ABSTRACT

Image morphing, which is a 2D imaging technique, allows smooth transition between images. However, a limitation of existing image morphing techniques is the lack of user interaction. Another limitation is that shape warping often causes distortion due to barycentric mapping. In this paper, we present a novel 3D morphological technique to address these problems. A new concept of relief occlusion-adaptive meshes is introduced, allowing a user to change the viewpoint of the output images in the morphing process. By making use of the intrinsic geometric relationship among the reference images for projective transformation, the distortion due to barycentric mapping is prevented. Unlike other warping-based view transferring techniques, our morphological technique provides very smooth transition between reference images and supports both rigid and non-rigid scenes.

1. INTRODUCTION

Image morphing [3,4,7] is a technique for producing smooth transition between images. Novel views are generated by combining shape warping with attribute blending for graphical objects (GO) [5]. During the morphing process, the source image is gradually deformed and faded out, while the destination image is gradually faded in. The output images generated depend only on a single parameter \( \lambda \in [0,1] \). An animation video can be created by mapping time \( t \) with \( \lambda \) and varying its value. This process is simple and efficient as both shape warping and attribute blending are performed in 2D image domain. However, these 2D GOs are in fact projection of 3D GOs. Thus, 3D distortions or unnatural image transitions may occur if naive image morphing techniques are used. View morphing [10] extends the concept of image morphing and achieves more smooth and natural image transition by considering 2D image point as a projection of 3D scene. However, it does not address the image partitioning and visibility problems.

In [8], an automatic image interpolation/morphing technique was proposed. It introduces the joint view triangulation (JVT) method for image partitioning and matching. The major limitation of the JVT method is that the viewpoint is restricted to the line joining the viewpoints of the source and the destination images.

In order to warp points to arbitrary views, [9] has derived the warping equation to transfer image points from a reference image to any arbitrary view, given the generalized disparity of the reference image. In [2], a method for transferring points in tensor space with full correspondences in the reference images is introduced. By applying the transformation operator on a given seed tensor of three reference views, it can obtain a chain of warping functions from a set of reference images to warp image points. Since these techniques perform only point warping but not attribute blending, they do not address the problem of smooth transition between reference images.

In general, existing image morphing techniques have restricted output viewpoint and thus arbitrary view synthesis is not supported. In addition, the output of the morphological function depends only on a single input parameter (usually on time \( t \)), which limits the application of the techniques to video or other time dependent applications. Hence, the user only has a limited interaction with the system to obtain a particular output image from the morphological function. In addition, although some image point warping techniques provide solutions for transferring image points to any arbitrary views, they require the disparity or correspondence information of all transfer points, which may be achievable for rigid objects but not for non-rigid scene.

In this paper, we present a novel 3D morphological technique. By using relief occlusion-adaptive meshes, the new method supports arbitrary view synthesis (including translation and rotation) with three reference images, view interpolation, and view extrapolation. It also provides smooth transition among the reference images and supports both rigid and non-rigid scenes.

There are a few problems that we need to address here. First, we need to relax the viewpoint constraint from the line joining the source and the destination viewpoints to any arbitrary viewpoint and view position. Second, since most images will likely contain some occluded or non-rigid regions, it may not be possible to establish...
correspondences for them. We need to have a representation scheme for the morphing operation to be performed in 3D spatial domain with 2D arbitrary images. Third, we need to consider the distortion and fold over problems for triangulation with barycentric mapping. Finally, we need to minimize the angular deviation of the ray between the reference images and the novel view. We will discuss these issues and our proposed solutions here.

The rest of the paper is organized as follows. Section 2 presents in detail our relief occlusion-adaptive meshes and the corresponding morphing technique. Section 3 shows some experimental results of the proposed technique. Section 4 concludes the paper with a discussion on possible future work.

2. METHODOLOGY

In this section, we first introduce our relief occlusion-adaptive meshes as the representation scheme. We then discuss the proposed morphing technique in detail.

2.1 Relief Occlusion-Adaptive Meshes

Representation scheme for registered images is indispensable for morphological operations. However, the fundamental limitation of using standard mesh models is that it enforces continuity of motion across the whole image and hence, object discontinuity cannot be represented. To represent object discontinuity and occluded regions, we apply the concept of occlusion-adaptive meshes [1] here such that occluded or unmatched regions can be represented adaptively without the need for corresponding regions in another image. We partition the image into triangular patches and match them with other reference images. The set of matched patches \( M \) and the set of non-matched patches \( N \) are represented separately, where \( M \cap N = \emptyset \). In addition, in our relief occlusion-adaptive mesh scheme, we compute the generalized disparity [9] for each node such that the nodes can be transferred to arbitrary viewpoints with the warping equation, supporting arbitrary view synthesis.

