An Efficient Virtual Aesthetic Surgery Model Based on 2D Color Photograph

Hyun Park\textsuperscript{1}, Kee Wook Rim\textsuperscript{2}, and Young Shik Moon\textsuperscript{1}

\textsuperscript{1} Department of Computer Science and Engineering, Hanyang University, 1271 Sa-Dong, Ansan, Kyunggi-Do 425-791, Korea (hpark, ysmoon)@cse.hanyang.ac.kr
\textsuperscript{2} Department of Knowledge and Industrial Engineering, Sun Moon University, Chung-Nam, Korea rim@sunmoon.ac.kr

Abstract. In this paper, we propose a virtual aesthetic surgery (VAS) system using a deformation technique based on a radial basis function (RBF) and blending technique that combines the deformed facial component with the original face. The proposed VAS system is composed of three main steps. First, various deformation templates are matched to facial components by a multi-resolution active appearance model (MAAM), which is trained by 2D color face images. Next, the VAS system computes the degree of deformation for lattice cells on the free-form deformation (FFD) using the proposed RBF. The deformation error is compensated for by the coefficients of the mapping function, which is recursively solved by the singular value decomposition (SVD) technique using the sum of squared error (SSE) between the deformed control points and target control points on the base curves. Finally, the deformed facial component is blended with the original face using a blending ratio that is computed by the modified Euclidean distance transform. Experimental results show that the proposed deformation and blending techniques are very efficient in terms of smoothness, accuracy, and distortion.

1 Introduction

Recently, many people have shown an interest in their facial appearance due to images propagated by the mass media. The development of modern medical technology can now satisfy an individual's needs for altering their appearance through aesthetic surgery.

Even if a face is only slightly deformed, the overall facial appearance may look more affected [1]. Thus, individuals who want to undergo aesthetic surgery need a VAS system that can realistically predict the appearance of altered features after surgery. The proposed VAS system consists of deformation template tools and various filters (double eyelid filter, skin care filter, etc.) for aesthetic effects.

Two-dimensional (2D) or three-dimensional (3D) face deformation models, which are the basic techniques of a computer-based virtual surgery system, typically fall into two main categories: (i) physically/anatomically based models and (ii) non-physically based models using morphing or special deformation [2].

© Springer-Verlag Berlin Heidelberg 2005
Physically-based models attempt to model the structure and function of the muscles and skin of the face using a mass-spring or finite element model [1], [2]. Since physically-based models can predict the results of virtual surgery with high accuracy, these models are used for plastic reconstructive surgery systems, which can virtually correct craniofacial deformities [3], [4], [5]. However, physically-based models have some drawbacks. Typically, it is very difficult to define physical models for faces, they require a large amount of computation time, and do not properly handle facial color and textures. These models are used on a limited basis for facial surgery of some parts of faces.

Non-physically based models use generic animation techniques, such as morphing and spatial deformations, to create facial expressions or deformation. Lee et al. described a morphing technique based on a multilevel free-form deformation (MFFD). The MFFD model is controlled by a set of feature points that place positional constraints on the MFFD lattice [6], [7]. The MFFD model affects or distorts facial components other than the deformation region because it uses only a global deformation. Also, MFFD does not properly handle the facial texture distortion. Lin et al. described a deformation technique that used radial basis functions (RBFs) and the displacements of feature points for facial expression and animation [8]. This model also has drawbacks: it does not produce an accurate facial deformation and the facial components outside of the deformation region are also distorted. Noh et al. described a deformation technique for 3D facial expression. This deformation technique, after defining deformable regions on lattice cells for a particular face, computes the degree of deformation by using RBFs and the displacements of feature points [9]. Although this deformation technique has locality, it has the same drawbacks as Lin’s method.

The VAS system requires four main characteristics: (i) convenience in representing virtual surgery regions and using deformation tools; (ii) smoothness and accuracy in deforming facial components (eye, nose, mouth, jaws, etc.); (iii) locality that does not affect or distort other facial components outside of the deformation region; and (iv) preservation of the facial color texture.

To satisfy these four characteristics, the proposed VAS system is composed of three main steps. First, various deformation templates are semi-automatically matched to facial components by the multi-resolution active appearance model (MAAM) that is trained by 2D color face images. As in the original AAM (active appearance model) method, this appearance model will be constructed with multiple resolutions to provide coarse-to-fine fitting. Next, the VAS system computes the degree of deformation for lattice cells on the free-form deformation (FFD) using the proposed RBF. As RBFs can compute the degree of deformation efficiently with less computing time, RBFs are typically used for facial deformations that require real time processing. The deformation error is compensated for by the coefficients of the mapping function, which is recursively solved by the singular value decomposition (SVD) technique using the sum of squared error (SSE) between the deformed control points and target control points on base curves. Finally, a blending technique is proposed to minimize the distortion of facial color textures, which are important in the VAS system. The blending technique combines the deformed facial component with the original face by using a blending ratio that is computed by the modified Euclidean distance transform.
In this paper, the proposed facial deformation framework uses a flexible framework that is proposed by Pixar Animation Studios and Princeton University [10]. Figure 1 shows the total framework of the proposed VAS system.

