

# 基于方法的表情生成技术综述

齐敏<sup>1)</sup> 李大健<sup>2)</sup> 王毅<sup>1)</sup> 樊养余<sup>1)</sup> 郝重阳<sup>1)</sup>

<sup>1)</sup>(西北工业大学电子信息学院,西安 710072) <sup>2)</sup>(西北工业大学365研究所,西安 710072)

**摘要** 表情生成技术是智能人机接口领域的重要研究内容之一。基于方法的表情生成技术多数以图像为研究对象,不需要3维建模,同时又具有丰富的纹理细节,在表情生成的逼真度和降低复杂性方面具有相当的优势,一些方法也可扩展至3维应用。本文根据各种算法的技术本质将已公开发表的研究总结归纳为渐变技术、表情映射技术、统计学方法、2维网格法、面色表情和表演驱动技术等几类,从概念、理论和技术方法等方面对典型成果进行讨论,分析了不同算法的特点和存在的问题以及今后的发展方向,为开展相关研究提供参考。

**关键词** 表情生成 渐变 表情映射 表情比例图 面色表情 表演驱动

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2009)11-2181-10

## A Survey of Facial Expression Generation Techniques Based on the Study Method

QI Min<sup>1)</sup>, LI Da-jian<sup>2)</sup>, WANG Yi<sup>1)</sup>, FAN Yang-yu<sup>1)</sup>, HAO Chong-yang<sup>1)</sup>

<sup>1)</sup>(College of Electronic Information, Northwestern Polytechnical University, Xi'an 710072)

<sup>2)</sup>(Institute 365, Northwestern Polytechnical University, Xi'an 710072)

**Abstract** The expression generation is one of important research contents of intelligent human-computer interface. Because images are the research objects in the majority, the facial expression generation techniques based on the study method don't need to model for 3D faces and potentially can be of highly realistic expressions. These are of considerable advantages in reality improving and complexity reducing. Some methods can also be used to 3D models. The paper summarizes the published researches from the aspects of morphing, expression mapping, statistical method, 2D meshes, facial color expression and performance-driven. The conceptions, principals and the methods of typical achievements are discussed. The features and existing problems of different algorithms as well as the development trend are analyzed. It provides some useful information and gives a reference to related works.

**Keywords** expression generation, morphing, expression mapping, expression ratio image, facial color expression, performance-driven

## 1 引言

人脸是人类最重要的生物特征之一,包含着极其丰富的信息,以人脸为主体的识别、跟踪、表情动画等工作是近年来的研究热点,其中表情生成是智能人机接口领域的重要研究内容之一。客观上,廉

价桌面系统的处理能力飞速提高;主观上,人们对于具有全新沉浸感的人机交互方式的需求也越来越强烈,这两方面的因素成为表情生成技术发展的重要推动力。表情生成技术在智能人机接口、虚拟现实、计算机辅助训练、远程教育和远程会议等诸多方面有着广泛的应用前景。

由于表情的丰富性,以及人脸的复杂性和易变

基金项目:国家高技术研究发展计划(863)项目(2007AA01Z324);西北工业大学行动计划项目(52103020301-0800-07XE0126)

收稿日期:2008-08-25;改回日期:2009-08-03

第一作者简介:齐敏(1968~),女,副教授,硕士生导师。2000年于西北工业大学获系统工程专业博士学位。主要研究方向为图形图像工程、虚拟现实等。E-mail: drqimin@nwpu.edu.cn

性,具有真实感的人脸表情生成技术一直是一个富有挑战性的研究课题。人脸表情是人类对自身脸部行为的一种主观定义。在人脸能够产生的 5 万多种不同的表情中,仅有 30 多种有代表性的表情可以通过人类语言进行定义和描述,大部分表情是无法用语言表达的。同时在另一方面,人脸不同生理层之间以及同一生理层的不同区域之间都存在相互作用,模拟人脸表情的复杂和困难程度从此可窥一斑。

根据研究的需要,人们通常将人脸表情分为 6 个基本类型:中性(neutral)、喜悦(happiness)、悲哀(sadness)、愤怒(anger)、恐惧(fear)和惊讶(surprise),其典型特征如图 1 所示<sup>[1]</sup>。许多研究中也将会表情分为 7 个基本类型,此时在上述表情中增加一个厌恶(disgust)表情。其余表情通过不同方法由这些基本表情合成。

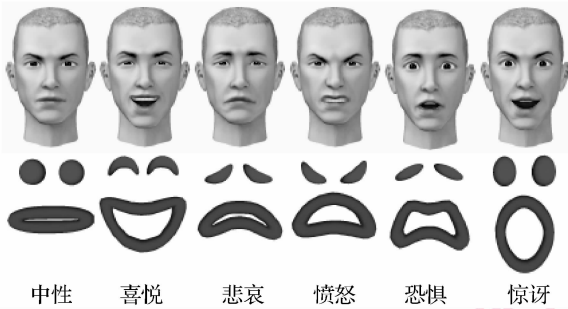


图 1 6 种基本表情及其典型特征

Fig. 1 The six basic facial expressions and their typical features

## 2 表情生成的生理学依据

人脸模型是表情生成的基础。从解剖学角度分析,包括额骨、颧骨和上颌骨及下颌骨等在内的颅骨决定着人脸的整体外观轮廓,其中颧骨和下颌骨起到决定作用。肌肉等软组织依附在颅骨之上,也会影响脸型,其各种变动产生人脸的相应表情和动作,也就是说,颅骨和肌肉组织的形态基本上决定了人脸的造型特征,两者的结构如图 2 所示。

因此,通常人脸模型都是建立在头部颅骨与肌肉组织的结构特点之上,其参数和特征点的定义离不开生理解剖学基础。从外表来看,人脸部的眼睛、嘴巴、下巴以及人脸外轮廓等部分都含有一人区别于其他人的许多重要信息,是产生表情时面部变形的重要依据。