Our morphological operation is by constructing relief occlusion-adaptive meshes among three reference images \( I_0, I_1, I_2 \in \mathbb{R}^2 \). We first extract a set of feature points for image \( I_0 \) and then partition the image through delaunay triangulation on the feature points. The partitioned triangular patches in image \( I_0 \) are matched with those in images \( I_1 \) and \( I_2 \) to obtain a set of matched triangular patches. For the non-matched regions, we perform edge-constrained delaunay triangulation on each image.

In order to compute the generalized disparity for each node, we use linear triangulation method to get the maximum likelihood estimate of a 3D point \( X \) from its 2D correspondences \( x, x' \) and \( x'' \) in reference images \( I_0, I_1, \) and \( I_2 \), respectively. Let \( P, P' \) and \( P'' \) be the 3x4 camera matrices of images \( I_0, I_1, \) and \( I_2 \), respectively. Then, we can form the equation: \( AX = 0 \), where

\[
A = \begin{bmatrix}
p_3 & -p_1 & 0 & 0 
p_3 & -p_1 & 0 & 0 
p_3 & -p_1 & 0 & 0 
p_3 & -p_1 & 0 & 0 
p_3 & -p_1 & 0 & 0 
p_3 & -p_1 & 0 & 0 
\end{bmatrix}
\]

\[
P = \begin{bmatrix}
p' 
p' 
p' 
p' 
p'' 
p'' 
\end{bmatrix}
\]

\[
x = (x, y, 1)^T
\]

We do singular value decomposition on \( A \) and get the least square solution of \( X \) by finding the solution as unit singular vector corresponding to the singular value of \( A \). After \( X \) is obtained, the generalized disparity, \( \delta(x) \), for 2D image point \( x \) can be calculated as:

\[
\delta(x) = \frac{\|Px\|}{\|C - X\|}, \text{ where } C = \begin{bmatrix} C \\ 1 \end{bmatrix}, X = \begin{bmatrix} X \\ 1 \end{bmatrix}
\]

2.2 Shape Warping

When the user’s viewpoint is changed, features points are first warped from the reference images to the output image with the warping equation defined as follows:

\[
x' = \delta(x)N^{-1}(C - C') + N^{-1}N_x
\]

where \( x \) and \( x' \) are points in the reference image and the output image, respectively. \( C \) and \( C' \) are the centers of projection of the reference image and of the output image. \( N \) and \( N' \) are the 3x3 projection matrices of the reference image and of the output image. \( \delta(x) \) is the generalized disparity at point \( x \). This warping equation warps feature points without the need to perform 3D perspective transformation.

2.3 Projective Transformation

After the feature points are warped, images points within each triangular patch can be transferred. However, the use of barycentric mapping [6] for transformation may lead to potential transformation distortion. An alternative method is to use projective transformation, which can model the transformation of G0s in 3D spatial domain and thus avoid distortion. However, it requires 4 matches (\( u_i, u_i' \)) between two images to compute the plane homography \( H \) (3x3 matrix) for the transformation as the degrees of freedom (DoF) of \( H \) is 8. Thus, the shape of the patches needs to be quadrangle instead of triangular.

Here, we address this homography computation problem by proposing a new way of calculating the plane homography from triangular patches based on intrinsic geometric relationship among the reference images. Referring to figure 1, since a 3D triangle may define a plane in the projective domain, the plane homography can be obtained even the shape of the patch is triangular.

Let \( H_1 \) be a 4x3 transformation matrix that transforms a homogenous 2D point into the projective 3D space and \( H_2 \) be a 3x4 transformation matrix that transforms a
homogenous 3D point into a 2D point in image 2. The plane homography $H$ is then given as:
$$ x' = Hx = H_1H_2x $$
(1)

$H_1$ of equation (1) can be given by:
$$ H_1 = \begin{pmatrix} (I + \frac{C_1^T}{\pi^T C_1 + 1}) N \\ \frac{\pi^T C_1}{\pi^T C_1 + 1} \end{pmatrix} $$
(2)

where $C$ is the centre of projection, a homogenous 3D point. $\pi$ is the projective 3D plane formed, which can be computed from three projective 3D points, and $\pi^T = [\pi^T, 1]$. $H_2$ is the camera matrix of image 2, $P^*$. 