![Framework of the proposed VAS system](image)

**Fig. 1. Framework of the proposed VAS system**

### 2 Search-Based Matching of Various Deformation Templates

In the VAS system, it is not easy for an operator to indicate facial components manually using a pointing device such as a mouse. Therefore, a semi-automatic matching step for deformation templates is provided to help the operator indicate the deformation regions for virtual aesthetic surgery easily and efficiently.

#### 2.1 Training of Face Images

The data set consists of 100 still images of 100 different frontal-view human faces, all without glasses and with a neutral expression. Images are acquired in 1187 x 1190
bitmap color format. As shown in Figure 2(b), the facial structures are manually annotated using 72 total landmarks of the eyebrows, eyes, nose, mouth, and jaw. In this paper, the training method for AAM uses the Jacobian learning scheme. As in the original AAM method, these AAMs are built at each level of a scale-pyramid for coarse-to-fine fitting based on multi-resolution [11], [12].

Fig. 2. Face image set for training the AAM. (a) Frontal human color texture. (b) Facial landmarks.

Fig. 3. Average texture and shape on a trained AAM. (a) Average texture. (b) Average shape.

Fig. 4. Search-based initialization of deformation templates using multi-resolution AAM. (a) Original image. (b) Search result in level 2 (50% scaled down of level 1). (c) Search result in level 1 (50% scaled down of level 0). (d) Final search result in level 0 (not scaled down).

2.2 Template Matching by Multi-resolution Active Appearance Model

In this paper, to improve the efficiency and robustness of the matching algorithm, various deformation templates were matched by using the MAAM based on color images. This involves searching for the object in a coarse image first, then refining the location in a series of finer resolution images [13].

As shown in Figure 4, a multi-resolution pyramid is scaled into three levels. Figure 4(d) is the final search result in which parameters of the combined model for

3 Proposed Del

First, we briefly describe the degree of three characteristics for the system, we define the more accurate, greater accuracy.

3.1 RBF Interpol

RBFs define the interpolation basis functions, each being a smooth, continuous function.

The RBF mapping is given in equation: $f(x) = \sum_{i=1}^{n} r_i \phi_{g_i}(x)$

where $k$ is the distance between points, $r_i$ is the Euclidean distance, $\phi_{g_i}(x)$ is the global basis mapping to the basis.

3.2 Radial Basis Functions

In this paper, we used a combination of the B-spline technique of the AAM. [1] It should be able to correct the control point [2] It should maintain the identity of the face to the basis [3] It should have a higher resolution, smoother than a

Inverse multi-resolution has the above the
AAM were optimized (translation, scaling, rotation, texture model parameters, and shape model parameters).

3 Proposed Deformation Technique for Virtual Aesthetic Surgery

First, we briefly describe the properties of the RBF interpolation that actually computes the degree of lattice cell deformation on the FFD. Second, after suggesting the three characteristics that are required for the deformation technique using the VAS system, we define an RBF mapping function to satisfy those characteristics. Finally, we describe the technique that compensates for deformation errors, which results in greater accuracy.

3.1 RBF Interpolation

RBFs define the interpolation functions as a linear combination of radial symmetric basis functions, each of which is centered on a particular control-point. RBFs are smooth, continuous functions that provide at least $C^1$ continuity [7].

The RBF mapping function can be decomposed into a global and local component, as given in equation (1). Although the two components are distinct, they can be computed almost simultaneously.

$$f_k(x) = P_{nk}(x) + \sum_{i=1}^{n} A_i g(r_i), \quad r_i = \|x - x_i\|$$

(1)

Here $k$ is the dimension, $m$ is the degree of polynomial, $n$ is the number of control points, $r_i$ is the Euclidean norm between a point $x$ and a source control point $x_i$, and $P_{nk}(x)$ is the global component. If the degree of polynomial is 1, it yields the global component to an affine transformation, as in equation (2) [14].

$$P_{1k}(x) = a_{0k} + a_{1k}x + a_{2k}y$$

(2)

3.2 Radial Basis Function for VAS System

In this paper, we suggest three characteristics that are required for the deformation technique of the VAS system.

[1] It should achieve the consistency of deformation regardless of the number of control points on base curves (source curve and target curve).

[2] It should maintain the visual smoothness and the accuracy of deformation according to the base curves.

