NPAR by Example: line drawing facial animation from photographs

Yuan Luo, Marina L. Gavrilova and Mario Costa Sousa
Department of Computer Science
University of Calgary
Calgary, Alberta, Canada T2N 1N4
yluo,marina, mario@cpsc.ucalgary.ca

Abstract

This paper presents a new method for generating non-photorealistic (NPR) facial expression animations. The method consists of two main processing parts. First, a particular NPR style portrait is created from a frontal facial photograph. Facial expressive details, like: expressive wrinkles,fine creases, even freckles, can be included in the synthesized portrait, which makes it more expressive. Second, with two different facial expression portraits, a metamorphing algorithm using distance transform is utilized to produce the 2D animation. Our morphing process is fully automated, no control point or any manual work is required. Finally, a continuous facial expression animation is produced. The resulted animation is useful in a number of potential applications, ranging from entertainment and education to teleconferencing and psychology research.

1. Introduction

Facial animation, as an identifiable area of computer graphics, has long fascinated computer graphics researchers. Historically, the earliest attempts to model and animate realistic human faces date back to the early 1970s[1]. Since then, numerous of research papers have been published on this topic. Yet there is still room for improvement.

NPR animation of faces offer several advantages over attempts to work with photorealistic images. First, viewers do not expect a faithful replica of the speaker in real applications, like: teleconferencing or computer games. Also, line drawing, caricature and many NPR style pictures require less storage space than images, which makes them more suitable for low bandwidth network applications. They are also easy to compress for privacy. Finally, an animated figure has an engaging quality that is often more entertaining than a live video clip.

In this paper, we describe a new method for automatic animation of NPR faces from sample images. Our system takes two human face images as input and outputs a particular NPR style facial animation. First, a particular NPR style portrait is created from a frontal facial photograph. Many NPR techniques have already been proposed to generate digital artwork. However, detail information, such as expressive wrinkles and creases,is usually missed in the generated pictures. As an extension, we propose a segmentation and tracking method to map those expressive lines, which makes the portrait more expressive. Then, a metamorphing algorithm using distance transform is utilized to produce the resulted animation. Our two main contributions include: (1) As a difficult and unsolved problem in NPR, detail information is usually missed in generated pictures. We solve this problem by extracting expressive lines from the original input image, which can be mapped to the generated NPR sketch with an application-dependent style. (2)Our metamorphing algorithm is fully automated to create the 2D animation. Unlike using control points or control lines, distance transform is utilized here to find pixel correspondence automatically.

1.1 Related Work

Two research fields relevant to this paper are NPR picture generation and image morphing techniques.

Many systems have been created to emulate watercolor and impressionism. More relevant to our work, however, are the NPR results of pen-and-ink and technical illustration. NPR techniques have also been used to depict facial images with an artistic style. Li[2] utilized morphological processing techniques to extract facial sketch from a facial image.This approach only connects the feature points of facial parts, which produced a stiff resulted sketch. Combining a flexible sketch model with the non-parametric sampling method, Hong [3] [4] presented an approach to automatically generate sketches from input images.Different from other similar algorithms, Hong synthesizes hair style into face models in this paper. Adapting a modified model

Proceedings of the International Conference on Computer Graphics, Imaging and Visualisation (CGIV'06)
0-7695-2606-3/06 $20.00 © 2006 IEEE

of human brightness perception to photographs, Gooch etc. tried to keep certain facial features in order to preserve recognizability, then an image warping technique was applied to produce caricatures[5]. However, all these methods only consider the basic components of human faces. As the limitation claimed by those authors, fine details, such as: aging wrinkles, richer feature lines of expressions, injury and freckle can't be preserved by these approaches. That is the motivation of our efforts on developing the Expressive Portrait Generator, which offers the opportunities to include those detail information in the generated picture.

Once the still picture is achieved, we usually consider to create an animation which is more interesting. As one main means to create computer animation, morphing/metamorphing or the gradual and continuous transformation of one shape into another is a topic of great importance in computer graphics. Image metamorphosis between two images begins with an animator establishing their correspondence with pairs of feature primitives. The feature correspondence is then used to compute mapping functions that define the spatial relationship between all points in both source and target images. Typically, to get good morphing sequences a number of equal control points are manually selected in both images and then a one-to-one mapping is defined[6][7][8]. Because this approach requires extensive manual interaction for both to select the control points and compute the mapping, it's too computational expensive for our system. Distance transform, also referred as feature transform, is a map that assigns to each pixel the feature that is nearest to it. In order to find the best pixel correspondence, we utilize distance transform as the mapping function for morphing two input NPR portraits. Given two images with the same size, our morphing algorithm guarantees that each pixel will travel the shortest distance to a corresponding pixel,which exhibits as a smooth and visually continuous facial animation.

