Morphing Human Faces: Automatic Control Points Selection And Color Transition

Stephen Karungaru, Minoru Fukumi, and Norio Akamatsu

Abstract— *In this paper, we propose a morphing method by which face color images can be freely transformed. The main focus of this work is the transformation of one face image to another. This method is fully automatic in that it can morph two face images by automatically detecting all the control points necessary to perform the morph. A face detection neural network, edge detection and medium filters are employed to detect the face position and features. Five control points, for both the source and target images, are then extracted based on the facial features. Triangulation method is then used to match and warp the source image to the target image using the control points. Finally color interpolation is done using a color Gaussian model that calculates the color for each particular frame depending on the number of frames used. A real coded Genetic algorithm is used in both the image warping and color blending steps to assist in step size decisions and speed up the morphing. This method results in "very smooth" morphs and is fast to process.*

Keywords— Color transition, Genetic algorithms morphing, warping.

I. INTRODUCTION

Image metamorphosis, or morphing for short, is commonly referred to as the animated transformation of one digital image to the other. Image morphing finds numerous applications in many fields including computer vision, animation, art and medical image processing. Image warping can be defined as a method for deforming a digital image to different shapes. Practically, this can be simulated using an image drawn on an elastic surface. By moving the corners of the elastic material to new positions, the image will deform accordingly. Image morphing combines image warping with a method that controls the color transition in the intermediate images produced. To morph one image to another, new positions and color transition rates for the pixels in each of the images in the sequence must be calculated. Three processes are involved; feature specification, warp generation and transition control [1]. The feature specification process, defines the control points to be used when the image is being warped. This is a difficult process that in most cases is performed manually. Warp generation is an algorithm that calculates and transforms the pixels in one image to new positions in the other image. Many algorithms have already been proposed to do warping

as briefly described in the next section. Once the pixels are in position, transition control blends in the colors between the two images. Transition control has also received a lot of attention. Originally, cross-dissolve was the color blending method of choice, but this method produced undesirable artifacts referred to as "ghosts" due to the computed warp function [1].

In this paper, feature specification is performed automatically using a combination of a face detection neural network, edge detection and smoothing filters. A triangulation method is used as the warp algorithm while a method based on the Gaussian function is applied in color transition control. In the latter two methods, we introduce a genetic algorithm to decide on the warp and blend step size and therefore speed up the morphing process.

If the source and target images are the same but the control points in the target image are freely chosen, the deformation of the source image is possible. Deformation can be achieved using only the warping method described in this paper. In this case, no intermediate images are produced and therefore transition control is not required.

II. RELATED WORK

Over the years, many methods have been proposed to perform image morphing. Of these, the following are worth mentioning. Triangulation based methods achieve interpolation through a triangulation of control points by first dissecting the defining space into a number of triangles with the given control points being the corners of the triangle. Each of the resulting triangles is then interpolated independently [2]. Obviously, given any number of control points, many triangulations are possible. To avoid thin, poorly shaped triangles and therefore find the optimal triangulation, Delaunay triangulation [3] is commonly used.

Field morphing handles correspondence by means of line pairs. A line of corresponding lines in the source and target images defines a coordinate mapping between two images, [4]. The Shepard approach of scattered data uses a weighted average of the data values at the data points, with weights dependent on the distance of the points from the given control point, [5]. Other methods include the radial basis functions, nearest neighbor interpolation, inverse distance weighted method etc. A recent comprehensive study of these methods is available in [6].

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:1, No:1, 2007

III. Morphing

The morphing problem can be defined as follows: Given two face images, transform one into the other as smoothly (visually pleasing) and as fast as possible. The problem involves image warping and color interpolation. The procedures described here, which collectively produce the morph, are control point extraction, warping and transition control.

A Control Points Extraction

The choice and number of control points determine how accurately two images can be warped. The use of many control points will usually result in a very good warp. However a high number of control points will reduce the speed of the system since the number of triangles increase, which in turn increases the computation time. Given an image, this work focuses on the warping of the human face. Therefore, the control points selected are all within the face region. Five control points, center of both eyes (C1, C2), tip of the nose (C3) and both corners of the mouth (C4, C5), are consequently defined in this system as shown in Fig. 1(left).
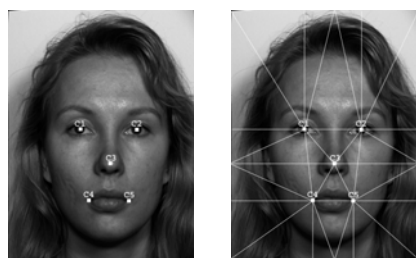
Fig. 1. Left: Location of the control points. Right: The 32 triangles used for warping. There are six triangles between the control points chosen.