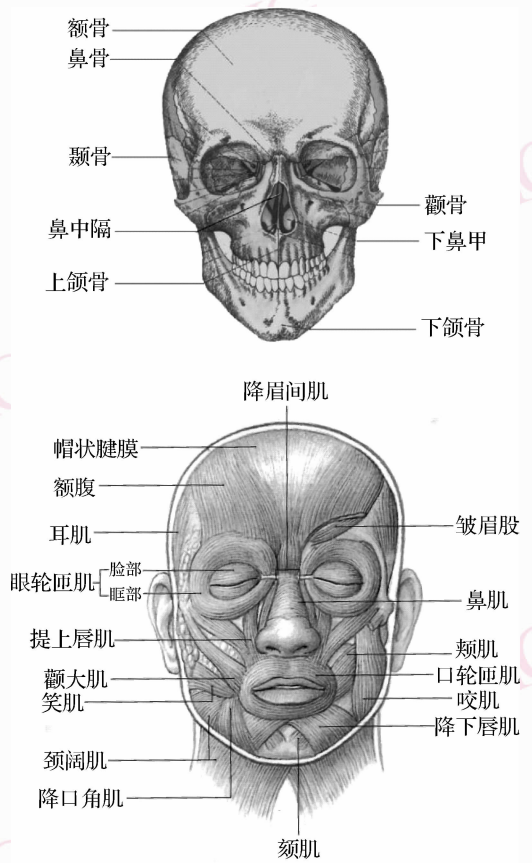


图 2 颅骨与肌肉组织结构

Fig. 2 The structures of the skull and muscle

## 3 表情生成技术分类

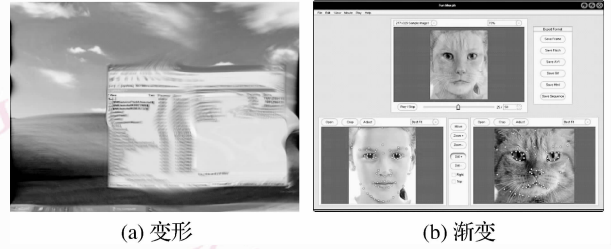
表情生成算法研究的对象有 2 维脸部模型和 3 维头部模型两大类,但是许多方法对 2 维和 3 维模型的应用并没有严格的界限。较多的研究以某种 3 维人脸头部模型为基础,如弹性模型、参数模型、肌肉模型、伪肌肉模型以及脸部动作编码系统(FACS)和 MPEG-4 的动画模型等,表情的生成依赖于模型的表示方法,通常两者是结合在一起的。这一类研究可以有自由的视点,某些方法可以获得精确的模型数据和面部表情的变形,但一般手工干预比较多,在表情生成的逼真度、建模和算法的复杂性及鲁棒性等方面存在较大阻碍。

另一类研究与 2 维图像的关系更密切一些,但一些技术也涉及 3 维方面,可以对它们按技术方法进行分类,本文称之为基于方法的表情生成技术。由于这类研究绝大多数以图像为主,不需要重构 3 维模型,避免了建模误差,同时又天然继承了图像细

节丰富的特点,在表情生成的逼真度和降低复杂性方面具有相当的优势,这也是人们致力于这一研究方向的重要原因。

通过对公开发表的文献资料中各种方法进行分析、比较和归纳,根据算法的实现原理将基于方法的表情生成技术进一步划分为 6 类,包括渐变技术、表情映射技术、统计学方法、2 维网格法、面色表情和表演驱动技术。表情生成技术的分类如图 3 所示。下面就基于方法的表情生成技术分别进行详细讨论。

所示。



(a) 变形 (b) 渐变

图 4 变形与渐变

Fig. 4 Warp and morph

生成表情的基本渐变技术的原理是定义一个在单位时间区间上的形变函数,通过对同一对象 2 种不同的表情图像进行帧间插值,计算生成中间状态特征点的 2 维位置坐标,从而产生 2 个指定的脸部表情图像之间的光滑过渡,即在 2 个已有的表情之间生成新的表情。基本渐变的缺点是:要求确定 2 个图像特征点之间点对点的对应关系;其次,当视点和姿势发生变化时,会产生不真实的脸部表情图像。

基于视点的渐变克服了图像变化对视点和头部姿势的敏感性,但图像中目标对象可视度的变化对变形结果有一定的影响,因此出现 3 维渐变。3 维渐变是 2 维图像渐变和 3 维几何模型变形相结合的产物,需要计算 3 维位置坐标和纹理空间坐标值,可以附加物体的物理特性描述。其实质是用 3 维插值实现脸部表情之间的形状变化,用 2 维渐变实现对纹理图像的变化。插值分为两种:几何插值和参数插值。几何插值直接更新人脸网格上顶点的位置;参数插值对控制表情的函数参数进行插值,从而达到间接控制顶点移动的目的<sup>[7]</sup>。3 维渐变获得了独立于视点的真实性,但是动画还是受预先定义的关键表情之间插补的限制。

此外,还有将 3 维模型映射到 2 维参数空间,通过变形其 2 维参数空间,达到实现 3 维模型变形的方<sup>[8]</sup>。

渐变的关键是帧间插值。线性插值由于简单而被广泛使用;复杂一些的情况可以使用余弦或类似的插值函数,使表情动画在开始和结束时有加速或减速的效果;使用双线性插值等其他技术能够表达更加丰富的表情<sup>[9-12]</sup>。

插值方法要求必须有同一个人的两幅表情图像或表情模型,当实际中只有一幅表情图像或模型时,单纯的表情渐变方法就无能为力了。

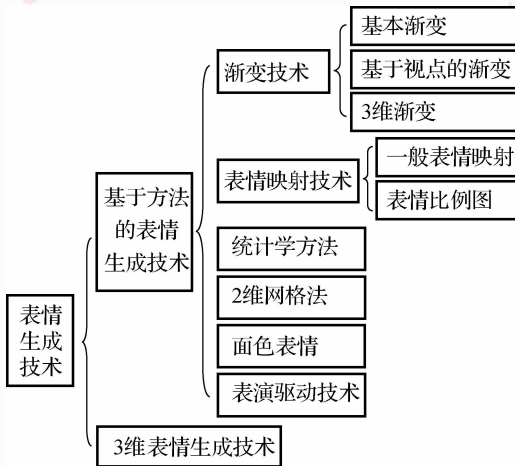


图 3 表情生成技术分类

Fig. 3 The classification of the facial expression generation techniques

## 4 基于方法的表情生成技术

### 4.1 渐变技术

将广泛使用的渐变技术用于生成人脸表情是一种直观的方法,按其特点可分为基本渐变 (morphing)<sup>[2]</sup>、基于视点的渐变 (view morphing)<sup>[3]</sup> 和 3 维渐变 (3D morphing)<sup>[4-6]</sup> 3 类。