2 System Overview

The goal of our work is to produce a 2D NPR animation of different facial expressions. The block diagram shown in Fig.1 depicts the two main parts of this system: Portrait Generator and Animation Creator. The purpose of Portrait Generator is to create line drawing style portrait from the input facial photograph. Different with other previous portrait generation methods, detail information, like: fine creases around the eyes, richer feature lines of expressions etc.. is extracted and can be mapped to the generated portrait if needed. Depending on the image scale and users' interests, the system can control the Level-of-Detail(LOD)of the generated portrait. For example, if users want to have a closer look at the area around the left eye, by zooming out that area, more fine lines will show up around the eye.

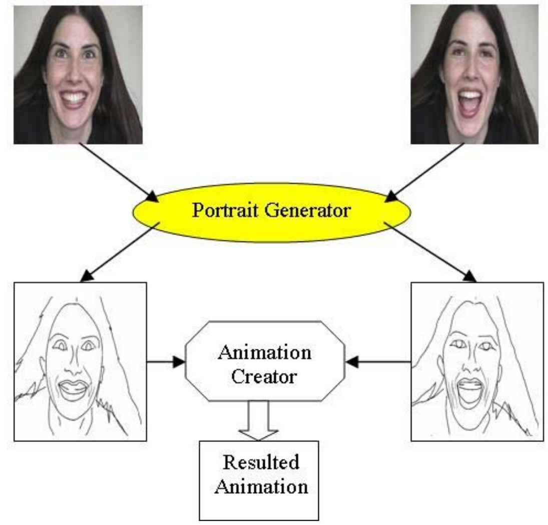


Figure 1. The framework of the whole system

In Animation Creator part, distance transform is utilized to create the morphing sequences between the two generated portraits. As the final result, a 2D line drawing style facial expression animation is produced. Detail explanation for each part will be given in following sections.

3 Expressive Portrait Generator

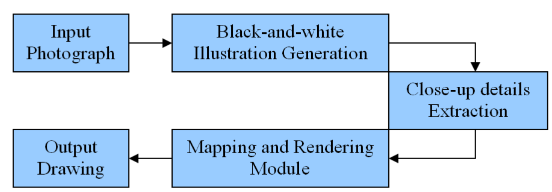


Figure 2. The flowchart of Portrait Generator

As an extension of the previous researches on portrait generation, we focuses on extracting those richer feature lines of human faces by image-space based techniques. In our method, we firstly get the black-and-white illustration from the input photograph, and then feature lines were extracted from 5 fixed windows. We enlarge these five areas in order to use edge detection techniques to segment those close-up detail lines. After thinning the segmented feature lines, a curvature tracking algorithm was applied to follow each line. On the mean time, information of those feature lines, such as: the x and y coordinates of the points along a feature line, the number of points in a feature line and the number of feature lines in a window etc.. is recorded. At the end, depending on users' interests, we could map fine detail features back to the generated portrait. For example, more detail feature lines will show up when zooming in some particular interested area; vice versa, an abstract portrait is given when zooming out. Since fine details like: wrinkles, creases and freckles etc. are given special treatment in our system, no matter what facial expression the subject has or how old it is, an explicit and detail portrait can still be achieved if needed. Figure2 shows various steps in the extraction and mapping phases. Detailed algorithms of each step are discussed in the following subsections.

3.1 Facial Two-tone Illustration

Creating a black-and-white illustration from a photograph can be done in many ways. A number of proposed methods are stroke-based and rely heavily on user input[9][10].In addition,stoke-based methods are mainly concerned with determining stroke placement in order to maintain tonal values across an object's surface. For our application, we prefer a method that could fully be automated and does away with tonal information. we choose our illustration algorithm based on binarization of luminance images, which is the simplest and most effective way to get a two-tone image. Figure3 shows the three steps: first,original photograph was taken by a high-resolution digital camera during regular luminance conditions; second, hue and saturation information was removed while only retaining the luminance;third, based on the histogram of this gray-level image, Otsu threshold was chosen to binarize it into a black-and-white illustration[11].

