

# 用于人脸表情图像与动画合成的图像变形方法

李旭东<sup>1</sup>

(1 浙江财经学院数学与统计学院,杭州市学源街 18 号,310018)

**摘 要:**基于传统图像变形的人脸表情合成方法,在对由闭口的表情图像到开口露齿的表情图像进行中间状态的表情合成时会失效。为克服这一不足,提出了一种更适用于人脸表情图像与动画合成的图像变形方法。该方法引入了掩蔽图像的概念,并提出了对掩蔽图像的变形方法,修正了传统图像变形方法的变形公式。实验结果显示,该方法能自然、逼真地合成出由闭口的表情图像到开口露齿的表情图像的中间状态表情图像。

**关键词:**图像变形;人脸表情合成;虚拟现实;计算机动画

**中图法分类号:**P237.4

根据所基于的图像特征的形式,图像变形方法可以分为三类:基于特征网格的方法<sup>[1-3]</sup>、基于特征线的方法<sup>[4]</sup>、基于特征点的方法<sup>[5-11]</sup>,其中,基于特征点的方法能产生自然、逼真的图像变形效果,并且对图像特征的控制非常方便和灵活,是目前较为常用的图像变形方法。

利用图像变形方法可以实现人脸表情的合成<sup>[12-14]</sup>,即对一个人的两个不同表情的脸部图像,通过图像变形方法合成出介于这两种表情之间的表情图像。这类人脸表情合成方法能有效地合成出自然、逼真的表情图像与动画,但是,由于传统的图像变形方法要求两个图像比较类似或者具有相同的拓扑结构,这就使得基于图像变形的人脸表情合成方法所能合成出的表情图像的范围受到了一定的限制,对某些表情图像的合成会失效。本文提出了一种更适用于人脸表情图像与动画合成的图像变形方法。

## 1 图像变形方法基础

一般情况下,图像变形可以分为三步:① 计算源图像和目标图像中的每一个像素点的变形偏移量,即通过图像上有限的特征点的坐标变化来确定整幅图像像素点的坐标变化;② 对源图像和目标图像的每一个像素点的灰度值分别进行重采样,以构造出其扭曲图像;③ 将源图像和目标图

像的扭曲图像线性组合成变形图像。

假设给定源图像  $I_0$  和目标图像  $I_1$ ,并设源图像和目标图像的  $n$  对对应特征点的坐标分别为:

$$p_0^{(i)} = \begin{bmatrix} p_0^{(i1)} \\ p_0^{(i2)} \end{bmatrix}, p_1^{(i)} = \begin{bmatrix} p_1^{(i1)} \\ p_1^{(i2)} \end{bmatrix}, i = 1, 2, \dots, n$$

则图像变形过程中  $t$  时刻的变形图像  $I_t$  相应特征点的坐标为:

$$p_t^{(i)} = (1 - t) p_0^{(i)} + t p_1^{(i)}, i = 1, 2, \dots, n$$

其中,  $t \in [0, 1]$ , 当  $t=0$  时,即为源图像所处的时刻;当  $t=1$  时,即为目标图像所处的时刻。则根据此时的特征点分布和源图像的特征点分布,利用插值方法即可获得源图像每个像素的变形偏移量,接着对像素点的灰度值进行重采样后即可获得源图像  $t$  时刻的扭曲图像  $I_t^{(0)}$ ;同理,可获得目标图像  $t$  时刻的扭曲图像  $I_t^{(1)}$ ,从而获得  $t$  时刻的变形图像<sup>[8]</sup>:

$$I_t = (1 - t) I_t^{(0)} + t I_t^{(1)} \quad (1)$$

## 2 本文的图像变形方法

### 2.1 掩蔽模型

为使图像变形方法能更有效地应用于人脸表情的合成,需对式(1)进行修正。为此,本文首先提出一个掩蔽模型。

掩蔽模型的基本思想是:通过构造出的掩蔽图像与待处理图像进行点对点的乘积,游离出待处理

图像中感兴趣的部分。掩蔽图像是每个像素值仅取为 0 或 1 的二值图像,其构造方法是:将待处理图像中感兴趣区域所对应的掩蔽图像区域的像素值设为 1,其他部分的像素值设为 0。如图 1 所示,对一个人脸图像,如果感兴趣的是脸部,则可以通过脸部区域对应的像素值为 1、其他区域为 0 的掩蔽图像与该人脸图像的点对点乘积获得。其中,“ $\otimes$ ”定义为矩阵对应元素与元素的乘积。