基本渐变指普通的 2 维渐变,包括源图像和目标图像对应点的变形 (warp) 和淡入淡出。有时将 warp 和 morph 均称为“变形”,实际上混淆了两者的概念。warp 是指对一幅图像进行某种运算,使图像本身发生变化,而 morph 是指从一幅图像到另一幅图像逐渐变化的过渡过程,本文用“变形”和“渐变”两个词加以区别。两幅图像间的 morph 方法是首先分别按照特征结构对两幅原图像做 warp 操作,然后从不同的方向渐隐渐显地得到两个图像序列,最后合成得到 morph 的结果。变形与渐变的区别如图 4

## 4.2 表情映射技术

表情映射是一种将某个人脸对象的表情重新定位到其他特定人脸上的方法,广泛应用于表演驱动的脸部动画中。可分为一般表情映射和表情比例图两类。

### 4.2.1 一般表情映射

一般表情映射(expression mapping)的基本思路是:给定某人的中性脸(neutral face)和表情脸(expression face)图像,确定两幅图像中的特征点,然后计算这两组特征点的差向量,并将它作用到另一个人中性脸的特征点上,使该中性脸依此进行图像变形(image warping),从而得到新的表情<sup>[13]</sup>。其实质是利用已经存在的顶点运动向量等数据,将其他人脸对象的表情映射到或者说定位到新的特定人脸上。

这类方法借助两幅参考图像的帮助实现了任意对象新表情的生成,弥补了单纯的表情渐变方法的缺陷。但它也有个明显的缺点,即整个过程仅针对人脸表情进行,没有考虑皮肤变形挤压产生的皱纹等变化丰富的表情细节,因而影响了表情的真实感程度。

改进的方法<sup>[14]</sup>在获得特征点差向量的基础上,根据少量已知的表情样本图像,推测出一些决定表情皱纹的细节点的移动,生成皱纹,以此提高逼真度。其中表情样本图像事先离线获得,每人 10 至 15 幅。特征点可以通过自动的方法计算出来,数量不多的情况下也可以直接手工操作确定。图 5(a)是这种方法所选取的 134 个特征点,左下角是嘴张开时的情况。由于表情皱褶位于脸部的不同区域,因此根据这些特征点将脸部划分为 14 个子区域进行计算,如图 5(b)所示,最后再进行无缝拼接,得到

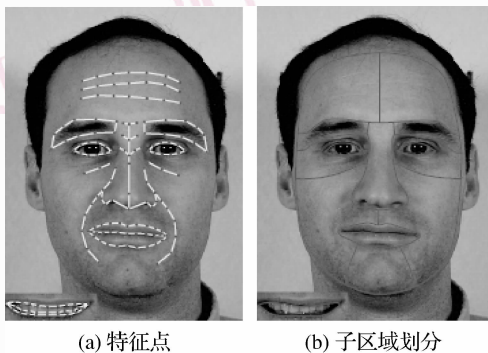


图 5 特征点和脸部子区域划分

Fig. 5 Feature points and face region subdivision

最终的表情图像。图 6 是将女性表情映射到男性脸上的效果。这种方法可以扩展应用到 3 维情况。

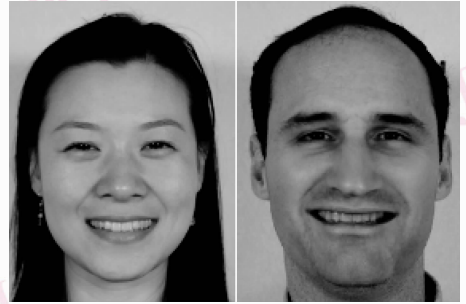


图 6 将女性表情映射到男性脸上

Fig. 6 A expressions of the female is mapped to the male

2001 年 Noh<sup>[15]</sup>首次提出了表情克隆(expression cloning)一词,Noh 的表情克隆在原理上仍属于表情映射,同时考虑了表情变换过程中脸部特征点的位置变化和脸部细节的改变,可以产生具有真实感的人脸表情动作。图 7 是 Noh 的表情克隆效果。针对 2 维图像和 3 维模型的克隆技术研究是近年的一个研究热点<sup>[16-19]</sup>。

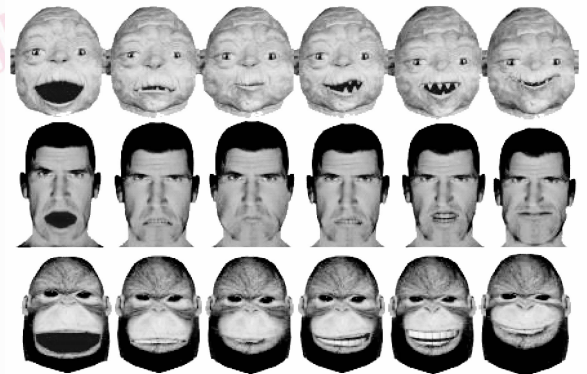


图 7 Noh 的表情克隆

Fig. 7 Noh's expression cloning

### 4.2.2 表情比例图

面部表情的体现不仅是特征点位置的移动,还有皮肤表面起伏引起的明暗效果及皱褶的变化,这些细节是非常重要的视觉信息,但是由于非常细小,合成起来很困难,其效果的体现很大程度上依赖于光照的作用。基于这样的认识,Liu 等人 2001 年在一般表情映射的基础上提出了“表情比例图”(ERI)的概念<sup>[20]</sup>,深入研究了由光照导致的细节变化。他们将人脸表面由法向改变而引起的光照变化用一种与表面反射率无关的方式提取出来,记录在表情比