The choice of these facial points as control points is base on the fact that some of their properties, for example, the distances between them, have strong features that can also be used in face recognition. In addition, these points can also be accurately detected and extracted in most face images. Moreover, given two faces of different sizes these control points can be used to normalize the faces to a given normal size. For each of the two images to be morphed, the control points are automatically detected and extracted using the following processes.

B Face Detection Neural Network

First, for each of the two images, the position of the face region in the images needs to be determined. This is accomplished using a neural network based face detector. The face detector, referred to as face detection neural network (FDNN) extracts the position of the face from an image [7]. The FDNN consists of a face locator, down sampler and a merger.

The face locator consists of a skin color detector and a neural network. At first, the skin color regions of the given image are detected using the skin color detector. The skin color detector was implemented using a YIQ color system threshold based method. The neural network, face detector, is the part of this system that does the actual face detection. The face detector is chosen to be a three layered back propagation trained neural network. The size of the training samples was set at 20x20 pixels because experiment showed that both the accuracy of the system and the speed were best using this size.

The error back propagation method is used to train the neural network. The system is trained to produce an output of 0.95 for a face and 0.05 for a non-face. Structural learning with knowledge is also used during training to reduce the size of the face locator and therefore improve the overall speed. The FDNN used independently has an accuracy of 97% when tested with images with complex backgrounds and including many people per image. In this paper, each image used contains only one face taken on a neutral background. In such a situation the FDNN accuracy is 100%.

C Filters

After the face position has been determined, the position of the facial features, and hence the control points needs to be extracted. For all images with the face facing directly forward, the control point C3, tip of nose, Fig.1(left), is directly determined using skin color detection step of the FDNN. Moreover the average position of the other features can be estimated using the same result.

The filters are used to conform and reinforce these positions. The filters are applied only inside the face regions detected by the FDNN. The first filter applied is the Laplacian of Gaussian, to smooth and detect the edges of the facial features. A large kernel (e.g. 15x15) used with this filter produces very good edge detection but the computation time rises exponential with the size of the kernel. After several experiments, a trade-off kernel of 7x7 was employed. Fig. 2 shows the results of this filter using a 7x7 kernel. All the edges are extracted as shown.

Fig. 2. Left: Edge detection results for the 7x7 laplacian of gaussian filter. Right: Edge detection results for the 7x7 laplacian of gaussian filter and the 3x3 medium filter

However, a lot of noise is produced. Therefore a 3x3 medium filter is then used to reduce this noise. Fortunately the cost of calculation when using a 7x7 laplacian of gaussian filter followed by a 3x3 medium filter was less than using a 15x15 Laplacian of Gaussian that produced similar results.

After noise reduction using the medium filter, the edges are grouped together (small regions and areas around the nose, the left and the right bottom corners are ignored). The regions at the top and the bottom are selected from which the control points C1 (center of the area on top-left), C2 (center of the area on top-right), C4 (bottom area, left-edge's center) and C5 (bottom area, right-edge's center) are extracted, Fig. 1(right).

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:1, No:1, 2007

The choice of the control points avoids very thin and poorly shape triangles. Such triangles produce an unnatural looking warp, due to pixel compression, [8].

D Genetic Algorithms

Accurate extraction of the lips and the eyes is vital for face recognition. We employ a genetic algorithm to reinforce the results of the segmentation of these regions by using the GA to assist in the matching of the general shapes and the segmented areas. The length of the GA chromosome used is 26. The first 5 pixels represent the height and the next 5, the width of the shapes. The remaining 16 chromosomes represent selected points in the shape (roughly equal distance from each other).

The size of the shape can thus be reduced or increased to fit the size of the segmented region. The genetic algorithm processes of reproduction crossover and mutation are described below.