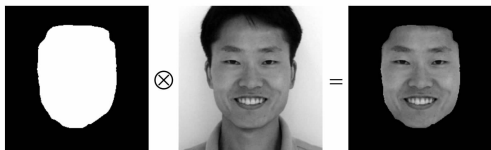


图 1 利用掩蔽图像实现脸部的游离

Fig. 1 Dissociation of Face by Mask Image

## 2.2 图像变形公式的修正

为解决传统方法的不足,需要将人脸表情图像的嘴部游离出来单独考虑。对于嘴部区域,本文限定其为上下嘴唇外轮廓所围的区域,并且为了方便计算和操作,将区域根据轮廓特征设定为凸多边形区域。

设源图像  $I_0$  和目标图像  $I_1$  的嘴部区域对应的掩蔽图像分别为  $M_0$  和  $M_1$ ,  $t$  时刻的变形图像  $I_t$  对应的掩蔽图像为  $M_t$ 。需要注意的是,  $M_t$  可以直接由  $M_0$  和  $M_1$  通过图像变形方法获得,而无需再通过对变形图像  $I_t$  嘴部区域的对应标定获得。另外,记  $E$  为每个元素均为 1 的矩阵,则式(1)等价于:

$$I_t = (1-t)[M_t \otimes I_t^{(0)} + (E - M_t) \otimes I_t^{(1)}] + t[M_t \otimes I_t^{(1)} + (E - M_t) \otimes I_t^{(1)}] \quad (2)$$

若记

$$\begin{aligned} I_{t,1} &= (1-t)M_t \otimes I_t^{(0)} + tM_t \otimes I_t^{(1)} \\ I_{t,2} &= (1-t)(E - M_t) \otimes I_t^{(0)} + t(E - M_t) \otimes I_t^{(1)} \end{aligned} \quad (3)$$

则式(2)可简记为:

$$I_t = I_{t,1} + I_{t,2}$$

即变形图像  $I_t$  分割为嘴部区域图像  $I_{t,1}$  和其他区域图像  $I_{t,2}$ 。

进一步地,假设源图像为闭口的表情图像,目标图像为开口露齿的表情图像,可知,其中间变形图像应该也是个开口露齿的表情图像,而开口和露齿的程度则取决于时刻  $t$ 。为解决中间变形图像出现重影现象,可将式(3)中  $t$  时刻源图像的扭曲图像  $I_t^{(0)}$  替换为  $t$  时刻目标图像的扭曲图像  $I_t^{(1)}$ ,即修正式(3)为:

$$I'_{t,1} = (1-t)M_t \otimes I_t^{(1)} + tM_t \otimes I_t^{(1)} = M_t \otimes I_t^{(1)}$$

从而实现了式(1)的修正:

$$I_t = I'_{t,1} + I_{t,2} \quad (4)$$

即

$$I_t = M_t \otimes I_t^{(1)} + (1-t)(E - M_t) \otimes I_t^{(0)} + t(E - M_t) \otimes I_t^{(1)} \quad (5)$$

## 2.3 掩蔽图像的变形

修正后的图像变形式(5)中比较关键的因素是  $t$  时刻变形图像  $I_t$  对应的掩蔽图像  $M_t$ 。由于  $M_t$  可以直接由  $M_0$  和  $M_1$  通过图像变形方法获得,而  $M_0$  和  $M_1$  是个简单的图像,其对应的区域是个多边形区域,因此,对  $M_0$  和  $M_1$  的变形可以采用简单的方法实现。

设掩蔽图像  $M_0$  和  $M_1$  对应的多边形的顶点分别为:

$$q_0^{(i)} = \begin{bmatrix} q_0^{(i1)} \\ q_0^{(i2)} \end{bmatrix}, q_1^{(i)} = \begin{bmatrix} q_1^{(i1)} \\ q_1^{(i2)} \end{bmatrix}, i = 1, 2, \dots, m$$

则  $t$  时刻掩蔽图像  $M_t$  对应的多边形的顶点可以构造为:

$$q_t^{(i)} = (1-t)q_0^{(i)} + tq_1^{(i)}, i = 1, 2, \dots, m$$

于是将获得的多边形所对应的区域内像素值填充为 1,其余填充为 0,即获得了掩蔽图像  $M_t$ 。

## 2.4 图像变形方法描述

根据前面的叙述和讨论,本文用于人脸表情合成的图像变形方法可描述如下:

- 1) 输入源图像  $I_0$  和目标图像  $I_1$ 。
- 2) 设置源图像和目标图像的特征点,即  $p_0^{(i)}$  和  $p_1^{(i)}, i = 1, 2, \dots, n$ 。