例图里,在表情比例图中捕捉表情变化引起的光照变化信息,结合人脸图像的稀疏特征表达,获取表情变化前后的形状变形与纹理变化细节信息,并将这种信息映射到其他人脸图像中,实现逼真表情变换。

基本思路是给定某人的中性脸和表情脸图像,计算两幅图像各对应像素光亮度之比或  $R, G, B$  3 个成分的比值,然后结合表情变化前后特征点的移动,将这组比例作用到另一个人的中性脸图像上,进行变形操作,从而获得另一个人的表情图像。图 8 是 Liu 的表情比例图,其中左图是中性表情,中间的图是蹙眉表情,右图是蹙眉表情的 ERI。



图 8 Liu 的表情比例图

Fig. 8 Liu's expression ratio image

ERI 本质上是基于视点的光照变化信息提取,它捕捉了人脸表面的深度信息,可以反映面部几何变形引起法向量改变而导致的光影变化。将其作用于不同的中性脸图像,可以产生逼真的表情效果,还可用于不同光照的细微表情合成。

这种方法也可以扩展应用到有纹理的 3 维模型,此时只需将 ERIs 应用于 3 维模型所用纹理贴图即可。同时也可以应用于无纹理,仅有明暗效果的 3 维模型,包括人工合成模型之间的表情映射和真实人脸模型到人工合成模型的表情映射。

ERI 可以描述细节褶皱,但是维数太高。改进的方法用主分量分析(PCA)法将高维空间的 ERI 压缩成低维空间的特征 ERI(EigenERI),以更便于可视化和特征化的 EigenERI 充当褶皱量化参数,最后用 PCA 将 EigenERI 恢复成 ERI,进行表情合成<sup>[21]</sup>。

ERI 方法解决了一般表情映射无法合成表情细节的缺陷,但仍然存在问题:要获得某个人某种特定表情脸,必须有一幅已知的表情脸图像作为一个样本与之对应。但在表演驱动动画中,人脸表情的变化是极其丰富的,因此需要有大量的样本来供选

择,这显然是很困难的。

针对这一问题,出现了参数化表情映射(parameterized expression mapping)技术<sup>[22]</sup>。这种方法利用图像的小波金字塔分解技术,对 ERI 算法进行拓展,通过引入参数,由一幅中性脸图像和几幅最大程度表情脸图像生成一系列带有细节的中间表情图像,并且表情夸张程度可通过参数控制。若将表情图像作为纹理映射到 3 维人脸模型上,就可以拓展到 3 维人脸动画上。

#### 4.3 统计学方法

统计学方法以建立人脸 2 维图像样本数据库为基础,基本思想是利用样本库中的人脸图像,以线性组合或其他组合方式表示新的人脸。这种表示建立在一个人脸空间上,该人脸空间中的基底可以直接由训练集内的样本图像表示,也可以用主分量分析或独立分量分析(ICA)得到的抽象基底表示。通过总结人脸对象的一般规律,对特定人脸图像进行模型匹配与表达,就可以结合不同属性特征的人脸图像数据库实现不同的脸部图像处理效果<sup>[23-27]</sup>。

基于稠密特征对应的人脸表情生成算法<sup>[24]</sup>建立形状与纹理分离的人脸线性统计模型,以形状矢量和纹理矢量的矢量化方式表示人脸图像。矢量化表达是指将特征点集(如图像的像素点)依次排列成一个有序的矢量,通过相对于一个确定的参考图像完成特征对准,从而使每个矢量的相应维都代表同一个特征点。不同人像间的形状和纹理都存在差异性,形状的差异使每个像素点存在一个 $(\Delta x, \Delta y)$ 的位置偏移量,纹理矢量描述的则是消除了人脸形状差异后的纹理信息,即经过对准后的特定人脸纹理。

生成表情时,同时考虑脸部的形状变化和相应的纹理变化,充分利用人脸表示的矢量化机制,借助于 Liu 等人提出的表情比例图产生皱纹、阴影、遮挡等细节变化。该方法比 ERI 方法的变形操作次数少,不需要对表情比例图滤波,在线处理的数据量大降低,提高了实时处理的速度。

形状外观关联映射(shape-appearance dependence mapping, SADM)<sup>[25]</sup>方法又称形状纹理关联映射方法<sup>[26]</sup>。该方法通过对脸部表情视频序列的统计分析,建立人脸参数模型,根据关联思想建立映射,由脸部形状特征点的运动得到动态纹理,从而产生表情。算法中的“外观”是指脸部外观,与动态纹理紧密联系。动态纹理定义为表情变化时,面部状态短

时变化引起的特征纹理,如额头皱纹、酒窝、颧骨阴影、嘴部的牙齿舌头以及眼部区域的细微变化等,是相对于痣、雀斑以及人脸肌肤衰老产生的皱纹等不随面部运动变化的静态纹理而言。

算法中的人脸参数模型使用已被较多采用的 Cootes 等人<sup>[27]</sup>提出的用主分量分析法建立的人脸图像参数化统计模型。该模型将人脸外观分为形状模型和灰度模型两个独立的部分。形状模型表示人脸上主要特征点的相对位置与空间关系,描述了平均形状和各种形状变化模式。灰度模型反映了灰度纹理信息,主要是图像整体的纹理模型,有时为了描述沿人脸轮廓和重要特征边缘的灰度变化,还需建立局部的轮廓灰度模型。形状模型和灰度纹理模型共同描述了人脸图像的总外观。

关联映射方法认为在图像中人脸特征的形状与其纹理有一定的依赖关系,基于这样的思想,找到从形状到纹理的关联关系,并以映射的数学形式表达出来,从而可以根据形状参数自动地生成真实自然的脸部动态纹理,相应的表情自然也就出现了。这一方法也可以集成到基于线框模型的头脸模型中产生真实的动画效果。