[3] It should have a locality to limit the range of deformation influence so that other facial components are not affected.

Inverse multiquadric RBF has local attributes similar to a Gaussian RBF, but it is smoother than a Gaussian RBF. Therefore, we define a deformation technique that has the above three characteristics by using an inverse multiquadric RBF. As shown
in equation (3), the inverse multiquadric RBF is represented by a formula with a stiffness constant \( s_i \) that regulates the local or global effects of the control points, where \( \mu \) is a constant [9], [14], [15]. In this paper, we design an inverse stiffness that increases the deformation force as the control points are spread apart, and decreases the deformation force when the control points are closer to each other.

\[
g(r) = (r^2 + s_i^2)^{-\mu}, \quad \mu > 0
\]  

(3)

The inverse stiffness is computed by equation (4), according to the adaptive method proposed by Ruprecht and Müller [16].

\[
Inverse \ s_i = \max_{i \neq j} \left( \frac{x_j - x_i}{N} \right)
\]

(4)

Here, \( x_{source \ position} \) is a control point on the base curve (source curve), and \( N \) is the number of control points on the base curve.

3.3 Compensation of Deformation Error

For facial deformation, 2(n+3) coefficients of the basis function and the polynomial are solved by the singular value decomposition (SVD) technique. The coefficients are used for a lattice cell deformation on the FFD.

\[
SSE = \sum_{i=1}^{n} \epsilon_i
\]

Here, \( x_{deformed \ position} \) is the position of control points on lattice cells on the FFD iteratively until the SSI components are deformed lattice cells.

4 Proposed Blend

The following section describes the computation of facial color, texture, and the deformation region.

4.1 Blending Ratio Calculation

In our VAS system, the components outside of the RBF with a perfect local area or distort some facial component in the input data. Global deformation makes the result image visually smooth in the fitting of global deformation.

First, blending mask formation templates. I distance transformation from each point [17], component with the other components.

As shown in equation (5), to compute the blend:

\[
D_{fuel} \quad B_{ratio} \quad I_{result} \quad B_{ratio}
\]

As shown in equation (6), to compute the blend:

\[
D_{fuel} = \frac{1}{1 + e^{-B_{ratio} \cdot I_{result}}}
\]
\[ \text{SSE} = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (x_i^{\text{deformed position}} - x_i^{\text{target position}})^2 \]  

Here \( x_i^{\text{deformed position}} \) is a position of deformed control points and \( x_i^{\text{target position}} \) is a position of control points on base curve (target curve). Therefore, the deformation error of lattice cells on the FFD is compensated by the coefficients that are solved by SVD iteratively until the SSE value becomes less than a threshold value. Finally, facial components are deformed by a two-pass spline mesh warping, based on the compensated lattice cells.

### 4 Proposed Blending Technique for Virtual Aesthetic Surgery

The following section describes the blending technique used to minimize the distortion of facial color, texture, and the distortion of facial components in the remainder of the deformation region.

#### 4.1 Blending Ratio Computed by Modified Euclidean Distance Transform

In our VAS system, the deformation technique does not affect or distort facial components outside of the deformation region. However, since it is difficult to define a RBF with a perfect locality, the deformation of a specific facial component may affect or distort some facial components. Generally, a global deformation makes the resulting image visually smooth, but globally distorted. On the other hand, a local deformation makes the resulting image without global distortion, but does not produce a visually smooth image. In this paper, we suggest a blending technique that has the advantages of global deformation and local deformation.

First, blending mask regions on the face are automatically computed by using deformation templates. Next, the blending ratio is computed by the modified Euclidean distance transformation (EDT) that computes the distance to the closest boundary from each point [17]. Finally, the blending technique combines the deformed facial component with the original face by using the blending ratio.

As shown in equation (6), the proposed blending technique uses the modified EDT to compute the blending ratio.

As shown in equation (6), the proposed blending technique uses the modified EDT to compute the blending ratio.

\[ D_{\text{Euclidean}} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \]

\[ B_{\text{ratio}} = \text{Scaling Factor} \cdot D_{\text{Euclidean}} \]

\[ I_{\text{result}} = (255 - B_{\text{ratio}}) \cdot I_{\text{original}} + B_{\text{ratio}} \cdot I_{\text{deformed}} \]

\[ B_{\text{ratio}} = \text{Scaling Factor} \cdot D_{\text{Euclidean}} \]
Here $D_{Eulidean}$ is the distance to the closest boundary from each point, $B_{ratio}$ is the blending ratio on the facial mask region, $I_{original}$ is the original face, $I_{deformed}$ is the deformed facial component, $I_{result}$ is the final result that combines the deformed facial component with the original face using the blending ratio, and the Scaling Factor is a ratio that varies the $B_{ratio}$ value in the range from 0 to 255.