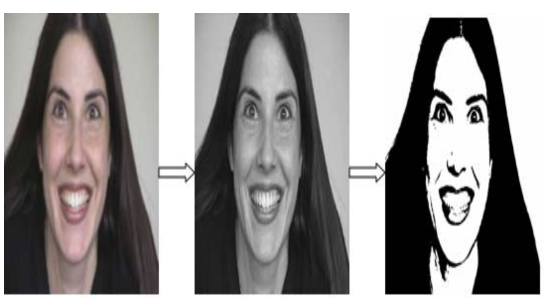


Figure 3. original photograph(left),gray-scale image(middle),two-tone illustration(right).

From the two-tone illustration, we can see that silhouettes of basic components of the face,like: eye brows,eyes,mouth,nose etc.are clearly preserved, which quite satisfied our requirements for this step.

3.2 Close-up Details Extraction

Based on the observation of many human faces, aging wrinkles or richer feature lines during different expressions are usually appearing in some certain regions, so we divide the face into five subregions: left eye area(LE),right eye area(RE),left cheek area(LC),right cheek area(RC) and forehead area(FH),like Fig.3.

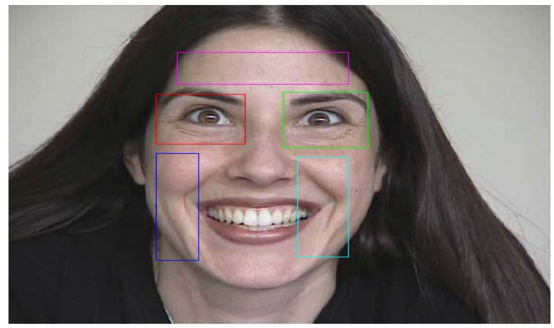


Figure 4. Five interested subregions.

Within each subregion, edge operators were utilized for detecting detail feature lines. Because it's impossible to draw a window which only contains detail feature lines, some normal contours of eyes,mouth etc. will also be included in the result of edge detection. Thus, those contours will be deleted manually to help us to track only those richer lines. Then morphological thinning operation is applied to get a thinned line image, in which detail feature lines will have maximum one-pixel thickness. Finally, we track those feature lines in the thinned line image, and record all the information we need for mapping.This extraction and tracking process is shown Figure5 and will be detailed in the following subsections.

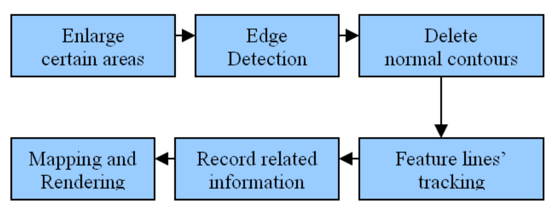


Figure 5. Several steps of extraction and tracking process

3.2.1 Edge Detection For Feature Lines

There are already many existed edge detectors, most popular of which are Sobel, Robert,Canny etc. We can classify these edge detectors into three categories: first-order differential operators, zero-crossing detection of second-order derivatives and more complex heuristic algorithm like: Canny operator. In this project, we compared the edge detection results of five typical operators:Sobel,Prewitt,Roberts,LoG and Canny. The results of extracting edges in LE are shown in Fig.6 as an example.