利用情感参数控制表情图像生成的方法<sup>[26]</sup>借助人脸表情图像数据库,训练出从情感状态到表情图像的映射,也就是从情感参数到表情图像参数的映射——情感函数,采用二次多项式形式拟合得到。情感函数以任意人的中性脸为输入,根据情感参数的控制,为这个人生成相应的表情图像。

情感函数描述了脸部表情随内心情感变化的方式,分为通用情感函数和专用情感函数。通用情感函数用所有人数据的混合集训练得到,可以用一种共同的风格为任意人生成表情图像。专用情感函数用某个特定人的数据集训练得到,有利于强调表现某个人以及和他长相相似的人的个性化表情效果。

这种方法采用的表情图像参数化统计模型是相对参数模型。在建立模型时采用相对量训练模型,以每个人的中性脸为参考点,将表情脸与中性脸做减法得到形状相对量,做除法得到纹理相对量,这样能够以独立于特定人的方式提取出表情变化量。然后对这些相对量应用主分量分析降维,并得到参数化表示。这样,即使待合成人脸的肤色、相貌距训练图像集很远时,其表情变化也能被参数模型很好地表示出来。图 9(a)是一张中性脸和一张表情脸,图 9(b)左边是形状差异,右边是纹理差异。

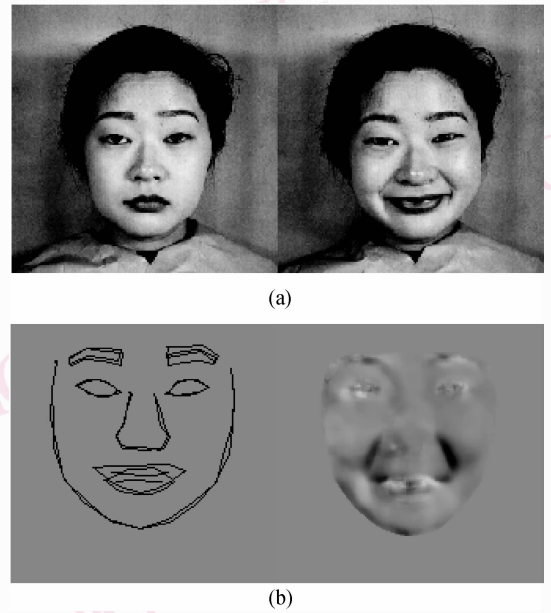


图 9 表情脸的形状变化量和纹理变化量  
Fig. 9 The shape variation and texture variation of the face with expression

这种方法可以用任意人的一张中性脸图像为其合成情感状态可控的表情图像。图 10 是用情感函数生成的表情示例,第 1 行是一个人的典型情感样本图像,第 2 行和第 3 行是利用第 1 行人各张图片所对应的情感参数和其专用情感函数为蒙娜丽莎和

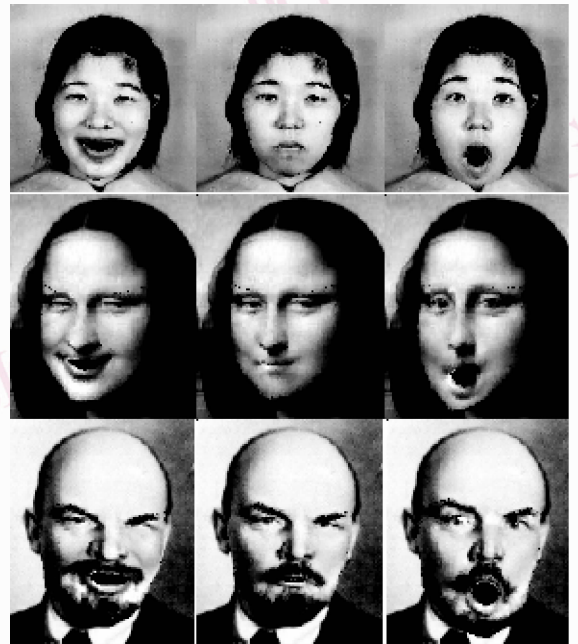


图 10 为蒙娜丽莎和列宁合成的表情图像  
Fig. 10 The synthetic expressions of Mona Lisa and Lenin

列宁做合成的表情效果,可以看到第 1 行人的表情风格映射到了第 2 行和第 3 行人的脸上。同时注意到,当图像中的人脸姿态与样本图像差别较大时,生成表情的逼真度下降。

#### 4.4 2 维网格法

2 维网格法同样考虑了人脸图像的两个主要特征:形状和纹理,形状包括脸部轮廓和五官位置等外观信息,纹理则表示基于该形状特征的图像灰度信息。

人脸的 2 维三角形网格模型是根据头的形状以及脸的特征结构抽象出的 2 维结构,脸部的表情运动由三角形的位置及大小的变化来合成,表现为三角形顶点的改变,它来源于表情分析中的参数。纹理的变化主要体现在图像某些区域中像素灰度的改变。因此表情合成的任务是根据表情参数修改 2 维模型并应用纹理映射技术重新分配图像中变化部分的像素灰度的处理过程<sup>[28]</sup>。

将脸部不同的表情变化用 2 维几何网格描述后,就可以使用 PCA 方法将网格序列用参数化模型表示,极大地压缩了数据量。在重构表情图像时,利用少量参考帧提供的场景纹理信息,结合网格参数模型提供的几何信息,使用图像变形与渐变技术,即可生成特定人脸的逼真表情图像。这种方法的目的同样是为了避免使用 3 维模型的困难,同时将 2 维模型中的脸部信息与背景信息、几何信息与纹理信息明确地分开,用经济有效的方式对不同的信息进行组织与操作,有利于由简单表情组合成复杂表情,适应不同光照条件产生的不同效果以及应用到具有不同背景の場合<sup>[29]</sup>。

此外,在如图 11 所示的正面和半侧面人脸图像上建立 2 维网格形状模型,可以借助于 MPEG4 3 维模型的形变参数以及图像插值与变形技术,合理地解决遮挡问题,实现一定程度上姿态可变的人脸动画和表情<sup>[30]</sup>。