During reproduction, we use the top 50% of the fittest individuals to reproduce 75% of the next population. The remaining 25% of the next population is reproduced by selection of their parents at random from the whole initial population. This method improves the search space by ensuring that we not only retain the best individuals for reproduction but also explore the rest of the population for other possible candidates. The selection of parents is performed using the roulette wheel method.

The crossover point is determined depending on the total fitness of the parents. The higher the fitness the lower, in the chromosome string, the point of crossover is.

The fitness function is designed such that the search converges as fast as possible. Fast convergence is possible if the size of the shape matches the size of the segmented region fast. The fitness function is given as

$$Fitness = \frac{(x - x_i)^2 + (y - y_i)^2 + \sum (z - z_i)^2}{(x + y + z)^2}$$

where: x, (x_i), y, (y_i) and z, (z_i) are the height, width and the points selected from the shape (segmented area) respectively.

E Warping

Control points extraction is performed for both the source and target images. Once these points are found, the images are ready for warping.

The source and target images are of the same size, but the faces in them can be of different sizes. The warping method used in this work is triangles based interpolation. The images are divided into several triangles using the control points already found. There are 32 triangles per image as shown in Fig. 1 (right).

Note that apart from the four corners of the image, all the other points, on the edges of the image, that complete the triangles are all related to the five control points and therefore easy to evaluate. To evaluate the values of these points all that needs to be known is the size of the image. For two corresponding triangles, one from the source image and the

other from the target image, the warping transformation from one to the other can be performed.

Let points P_1, P_2, P_3 on the source image be located at $\mathbf{x}_1 = (u_1, v_1)$, $\mathbf{x}_2 = (u_2, v_2)$ and $\mathbf{x}_3 = (u_3, v_3)$. Also let points Q_1, Q_2, Q_3 on the target image be located at $\mathbf{y}_1 = (x_1, y_1)$, $\mathbf{y}_2 = (x_2, y_2)$ and $\mathbf{y}_3 = (x_3, y_3)$. The points on the source image can be mapped to those on the target image using eqs. 1 and 2.

$$x = a_{11}u + a_{21}v + a_{31} \tag{1}$$
$$y = a_{12}u + a_{22}v + a_{32} \tag{2}$$

The coefficients a_{11}, a_{12}, a_{21}, a_{22}, a_{31} and a_{32} can be found by solving the equation below.

$$\begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 1 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix} \begin{bmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \\ u_3 & v_3 & 1 \end{bmatrix}^{-1} \tag{3}$$

Warping can then be done using target to source mapping. In target to source no pixels on the target image are missed and therefore no color interpolation is necessary.

F Color Transition

A morph contains a sequence of intermediate images from the source image to the target image. Color transition is the method that determines the rate of color blending across the sequence. The choice of this rate determines the quality of the morphs.

Interesting morphs can be created depending on whether the color-blending rate changes locally or globally. The rate of color blending is usually based on weights. Such weights are selected to smoothly complete the transition between the images on the sequence. In this paper, the weights are calculated using a one-dimension Gaussian function. This method is implemented as follows:

Given two corresponding target and source pixels, first, calculate the difference in color between them and then set the Gaussian function 1 for the target pixel and to 0 for the source image. The weight for each morph in the sequence is the calculated based on the color difference calculated before, the value of the Gaussian function at that point and the number of warps in the sequence. The color for each image in the morph sequence is then calculate using;

$$\psi_j = \psi_i - \omega_i * \Delta\psi_{ij} \tag{4}$$

Where: ψ_j is the color for the new warped pixel, ω_i is the weight and $\Delta\psi_{ij}$ is the color difference between the target and source pixels.

This procedure is repeated for every pixel in the image and for every image in the morph sequence. The source image changes from the original image, to the new image produced by the first warp and so on.

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:1, No:1, 2007

IV. RESULTS & DISCUSSION

Fig. 5 shows a selected morphing image sequence using the proposed method. The images used in this work were acquired from the physics based face database of the University of Oulu, [9]. The size of all the images is 540x400 pixels.