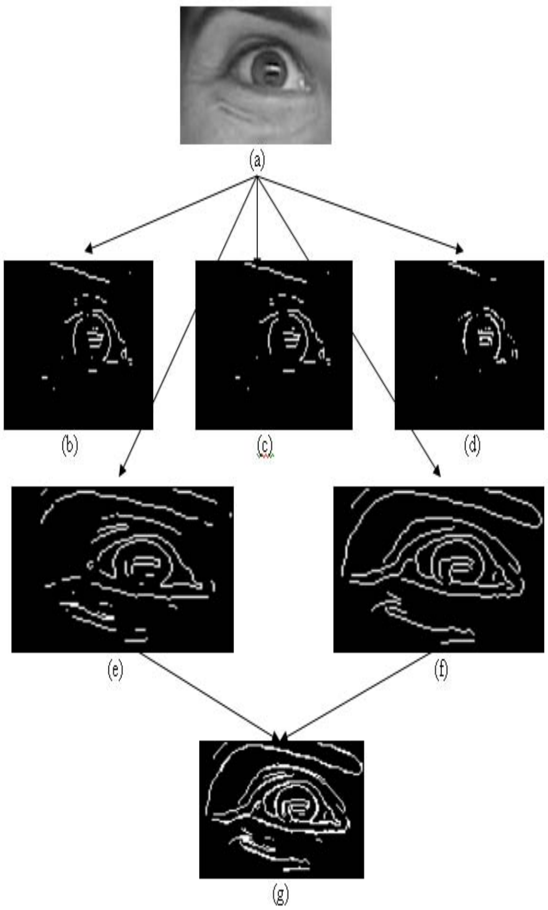


Figure 6. The edges of LE extracted by different edge operators:(a)the original LE image,(b)the result of Sobel operator,(c)the result of Prewitt operator, (d)the result of Roberts operator,(e)the result of LoG operator, (f)the result of Canny operator,(g)combined result of LoG and Canny, produced by addition of (e)and (f)

Since the feature lines are not uniformly in certain directions, all the results extracted by first-order derivatives operators:Sobel,Prewitt,Roberts,are not good enough. Many detail feature lines exactly what we need are lost. Compared with first derivatives, the results of LoG and Canny are much better: LoG covered all the possible features,including those minor features like those curvatures with really a short length and dots etc.;Canny connected some adjacent feature lines, so the resulted edges are smoother. In order to keep the continuity of feature lines and also cover all the possible detail features, we combined the results of LoG and Canny as the detected edges,shown in Fig.6(g).

3.2.2 Feature Lines' Tracking

After we get the binary edge results,the morphological thinning algorithm is applied to get a unitary image, in which the maximal width of the edge is one pixel. By thinning edges into skeletal pixels, which is an accurate representation of feature edges, we can also easily record the position of those feature lines. As illustrated in Fig.7, the thinned feature edges Fig.7(b)accurately represent detail feature lines Fig.7(a).

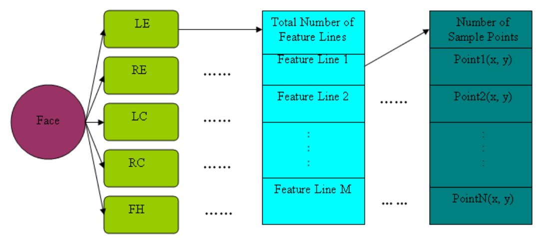


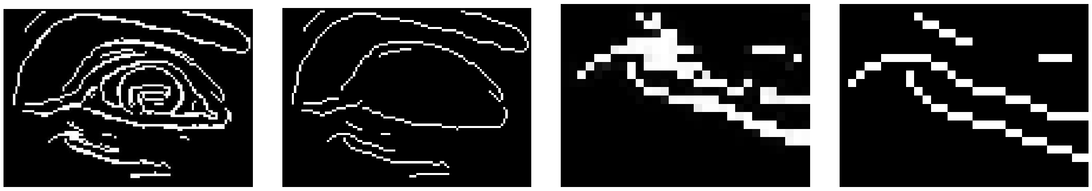
Figure 8. Datastructure "FLRec" used to save feature lines in a face image.

In the thinned image $I(x, y)$, we compute the sum of 8-neighbors of each pixel, and save those pixels whose sum equals to 7 as track starting points S_i. After creating a flag matrix $Flag(x, y)$ used to mark points and feature lines' recorder FLRec(as shown in Fig.8), we start tracking and save the x and y coordinates of sample points along each feature line. The tracking process is summarized in the following algorithms:

TRACKING$(I, S_i, Flag, FLRec)$
$InitializeFlag(x, y) = 0$
for $Point\ P \leftarrow S_i$
  do if $I(P) = 0$ and $Flag(P) = 0$
    then $TrackingLine(I, Flag, P, Sign, FLRec)$
      $Sign = Sign + 1$
      return Rec

TRACKINGLINE$(I, Flag, P, Sign, Rec)$
while $P\ \ is\ \ inside\ \ of\ \ I$
  do $record\ \ P\ \ in\ \ Rec$

Proceedings of the International Conference on Computer Graphics, Imaging and Visualisation (CGIV'06)
0-7695-2606-3/06 $20.00 © 2006 IEEE



(a) (b) (c) (d)

Figure 7. (a)The binary edge detection result of LE;(b)The thinning result of (a);(c)An enlarged area in (a);(d)Corresponding area to (c) in thinned image.

```
    Flag(P) ← Sign
    switch
      case P_fi = 0 :
          P ← P_fi
      case P_ei = 0 :
          P ← Pei
      case default :
          break
```

When tracking process is done with all starting points, feature lines information, such as: the total number of feature lines existed in the face image, the number of points in each feature line, and x,y coordinates of every point,are all recorded in FLRec, which can be used for mapping and rendering in the next step.