![Figure 1. Implicit geometry of a triangular patch.](image)

**2.4 Texture Blending**

In our method, we use three reference images. The texture blending coefficients to produce the output image from the reference images are determined according to the viewing angles of the output image and the reference cameras. A ray is first projected from the output view $D$ to the center of the patch (e.g., $p_1$). Rays are also projected from the reference cameras $C_1$, $C_2$ and $C_3$ to the patch. We refer to the angles between the output ray $(D, p_1)$ and each of the camera rays $(C_1, p_1)$, $(C_2, p_1)$ and $(C_3, p_1)$ as $\theta_1$, $\theta_2$ and $\theta_3$, respectively. We select the patches from the cameras with the smallest angle to ray $(D, p_1)$ and those to ray $(D, p_2)$ for texture blending. For example, for patch $p_1$, the patches from cameras $C_1$ and $C_2$ would be selected. However, for patch $p_2$, the patches from cameras $C_2$ and $C_3$ would be selected. The texture blending coefficients for patch $p_1$ can be obtained as follows:

$$ \lambda_1 = \frac{\theta_2}{\theta_1 + \theta_2}, \quad \lambda_2 = 1 - \lambda_1 \quad \text{and} \quad \lambda_3 = 0 $$

where $\lambda_1$, $\lambda_2$ and $\lambda_3$ are the texture blending coefficients of cameras $C_1$, $C_2$ and $C_3$ on patch $p_1$, respectively.

After computing the texture blending coefficients, the image attributes can be combined to produce the output image. Unmatched triangles are drawn first as they may contain occluded areas. Let the three vertices of a triangle in image 1 be $v_1, v_2, v_3$ and the area of the triangle in image 1 be $s_1 = ||v_1v_2v_3||$. The texture blending coefficient for each pixel of the triangle, $\lambda$, is approximated as the average value of the texture blending coefficients of its three vertices. The resulting value of a pixel in the output image $I$ is:

$$ I = \frac{\lambda_1 s_1 I_1 + \lambda_2 s_2 I_2 + \lambda_3 s_3 I_3}{\lambda_1 s_1 + \lambda_2 s_2 + \lambda_3 s_3} $$

(3)

Since illumination can be view-dependent, the texture for the matched patches is obtained from all three images instead of only one image to provide smooth blending between images.

![Figure 2. Selecting appropriate patches for the output view.](image)

**3. RESULTS AND DISCUSSION**

We have implemented the new method in Java and conducted a number of experiments on it. Due to page limitation, we only show one experiment here, which was conducted on a PC with a Pentium 4 2.2GHz processor. Three reference images with resolution 640x480 were used and are as shown in figure 3. An automatic tool similar to the joint view triangulation (JVT) method [8] was used to partition and to match the reference images. The registered images are then used for morphing. Figure 4 shows several arbitrary views generated by the proposed method. As can be seen, both interpolation and extrapolation are supported by the new method.

As pointed out by [10], the output quality of existing morphing techniques is significantly affected by the visibility of the reference images. Since the output arbitrary view may contain regions that are not visible in the reference images, it is difficult to predict the pixel flow to provide smooth transition of those regions. From figure 4, we can see that some regions in the three reference images are only visible in one image and therefore cannot be matched, e.g., the rectangular region. These regions may lead to a sudden change in the output contents. From the output images, we may see that the problem is effectively alleviated.

In terms of rendering time, it takes on average 0.5s to render an image of resolution 640x480. To improve output image quality, bilinear interpolation can be used in pixel warping. However, the rendering time will increase to approximately 0.7s. If we reduce the image resolution
to 320x240, the processing time can be reduced significantly to about 0.17s. We can see that the rendering performance depends mainly on the image resolution.

4. CONCLUSION AND FUTURE WORK

Although image morphing can produce smooth image transition and appealing visual effects, existing morphing techniques lack user interaction. In this paper, a novel 3D morphological technique and an appropriate representation scheme are presented. The technique allows morphing to be performed in 3D spatial domain, even though the reference images are in 2D. The new representation scheme of relief occlusion-adaptive meshes is designed such that object discontinuities can be represented and generalized disparity is included to allow images of arbitrary viewpoints to be synthesized.

Because of these advantages, the proposed morphological technique extend the use of morphing to applications with user interaction (e.g., 3D navigation). An important area of future work is to extend the technique for image-based rendering (IBR).

5. ACKNOWLEDGEMENTS

The work described in this paper was partially supported by a DAG grant from City University of Hong Kong (Project Number: 7100264).

6. REFERENCES


Figure 3. The input reference images 1, 2 and 3.

Figure 4. Images synthesized with the proposed method.