To morph between a source (a) and target (o) images, 13 intermediate images (b)-(n) were produced. Notice the smoothness of the face regions. Also note that halfway through the morph, images (h)-(j), faces of completely different persons are created.
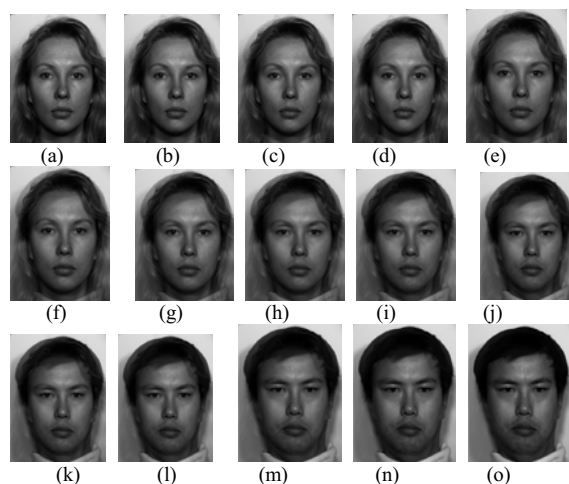
Fig. 5. Morphing: (a) source, (o) target (b)-(n) morph frames.

The warping method discussed in section IIIE can be used independently to freely deform images using all, one or any combination of the control points. During warping the source and target images are the same. Also, no intermediate images are produced, as is the case with morphing. Fig. 6 shows some of the images produced by the warping function.

Fig. 6. Image warping: (a) source (b)-(e) warped images.

This work was carried out using a Dell Optiplex Sx260 512MB personal computer. took about 2.01 seconds, while warping took 0.11 seconds to complete.

Five control points were automatically extracted to control the morph. The five control points were all selected from inside the face. This offers the advantage of morphing the face regions very smoothly. However as the distance from the control points increase the creation of shadows is apparent, especially around the hair regions. While skin color regions can be detected fairly accurately for different people, hair regions cannot due to different hair colors and complex backgrounds.

Using the Gaussian function as the color transition control increases the amount of computation time required, but has the advantage of producing better morphs. One other method tried was to use a constant color transition rate (equal percentage increments) on all the images based on the number of images in the morph. While this run a little faster than the Gaussian function method, its results were not visually better.

V. CONCLUSION

This paper describes a method using which face images can be automatically morphed. As described the only information required from the user is the size of the source and target images. The results show that the method produces smooth morphs using only 15 frames. Also the ability of the warping method to deform images is demonstrated.

This work uses only 5 control points to achieve a good result for the face region. These control points are extracted automatically, which reduces the parameters that must be inputted by the user. The time taken to produce a morph is 2.01 seconds. This includes the time taken to load the image to and from memory. Warping process took only 0.11 seconds.

One of the areas in this work that need improvement is the selection and extraction of the control points. More control points need to be defined to take care of the subjects hair and the background.

It is also hoped that the color transition algorithm will be improved so that the number of frames required to complete the morph are reduced. This work will also be extended to morphing between more than two images.

REFERENCES

[1.] G. Worlberg, "Image morphing: a survey," *The Visual Computer*, Vol. 14, pp. 360-372, 1998.
[2.] A. Goshtasby, "Piecewise linear mapping functions for image registration" *Pattern Recognition,* Vol. 19(6): pp. 459-466,1986.
[3.] M. Berg, M. Kreveld, M. Overmars and O. Schwarzkorf, "Computation geometry- Algorithms and Applications", *Springer*, 1997.
[4.] T. Beier and S. Nelly, "Feature-based image metamorphosis" *Proc. of SIGGRAPH,* pp. 35-42, 1992.
[5.] D. Rurecht and H, Muller, "Free form deformation with scattered data interpolation methods ", *Geometrical Modeling,* pp. 267-281, Springer Verlag, 1993.
[6.] I. Amidror, "Scattered data interpolation methods for electronics imaging systems: a survey", *Journal of Electronic Imaging,* 11(2), pp. 157-176, 2002.
[7.] S. Karungaru, M. Fukumi and N. Akamatsu, "Detection of human face in visual scenes." *Proc of ANZIIS,* pp.165-170, 2001.
[8.] G. Nielson, H. Hager and H. Muller, "Scientific Visualization", pp.433, IEEE, New York, 1997.
[9.] Soriano M, Marszalec E, Pietikäinen M, "Color correction of face images under different illuminants by RGB eigenfaces", *Proc. 2nd Audio- and Video-Based Biometric Person Authentication Conference,* pp. 148-153, 1999.