3.2.3 Feature Lines Mapping and Rendering

When artists create a portrait or any other line drawing sketches, there are two main particular aspects to take care of. First, there is the placement of lines, that is, their positions determined by the scene to be portrayed. Second, there is the actual look of the lines or how they are drawn with respect to their width, brightness, color, and so on. When mapping back those feature lines, these two problems have to be solved in order to make a uniform style portrait.

The first problem is perfectly solved by exactly mapping those detail lines back to the positions where they are in the input photograph.For each part: LE, RE,LC,RC and FH, we note the corresponding coordinates x_{cor} and y_{cor} with the original whole face image. Then we get each feature points' positions x^i and y^i from FLRec, so the original coordinates of these points are $x^i_{orig} = x_{cor} + x^i, y^i_{orig} = y_{cor} + y^i$. If it's not for a exaggerated caricature, that would be the most proper position for the placement of those feature lines.

For the second problem, those detail feature lines should be rendered in an application-dependent style. Using the line drawing sketch (Fig.9(b)) as an example, feature lines' thickness and intensity should be changed in order to get a uniform look with the original drawing. Like the result shown in Fig.9(c), curvature fitting is also applied to make the feature lines more smooth. The level of details is controlled by the user. Because the length of all feature lines is stored in FLRec, we are using the line's length as a threshold to decide whether it should be mapped back or not. For example, if the user only wants more details on the areas of LC and RC, it's better only to map those longer feature lines, as shown in Fig.9(c).

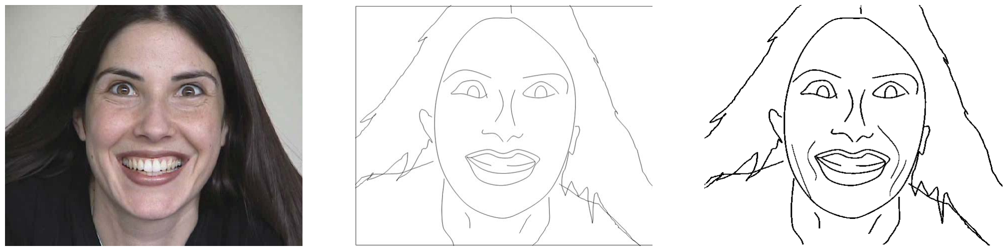
Following this way, our detail feature lines can be mapped to any NPR style portraits with different styles, which also offers chances for the users to have a closer look on particular areas. More results will be shown in Section4.

4 Morphing Algorithm Using Distance Transform

With two input photographs of different facial expressions, two NPR style portraits will be produced by Portrait Generator in the system. As the second step, we would like to create a smooth animation between these two NPR portraits. Given two images A and B, it is sometimes desirable in computer animation to create a smooth transition from one image to the other. This smooth transition is computed as a sequence of frames $S_0, S_1, \ldots S_k$, where $S_0 = A$ and $S_k = B$. Traditionally, to get good morphing sequences a number of equal control points are selected in both images, often by a person, and then a one-to-one mapping is defined. It can be difficult both to select the control points and compute the mapping.We present a novel approach to computing a morphing sequence, one that relies on distance transforms rather than control points.

Before we introduce our morphing algorithm, some definitions are made to facilitate the explanation. Let A, B be two bi-level $m \times n$ images, we treat the image matrix A as a set, that is $A = \{(i,j) \mid a_{ij} = 1\}$. B is defined in the same manner. In addition, the intermediate slides can be viewed either as bi-level images or as grey-scale images.