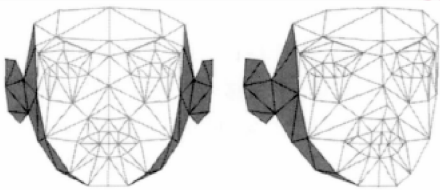


图 11 2 维脸部网格

Fig. 11 Two dimensional face meshes

#### 4.5 面色表情

面色表情是指脸色变化所体现出的表情。真实的人脸表情不仅仅包含人脸的变形和皱褶等纹理的变化,还包括因情绪变化而引起的面色变化,这方面的研究还不是很多<sup>[31-32]</sup>。较早的一个面色变化表情模型由 Kalra 等人提出,其中给出了一个情绪计算模型,用以模拟脸色变化,定义了最小可察觉颜色运动(minimum perceptible color action, MPCA),用来控制脸部不同部分的血液循环,进而产生面部颜色的变化。该方法在纹理图上定义了一些 Bezier 平面片形状的 MPCA 区域,通过像素估值计算每个像素的参数变化,该像素参数可以改变纹理图像的颜色属性。用这种技术可以再现脸色苍白和面红耳赤等。

Yamada 等人对大笑、恐惧和生气情绪时的面色变化进行了深入研究<sup>[33-35]</sup>,生成了上述 3 种情绪时反映面部表情的颜色变化。

首先,拍摄每种情绪变化时的脸部颜色,在 HSV(hue-saturation-value)色彩空间对其进行分析;测量鼻子和前额的温度,这两个地方温度变化显著,用以代表脸部皮肤温度,温度反映了情绪变化过程中血管血流量的变化;用 68 个不考虑眼睛和嘴巴区域的女性脸图像合成消除了个性差异的平均面色图像。

然后,分析脸色变化与脸部皮肤温度的关系,将脸部皮肤温度作为估计情绪,进而生成脸部表情颜色的指标。在不同情绪时,对平均面色图像在色度和饱和度上增加不同的修正量,得到增强的脸部颜色图像(enhanced facial color images),当色度和饱和度其中之一是主要影响因素时,则只考虑一个修正量,由此得到动态的面色变化。图 12 左边是平均面色图像,右边是生气时的面色图像。

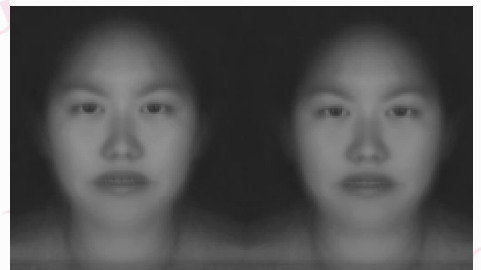


图 12 平均脸色图像(左)和生气时的脸色图像(右)

Fig. 12 An average facial color image(left) and the facial color under the emotional change of anger(right)

上面两种方法都是以血管血流量作为面色变化的生理学依据,用于模拟不同表情时面部的真实颜色。文献[36]研究了游戏中人物角色的脸色表现问题,开发了脸色引擎。脸色引擎是一个结合了人文学理论、情绪-颜色关系理论和情绪反应速度函数的情绪模型,与用血流量、皮肤温度这类表现肤色的方法完全不同。这种方法对生成真实自然的表情颜色并不适用,但是其中对不同情绪时面部色彩运用的研究具有一定的参考价值。

#### 4.6 表演驱动技术

表演驱动(performance-driven)是一种流行的脸部动画方法,最初是为了创造卡通形象。一般需要在真人脸上放置一些特殊的标识或传感器,其中的运动捕捉系统实时地检测、记录表演者脸部在3维空间中的运动轨迹,从而精确地跟踪脸部特征。用获得的人脸关键特征点的位移和头部姿势等表情参数控制表情的生成,例如将它们直接用于脸部动作编码系统(FACS)或MPEG-4的动画单元,或者利用径向基函数(RBF)、表情映射等方法产生脸部表情<sup>[14, 37-39]</sup>。

这种方法的数据逼真度高、细节丰富,已经成为高质量数据的主要来源,在动画制作中被广泛采用,但代价昂贵。存在容易造成动作变形,使用不够方便,整个过程中需要相当多的人工干预和枯燥调整的问题。研究减少表情生成误差<sup>[40]</sup>和人工干预的工作量<sup>[41]</sup>,同时保证逼真的动画效果是目前尚在努力的方向。

鉴于运动捕获的本质是记录运动数据再重放,而视频是记录运动的主要载体,因此重用已有的视频数据,以视频中的表演作为生成表情的驱动,成为一种思路,这样可以大大降低成本,同时提高表情动画生成的实时性。由此,根据真实的人脸录像序列跟踪每一帧人脸表情的方法成为一个新的研究方向。

最初的研究使用静态脸部纹理图像捕捉2维特征点的运动<sup>[42]</sup>,或从多张不同角度的脸部图像中捕捉几何模型的变化和纹理<sup>[43]</sup>,从而实现脸部表情动画。Snake动态轮廓线模型(ACM)常被用来获取脸部的特征变化<sup>[44]</sup>。

随后,基于多视频驱动和单视频驱动的研究相继出现<sup>[45-48]</sup>。根据视频驱动人脸动画,采用光流技术(optical flow)、活动形态模型或称主动外观模型(AAM)等各种方法在视频图像中捕捉对象的几何

和纹理变化,从而获取表情特征的变化数据<sup>[49-51]</sup>,然后和3维运动数据间建立映射关系,这种关系一般来说很难一一对应,是一个研究的重点。这样就可以利用从视频图像中提取的人脸动画参数,将表情重新定位到新的人物模型上。此外,如何进行运动编辑,使最终结果过渡自然,保证动画片段和原有纹理均能无缝连接也是需要考虑解决的问题。

其他还有神经网络法、基因算法等在人脸建模和表情动画方面也有应用<sup>[52-54]</sup>。

## 5 结 语

从表情生成的生理学依据、技术方法分类、典型算法等方面对基于方法的表情生成技术的研究成果进行了比较系统的归纳介绍。通过前面的讨论可以得到以下结论:

比较渐变技术和表情映射技术可以看到,渐变技术要求必须有同一个人的两幅表情图像或表情模型,才能生成此人两个表情之间的一系列新表情,受限制较大。表情映射技术虽然也需要某个人的两幅图像作为参考,即一个中性脸图像和一个表情脸图像,但是当根据这两幅参考图像得到位移和纹理的变化信息后,就可以将这个表情映射到任意多的其他人脸上,只需要其他人的一张中性脸图像就能为其生成相当逼真的表情图像。其中,ERI方法被较为广泛地采用,作为进一步研究的基础。表情映射技术的限制在于,对不同的表情需要相应的不同表情脸参考图像进行重新计算,虽然这方面有改进算法,但没有实质的进步,效果也有待于提高。

“表情克隆”是一个出现频率较高的词,其使用有泛化的趋势,主要是反映算法具有将表情重新定位到不同对象上的功能。其技术方法多种多样,大部分算法都同时关注了表情变化时面部的形状改变和表情褶皱等细节问题,强调产生表情动作的真实感。

在统计学方法和2维网格法中,多数都是从形状和纹理两个方面去研究问题,基本出发点大致相同。对人脸图像的描述有的采用相对量,有的采用绝对量,比如基于稠密特征对应的人脸表情生成算法和情感函数方法采用的是相对的形状模型和纹理模型,形状纹理关联映射方法和2维网格法采用的是绝对量模型。

人脸表情生成技术从最初研究不同表情时脸部

特征的形状变化开始,逐步发展到形状和表情纹理同时被考虑,近期面色表情的研究也开始受到关注,人们对于面部皮肤细节在表现表情方面的认识不断加深,研究能力不断提高。并且表情生成研究与情感计算和实时性等问题相接结合,研究的深度和广度都在不断扩展。基于方法的表情生成技术目前存在的主要问题是:

具体生成表情的算法一般针对正面图像,当发生侧转等姿态变化时,效果明显下降,这也是这类研究因图像本身特点而受制约的一个瓶颈问题,是与 3 维表情生成技术相比弱势的地方。如何解决图像中姿态变化引起的遮挡问题,生成姿态可变的逼真表情,仍需深入研究。

另外,在对光照变化、表情皱纹、面部颜色等细节方面考虑的同时,与 MPEG-4 人脸动画标准结合,可以在改善逼真度的同时,提高数据的压缩率,减少数据量,有利于扩大各类方法的应用领域,是一个值得研究的方向。

### 参考文献 (References)

- Pyun H, Kim Y, Chae W, *et al.* An example-based approach for facial expression cloning [A]. In: Proceedings of SIGGRAPH Symposium on Computer Animation [C], San Diego, California, USA, 2003: 167-176.
- Li Xu-dong. Image morphing method for facial expression image and animation synthesis [J]. Geomatics and Information Science of Wuhan University, 2007, 32(9): 796-799. [李旭东. 用于人脸表情图像与动画合成的图像变形方法 [J]. 武汉大学学报(信息科学版), 2007, 32(9): 796-799.]
- Fu T, Foroosh H. Expression morphing from distant view points [A]. In: Proceedings of International Conference on Image Processing [C], Singapore, 2004, 5: 3519-3522.
- Zhang P Y, Badler Norman I. Synthesis of 3D faces using region-based morphing under intuitive control: research articles [J]. Computer Animation and Virtual Worlds, 2006, 17(3-4): 421-432.
- Lu T, Mu D, Tsai M. Facial expressions for 3D game applications [A]. In: Proceedings of the Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IHMSP 2007) [C], Kaohsiung, Taiwan, China, 2007, 1: 341-344.
- Pighin F, Hecker J, Lischinski D, *et al.* Synthesizing realistic facial expressions from photographs [A]. In: Proceedings of the 25th annual Conference on Computer Graphics and Interactive Techniques [C], New York, NY, USA: ACM Press, 1998: 75-84.
- Sera H, Morishima S, Terzopoulos D. Physics-based muscle model for mouth shape control [A]. In: Proceedings of IEEE International Workshop on Robot and Human Communication [C], Piscataway, NJ, USA, 1996: 207-212.
- Chen D T, State A, Banks D. Interactive shape metamorphosis [A]. In: Proceedings of the 1995 Symposium on Interactive 3D Graphics [C], Monterey, California, USA, 1995: 43-44.
- Arai K, Kurihara T, Anjyo K I. Bilinear interpolation for facial expression and metamorphosis in real-time animation [J]. The Visual Computer, 1996, 12(3): 105-116.
- Zhang Han-ling. Study on Key Technology of Image Based Rendering [D]. Xi'an: Northwestern Polytechnical University, 2003. [张汗灵. 基于图像的绘制关键技术研究 [D]. 西安: 西北工业大学, 2003.]
- Zanella V, Vargas H. Active shape models to automatic morphing of face images [A]. In: Proceedings of the 4th WSEAS International Conference on Signal Processing, Computational Geometry & Artificial Vision [C], Tenerife, Canary Islands, Spain, 2004: 1-5.
- Guo H, Fu X Y, Chen F, *et al.* As-rigid-as-possible shape deformation and interpolation [J]. Journal of Visual Communication and Image Representation, 2008, 19(4): 245-255.
- Litwinowicz P, Williams L. Animating images with drawings [A]. In: Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques [C], New York, NY, USA: ACM Press, 1994: 409-412.
- Zhang Q, Liu Z C, Guo B N, *et al.* Geometry-driven photorealistic facial expression synthesis [A]. In: Proceedings of the 2003 Eurographics Symposium on Computer Animation [C], San Diego, CA, USA, 2003: 177-186.
- Noh J Y, Neumann U. Expression cloning [A]. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques [C], Los Angeles, California, USA, 2001: 277-288.
- Kong D H, Kang L, Yin B C. Cloning of facial expressions using spatial information [A]. In: Proceedings of Computer Graphics, Imaging and Visualisation: New Advances [C], Bangkok, Thailand, 2007: 365-369.
- Park B, Chung H, Nishita T, *et al.* A feature-based approach to facial expression cloning [J]. Computer Animation and Virtual Worlds, 2005, 16(3-4): 291-303.
- Yin L, Basu A, Yourst M T. Active tracking and cloning of facial expressions using spatio-temporal information [A]. In: Proceedings of the International Conference on Tools with Artificial Intelligence [C], Washington, DC, USA, 2002: 347-354.
- Theobald B J, Matthews I A, Cohn J F, *et al.* Real-time expression cloning using appearance models [A]. In: Proceedings of the 9th International Conference on Multimodal Interfaces [C], Nagoya, Aichi, Japan, 2007: 134-139.
- Liu Zi-cheng, Shan Ying, Zhang Zheng-you. Expressive expression mapping with ratio images [A]. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques [C], Los Angeles, California, USA, 2001: 271-276.
- Xie P S. Realistic Facial Animation Based on Contour Change-Wrinkle Synthesis [D]. Nanjing: Nanjing University of Information Science & Technology, 2006. [谢培树. 基于形变褶皱合成的逼真人脸动画 [D]. 南京: 南京信息工程大学, 2006.]
- Wu Y, Zheng W T. An algorithm for parameterized expression mapping [J]. Application Research of Computers, 2004, 2004(10): 117-121. [吴渊, 郑文庭. 一种参数化的表情映射方法 [J]. 计算机应用研究, 2004, 2004(10): 117-121.]
- Assadi A H, Eghbalnia H, McMenamin B W. Information processing of motion in facial expression and the geometry of dynamical systems [A]. In: Proceedings of SPIE Vision Geometry XIII [C], San Jose, CA, USA, 2005, 5675: 13-24.
- Zheng Nan-ning, Fu Yun, Zhang Ting, *et al.* Facial expression transformation, aging and invisible view reconstruction (1) [J]. Acta Electronica Sinica, 2003, 31(12A): 1955-1962. [郑南宁, 付昀, 张