3) 产生  $t$  时刻源图像和目标图像的扭曲图像,即  $I_t^{(0)}$  和  $I_t^{(1)}$ : ① 计算  $t$  时刻的特征点坐标  $p_t^{(i)}, i = 1, 2, \dots, n$ ; ② 根据  $t$  时刻的特征点  $p_t^{(i)}$  以及源图像和目标图像的特征点  $p_0^{(i)}$  和  $p_1^{(i)} (i = 1, 2, \dots, n)$ ,利用对像素坐标的二维空间插值方法和对像素值的重采样,分别获得源图像和目标图像在  $t$  时刻的扭曲图像  $I_t^{(0)}$  和  $I_t^{(1)}$ 。

4) 设定和构造源图像和目标图像的嘴部区域对应的掩蔽图像,即  $M_0$  和  $M_1$ 。

5) 对掩蔽图像  $M_0$  和  $M_1$  进行变形,获得  $t$  时刻变形图像嘴部区域对应的掩蔽图像  $M_t$ 。

6) 根据修正后的图像变形式(5)确定最终获得的  $t$  时刻变形图像  $I_t$ 。

## 3 实验结果

分别采用当前常用的基于径向基函数的传统图像变形方法<sup>[5]</sup>和本文的图像变形方法对嘴部图像实现变形。图 2 给出了对比实验结果,其中,图

2(f)和 2(j)分别为源图像和目标图像,图 2(a)和 2(e)分别为源图像和目标图像的相应特征点分布;图 2(b)~2(d)为利用传统方法得到的各个时刻的变形图像,图 2(g)~2(i)为利用本文方法得到的各个时刻的变形图像。从图中可以看到,利用传统方

法得到的几个中间变形图像有明显的重影,并且图像变形过程中产生的图像很不符合一个人从闭口到开口露齿的自然变化过程;而利用本文方法实现的图像变形过程比较符合客观实际,并且产生的中间变形图像自然、逼真。

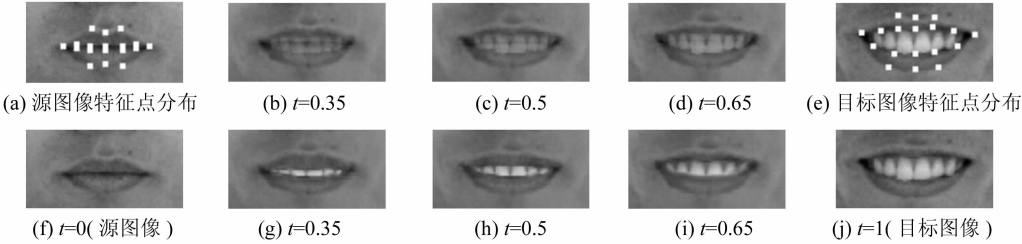


图 2 利用传统图像变形方法和本文图像变形方法得到的嘴部图像变形的对比实验结果

Fig. 2 Experimental Results of Mouth Morphing Images Obtained by Traditional Morphing Method and this Method

下面分别采用传统图像变形方法<sup>[5]</sup>和本文图像变形方法对整个人脸表情图像进行表情合成的对比实验。图 3 给出了实验所用的源图像和目标图像及其特征点分布,其对比实验结果如图 4 所示。图 4(a)~4(e)为利用传统方法得到的各个时刻合成的人脸表情图像,图 4(f)~4(j)为利用本文方法得到的各个时刻合成的人脸表情图像。从图中可以看到,利用传统方法得到的中间表情图像在  $t=0.2$  时与图 3(a)的源图像相比,看上去没有大的变化;在  $t=0.8$  时与图 3(b)的目标图像相比,看上去也没有大的变化,其他时刻的图像则在嘴部存在明显的不足。而利用本文方法实