During the morphing sequence, we would like each pixel $a \in A$ to move to its new location in B, but travel the short-



(a) (b) (c)

Figure 9. (a)The Input Photograph; (b)The Original Generated Portrait from (a);(c)The Portrait with feature lines mapped in LC and RC.

est distance. We can achieve this by sending each pixel to its nearest neighbor in B. In this way, the morphing problem between images A and B is reduced to animate the motion of $N = |A| + |B|$ pixels along straight line segments. To do this, we define two functions, $f : A \mapsto B$, and $g : B \mapsto A$. Each function maps pixels to a nearest neighbor in the other image. Each $a \in A$ is really an ordered pair and represents the location of a foreground pixel in A: $a = (x_a, y_a)$ and similarly for $b = (x_b, y_b)$. We define distance between points a and b in the Euclidean sense $d(a, b) = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2}$. Now we define f and g: $f(a) = \arg\min_{b \in B}\{d(a, b) \mid b \in B\}$, $g(b) = \arg\min_{a \in A}\{d(a, b) \mid a \in A\}$. Function f is called a *feature transform* of B and similarly g is a feature transform of A. $f'(a) = \min_{b \in B}\{d(a, b) \mid b \in B\}$ is a *distance transform*.

In the morphing sequence, each a will travel to its nearest neighbor in B, i.e. $a \mapsto f(a)$. This may cause some visual problems, as this mapping may not be one to one. Although all pixels in A will move, not all pixels in B may be involved in the transformation. Formally, $\mathrm{domain}(f) = A$, but $\mathrm{range}(g)$ need not equal B. Thus, a set of straight line paths in the reverse direction are created from B to A, using g as a guide.

Formally, define a set of line segments L as follows

$$\begin{aligned} L &= \{(a, f(a)) \mid a \in A\} \cup \{(g(b), b) \mid b \in B\} \\ &= \{(u, v)\}, \text{ for notational simplicity} \end{aligned}$$

So L is a set of line segments, each line represented by its endpoints u and v (u and v are each ordered pairs).

Now we are ready to compute the frames S_0, S_1, ... S_k. Define $S_i = \{u + \frac{i}{k}(u - v) \mid (u, v) \in L\}$,where $L = \{(u, v)\} = \{(a_1, b_1), (a_2, b_2), (a_2, b_3)\}$.

The maps f and g can be computed in linear time $O(mn)$, where A is an $m \times n$ image, using the algorithms in [12] and [13]. The algorithms in [12] present techniques using the L_1 and L_∞ norms, and [13] presents the technique using the L_2 norm.

Thus, the distance transform provides out a fast, smooth and visually continuous morphing process of facial expressions. A morphing sequence between a happiness and surprise facial expressions using this method is shown Fig.10.

5 Experimental Results and Analysis

As shown in Fig.11, for an input image(Fig.11(a)), a line drawing sketch is generated in Fig.11(b) and those expressive lines in the left cheek and right cheek, whose length is greater than 60 pixels are mapped back. Also, when users enlarge the areas around the eyes, more fine lines could be mapped back to show those details.From the results in Fig.9 and Fig.11, we can see that those detail feature lines definitely improve the expressiveness of the generated portrait. If only an abstract representation is required, users could also remove those detail features. By giving special treatment to those detail features, our method solves the limitation of previous picture generation systems.

For 2D morphing results, as we described in Section 4, in our experiment the k=60. That is to say, from one expression to another, there are 58 between frames, which make the morphing animation visually continuous. 5 key frames were selected to be shown in Fig.10. Also, like the results in Fig.12, our morphing algorithm can be applied to other NPR style portraits. With the Portrait Generator and Animation Creator together in our system, we offer users with great opportunities to control the level of details in the resulted animation.

6 Conclusion and Future Work

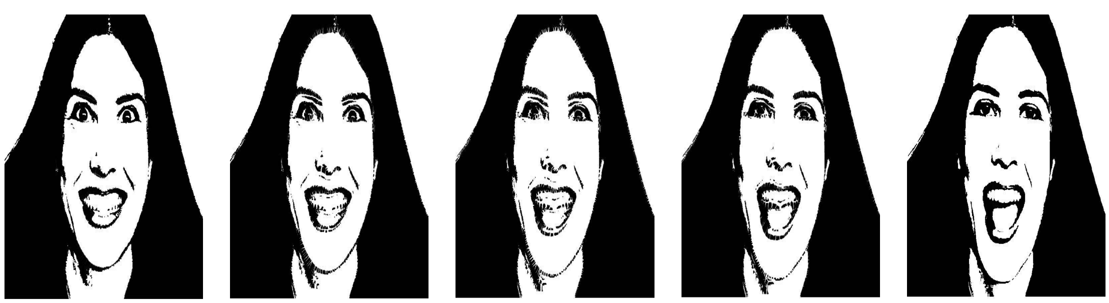
In this paper, a new method to produce NPR facial expression animation has been introduced. Two parts: portrait generation and metamorphing animation, consist of the whole system. Different with previous picture generation systems, facial expressive details, such as: fine creases and expressive wrinkles etc.., can be included in our generated