In Figure 6, the resulting image shows a visual smoothness without global distortion over all facial components.

Figure 6. Deformation using the blending technique. (a) Original image. (b) Deformation template on nose. (c) Deformation of nose. (d) Mask with blending ratio. (e) Blending of the deformed facial component and the original face. (f) Final result.

Figure 7 shows the degree of texture distortion on a face with a complex texture (pockmarks, pimples, blotches, etc.). Figure 7(b) shows that a local deformation distorts the texture around the nose region. Figure 7(c) shows that the proposed deformation and blending techniques are very efficient in terms of smoothness and distortion for color images of faces with complex textures.

Figure 7. Comparison of the degree of distortion for a color texture (a) Original image. (b) Local deformation. (c) Deformation using the blending technique.

5 Experimental

In this paper, we considered body splines (GEB) as an inverse multiquadric function used as a basis function. Gaussian RBF performed well but was difficult to choose the scaling factor of the proposed V2.

To satisfy the three criteria of the proposed V2, we used an adaptive stiffness:

\[ \sigma_i = \frac{1}{\sqrt{\sum_{j=1}^{n} (x_j - c)^2}} \]

Figure 8 shows the Gaussian RBF result for these three methods.

Figure 8. Comparison of deformation template (a) Original image. (b) Local deformation. (c) Result using the proposed RBF.
5 Experimental Results

In this paper, we compared the Gaussian RBF of equation (7) used in Gaussian elastic body splines (GEBS), the general Gaussian RBF of equation (8), and the proposed inverse multiquadric RBF with inverse stiffness [18]. When a Gaussian function is used as a basis function, the deviation value must be carefully chosen because the Gaussian RBF performs poorly without a good deviation value. However, it is very difficult to choose the best deviation value.

To satisfy the three characteristics that are required for the deformation technique of the proposed VAS system, the deviation value of Gaussian RBFs is calculated by an adaptive stiffness, as given in equation (9).

\[
g(r) = \frac{1}{(\sqrt{2\pi}\sigma)^3} e^{-\frac{r^2}{2\sigma^2}} \quad (7)
\]

\[
g(r) = \frac{1}{\sigma^2} \quad (8)
\]

\[
\sigma_i = \min_{i \neq j} \left( \frac{1}{|x_j - x_i|} \right) \quad (j = 1, \ldots, n) \quad (9)
\]

Figure 8 shows that the deformation accuracy of the proposed method is better than the Gaussian RBF of GEBS and the general Gaussian RBF. The deformation errors for these three methods are shown in Figure 8 (d), (f), and (h).

---

**Fig. 8.** Comparison of deformation results on various RBFs. (a) Original face image. (b) Deformation template. (c) Result using the Gaussian RBF of GEBS. (d) Deformation error in (c). (e) Result using the general Gaussian RBF. (f) Deformation error in (e). (g) Result using the proposed RBF. (h) Deformation error in (g).
In this paper, we estimate the deformation accuracy by using equations (10) and (11) [19].

\[
CDR(\text{Correct Deformation Ratio}) = 1 - DR
\]  
(10)

\[
DR(\text{Distortion Ratio}) = \frac{\text{Distortion Area}}{\text{Target Template Area}}
\]  
(11)

Table 1 shows that the proposed RBF is more accurate than Gaussian type RBF.

Table 1. Comparison of deformation accuracy for various RBFs

<table>
<thead>
<tr>
<th>Type of RBF</th>
<th>CDR (correct deformation ratio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian RBF of equation(7)</td>
<td>0.9619</td>
</tr>
<tr>
<td>Gaussian RBF of equation(8)</td>
<td>0.9608</td>
</tr>
<tr>
<td>Proposed RBF</td>
<td>0.9923</td>
</tr>
</tbody>
</table>

6 Conclusion

In this paper, we proposed a semi-automatic initialization of deformation templates, facial deformation, and blending techniques that are suitable for the VAS system. The choice of the best RBF to use depends primarily on the application objectives.

![Fig. 9. Final results of the proposed deformation and blending technique. (a) Original image. (b) Result that corrects the nose and jaw. (c) Original image. (d) Aesthetic result using an eyelid filter and jaw correction.](image)
We carried out experiments that compared various RBFs in order to develop a deformation technique for a VAS system. It has also been found that RBFs that have Gaussian function form are very difficult to control.

We have shown that the deformation technique using an inverse multiquadric RBF with inverse stiffness is suitable for a VAS system, and that the blending ratio computed by the Euclidean distance transform can be used to produce a good deformation result.

As shown in Figure 9, the proposed deformation technique achieves excellent deformation results on a face image as well as a natural image.

Acknowledgement

This research was supported by the MIC (Ministry of Information and Communication) of Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Assessment).

References