images, which enables 2D-image metamorphosis based on triangulation that results from our reconstruction method. The 2D-image morphing using Barycentric coordinate interpolation is possible with triangulation, which comes from an individualized head reconstruction from a generic head.

Various experiments show a potential area to develop a family and aging simulation from different ratios for 3D-shape interpolation and 2D-image morphing.

7. Acknowledgment

8. References

[1]   Exhibition On the 10th and 11th September 1996 at the Industrial Exhibition of the British Machine Vision Conference.

[2]   http://www.turing.gla.ac.uk/turing/copyrigh.htm

[3]   Takaaki Akimoto, Yasuhito Suenaga, and Richard S. Wallace, Automatic Creation of 3D Facial Models, IEEE Computer Graphics & Applications, Sep., 1993.

[4]   Peter J. Burt and Edward H. Andelson, A Multiresolution Spline With Application to Image Mosaics, ACM Transactions on Graphics, 2(4):217-236, Oct., 1983.

[5]   Horace H.S. Ip, Lijin Yin, Constructing a 3D individual head model from two orthogonal views. The Visual Computer, Springer-Verlag, 12:254-266, 1996.

[6]   Tsuneya Kurihara and Kiyoshi Arai, A Transformation Method for Modeling and Animation of the Human Face from Photographs, Computer Animation, Springer-Verlag Tokyo, pp. 45-58, 1991.

[7]   Yuencheng Lee, Demetri Terzopoulos, and Keith Waters, Realistic Modeling for Facial Animation, In Computer Graphics (Proc. SIGGRAPH), pp. 55-62, 1996.

[8]   L. Moccozet, N. Magnenat-Thalmann, Dirichlet Free-Form Deformations and their Application to Hand Simulation, Proc. Computer Animation, IEEE Computer Society, pp. 93-102, 1997.

[9]   Meet Geri: The New Face of Animation, Computer Graphics World, Volume 21, Number 2, February 1998.

[10]   Marc Proesmans, Luc Van Gool. Reading between the lines - a method for extracting dynamic 3D with texture. In Proceedings of VRST, pp. 95-102, 1997.

[11]   http://www.viewpoint/com/freestuff/cyberscan

[12]   Lee W. S., Kalra P., Magenat Thalmann N, Model Based Face Reconstruction for Animation, In Proc. Multimedia Modeling (MMM) '97, Singapore, pp. 323-338, 1997.

[13]   P. Kalra (1993) An Interactive Multimodal Facial Animation System, PH.D. Thesis, Ecole Polytechnique Federale de Lausanne.

[14]   T. Beier, S. Neely, Feature-based image metamorphosis, In Computer Graphics (Proc. SIGGRAPH) pp. 35-42, 1992.

[15]   J. Kent, W. Carlson , R. Parent, Shape Transformation for Polygon Objects, In Computer Graphics (Proc. SIGGRAPH) pp. 47-54, 1992.

[16]   Kiyoshi Arai, Tsuneya Kurihara, Ken-ichi Anjyo, Bilinear interpolation for facial expression and metamorphosis in real-time animation, Visual Computer, Volume 12, number 3, 1996.

[17]   S.-Y. Lee, K.-Y. Chwa, S.-Y. Shin, G. Wolberg, Image metamorphosis using Snakes and Free-Form deformations, In Computer Graphics (Proc. SIGGRAPH) pp. 439-448, 1995.