(a) frame1 (b) frame15 (c) frame30 (d) frame45 (e) frame60

Figure 10. Slices morphing from happiness to surprise in line drawing style, 15 frames as the interval

(a) (b) (c)

Figure 11. (a)The Input Photograph; (b)The Original Generated Portrait from (a);(c)The Portrait with feature lines mapped in LC and RC.



(a) frame1 (b) frame15 (c) frame30 (d) frame45 (e) frame60

Figure 12. Slices morphing from happiness to surprise in pen-and ink style, 15 frames as the interval

portrait, which makes the whole system more flexible. For animation part, a 2D morphing algorithm using distance transform is utilized here. Without any control points or lines, pixel correspondence and mapping function are created automatically by distance transform. Given two human face images as input, our system outputs a particular NPR style facial animation as the result, which can be applied to lots of practical applications. Currently, we are planning to improve and extend our work to three dimensional, which will make the resulted animation more interesting.

References

[1] F.I.Parke, *A Parameterized model for Human Faces*, PhD thesis, University of Utah, Salt Lake City, UT, December 1974

[2] Yuanzhong Li, Hidefumi Kobatake, *Extraction of Facial Sketch Image Based On Morphological Processing*, Proceedings of International Conference on Image Processing (ICIP '97), Vol.3, pp: 316-319, Washington, DC, 1997

[3] Hong Chen, Ying-Qing Xu, Heung-Yeung Shum, Song-Chun Zhu,Nan-Ning Zheng, *Example-based Facial Sketch Generation with Non-parametric Sampling*, *Computer Vision*, ICCV 2001. Proceedings of Eighth IEEE International Conference on Computer Vision, Vol.2, pp: 433-439, July 2001

[4] Hong Chen, Ziqiang Liu, Chunk Rose, Yingqing Xu, Heung-Yeung Shum, David Salesin, *Example-based Composite Sketching of Human Portraits*, Proceedings of the 3rd international symposium on Non-photorealistic animation and rendering, Annecy, France, pp.95-153

[5] Bruce Gooch, Erik Reinhard,Amy Gooch, *Human Facial Illustrations: Creation and Psychophysical Evalution*, ACM Transactions on Graphics,Vol.23, No. 1, pp: 27-44, January 2004

[6] Tatiana Surazhsky and Gershon Elber, *Metamorphosis of Planar Parametric Curves via Curvature Interpolation*, International Journal of Shape Modeling, Vol. 8, No.2, pp: 201-216, 2002

[7] Beier T, Neely S, *Feature-based image metamorphosis*, Computer Graphics, Vol.26, No.2, pp: 35-42, 1992

[8] Lee T, Lin Y-C, Lin L, Sun Y, *Fast Feature-based metamorphosis and operator design*, In Proceeding of Eurographics'98, Computer Graphics Forum, Vol.17, No.3, pp:15-22,1998

[9] Sousa,M.C., Buchanan,J.W., *Observational model of blenders and erasers in computer-generated pencil-and-ink illustration*, In Graphics Interface '99.ACM, pp: 157-166, New York, 1999

[10] Durand,F.,Ostromoukhov,V.,Miller,M.,Duraneleau,-F.,Dorsey,J., *Decoupling strokes and high-level attributes for interactive traditional drawing*, In Rendering Techniques 2001:12th Eurographics Workshop on Rendering, Eurographics, pp: 71-82

[11] Otsu, N., *A Threshold Selection Method from Gray-Level*,IEEE Transactions on Systems, Man, and Cybernetics, Vol.9, No.1, pp. 62-66, 1979

[12] J. R. Parker, *A System for Fast Erosion and Dilation of Bi-Level Images*, *Journal of Scientific Computing*, Vol.5, No.3, pp.187-198, 1991

[13] Marina L. Gavrilova, Muhammad H. Alsuwaiyel, *Two Algorithms for Computing the Euclidean Distance Transform*, *International Journal of Image Graphics*, Vol.1, No.4, pp.635-645, 2001