现的由闭口的中性表情到开口露齿的高兴表情的这一变化过程客观、自然、流畅,变化过程中每一时刻合成的表情图像自然、逼真,具有说服力。

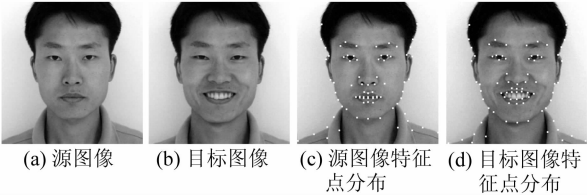


图 3 源图像和目标图像及其特征点分布

Fig. 3 Source Image, Destination Image and Their Feature Points

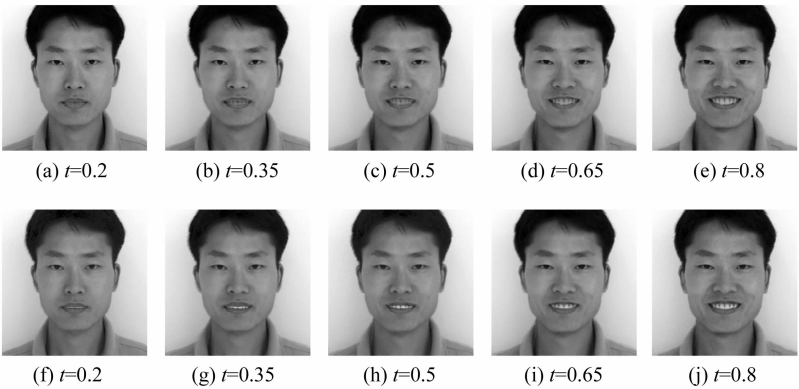


图 4 传统方法与本文方法应用于整个人脸表情图像的对比实验结果

Fig. 4 Experimental Results Obtained by Traditional Method and this Method with Whole Facial Expression Images

参 考 文 献

[1] Wolberg G. Digital Image Warping[M]. Los Alamitos: IEEE Computer Society Press, 1990

[2] 张静波,张桂林. 基于特征网格的面部图像变形算

法[J]. 中国图像图形学报 A, 2002, 7(3):256-260

[3] 朱桂斌,张邦礼,吴乐华,等. 基于 Delaunay 三角剖分的图像变形技术研究[J]. 中国图像图形学报 A, 2003, 8(6): 641-646

[4] Beier T, Neely S. Feature-Based Image Metamor-

phosis[J]. Computer Graphics, 1992, 26(2): 35-42

[5] Reisfeld D, Arad N, Dyn N, et al. Image Warping by Radial Basis Functions; Application to Facial Expressions[J]. CVGIP: Graphical Models and Image Processing, 1994, 56(2): 161-172

[6] Arad N, Reisfeld D. Image Warping Using Few Anchor Points and Radial Functions[J]. Computer Graphics Forum, 1995, 14(1): 35-46

[7] 翁妙凤,吴翰声,盛蔚. 人像变换的图像处理方法研究[J]. 计算机研究与发展, 1996, 33(9): 677-682

[8] 李旭东,张振跃. 非对称径向基函数与稳定边界图像变形算法[J]. 计算机辅助设计与图形学学报, 2004, 16(6): 747-752

[9] Lee S, Chwa K Y, Hahn J, et al. Image Morphing Using Deformable Surfaces[C]. Computer Animation'94, Geneva, Switzerland, 1994

[10] Litwinowicz P, Williams L. Animating Images with Drawings[M]. New York: ACM Press, 1994

[11] Lee S, Chwa K Y, Hahn J, et al. Image Morphing Using Deformation Techniques[J]. Journal of Visualization and Computer Animation, 1996, 7(1): 3-23

[12] Pighin F, Hecker J, Lischinski D, et al. Synthesizing Realistic Facial Expressions from Photographs [C]. The 25th Annual Conference on Computer Graphics and Interactive Techniques, New York, 1998

[13] Zhang Qingshan, Liu Zicheng, Guo Baining, et al. Geometry-driven Photorealistic Facial Expression Synthesis[C]. Eurographics/SIGGRAPH Symposium on Computer Animation, San Diego, CA, 2003

[14] 杜杨洲,林学阂. 基于形状外观关联映射的动态脸部纹理生成[J]. 软件学报, 2004, 15(2): 308-316

作者简介:李旭东,博士。现主要从事数字水印技术、人脸图像处理的研究。  
E-mail:leesheedong@163.com

# Image Morphing Method for Facial Expression Image and Animation Synthesis

LI Xudong<sup>1</sup>

(1 School of Mathematics and Statistics, Zhejiang University of Finance and Economics, 18 Xueyuan Street, Hangzhou 310018, China)

**Abstract:** Facial expression synthesizing method based on traditional image morphing will not take effect when synthesizing a facial expression image from an expression image with mouth closed and an expression image with mouth opened and teeth showed. In order to overcome this shortage, a novel image morphing method is proposed for image and animation synthesis of facial expression. In this method, the concept of mask image is brought in and the morphing method of mask image is presented, and then the formula of traditional image morphing is revised. The idea of this method is easy to understand and the method steps are simple. This method can effectively produce morphed images. Experimental results show that this method can naturally and livelily synthesize facial expression images from an expression image with mouth closed and an expression image with mouth opened and teeth showed. At this point, this method outperforms other existing facial expression synthesis methods based on traditional image morphing.

**Key words:** image morphing; facial expression synthesis; virtual reality; computer animation