- 婷等. 人脸的表情与年龄变换和非完整信息的重构技术(上) [J]. 电子学报, 2003, **31**(12A):1955-1962. ]
- 25 Du Y Z, Lin X Y. Dynamic facial texture generation based on shape-appearance dependence mapping strategy [J]. Journal of Software, 2004, **15**(2):308-316.
- 26 Du Y Z. Research on Statistics-based Realistic Face Image Synthesis [D]. Beijing:Tsinghua University, 2004. [杜杨洲. 基于统计学习的人脸图像合成方法研究[D]. 北京:清华大学, 2004. ]
- 27 Lanitis Andreas, Taylor Chris J, Cootes Timothy F. Automatic interpretation and coding of face images using flexible models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, **19**(7):743-756.
- 28 Gao Wen, Chen Xi-lin, Yan Jie, *et al.* Synthesis of facial behavior for virtual human[J]. Chinese Journal of Computers, 1998, **21**(8):694-703. [高文, 陈熙霖, 晏洁等. 虚拟人面部行为的合成[J]. 计算机学报, 1998, **21**(8):694-703. ]
- 29 Lin Xue-yin, Hong Peng-yu. Facial expression image synthesis by using 2D model with uncoupled geometric and textural information [J]. Acta Electronica Sinica, 1998, **26**(11):124-127. [林学茵, 洪鹏宇. 基于2D几何与纹理信息分离模型的表情分析与合成[J]. 电子学报, 1998, **26**(11):124-127. ]
- 30 Dai Peng, Xu Guang-yu. Facile face animation toward computer-mediated communication [J]. Journal of Computer-aided Design & Computer Graphics, 2008, **20**(6):793-800. [戴鹏, 徐光裕. 面向网上人际交流的便捷人脸动画[J]. 计算机辅助设计与图形学学报, 2008, **20**(6):793-800. ]
- 31 Kalra P, Magnenat-Thalmann N. Modeling of vascular expressions in facial animation [A]. In: Proceedings of Computer Animation '94 [C], Geneva, Switzerland, 1994:50-58, 201.
- 32 Tsumura N, Uetsuki K, Ojima N, *et al.* Correlation map analysis between appearances of Japanese facial images and amount of melanin and hemoglobin components in the skin [A]. In: Proceedings of SPIE Vol. 4299, Human Vision and Electronic Imaging VI [C], San Jose, CA, USA, 2001:252-260.
- 33 Yamada T, Watanabe T. Effects of facial color on virtual facial image synthesis for dynamic facial color and expression under laughing emotion [A]. In: Proceedings of 13th IEEE International Workshop on Robot and Human Interactive Communication [C], Okayama, Japan, 2004:341-346.
- 34 Yamada T, Watanabe T. Analysis and synthesis of facial color for the affect display of virtual facial image under fearful emotion [A]. In: Proceedings of the 2005 International Conference on Active Media Technology [C], Kagawa, Japan, 2005:219-224.
- 35 Yamada T, Watanabe T. Virtual facial image synthesis with facial color enhancement and expression under emotional change of anger [A]. In: Proceedings of 16th IEEE International Conference on Robot and Human Interactive Communication [C], Jeju, Korea, 2007:49-54.
- 36 Park K H, Kim T Y. Facial color adaptive technique based on the theory of emotion-color association and analysis of animation [A]. In: Proceedings of the 10th IEEE Workshop on Multimedia Signal Processing [C], Carins, Queensland, Australia, 2008:861-866.
- 37 Khanam Assia, Mufti Muid. Intelligent expression blending for performance driven facial animation [J]. IEEE Transactions on Consumer Electronics, 2007, **53**(2):578-584.
- 38 Cao X, Guo B. Real-time tracking and imitation of facial expression [A]. In: Proceedings of the Second International Conference on Image and Graphics [C], Hefei, Anhui, China 2002, **4875**:910-918.
- 39 Bickel B, Botsch M, Angst R, *et al.* Multi-scale capture of facial geometry and motion [J]. ACM Transactions on Graphics, 2007, **26**(3):33-1-33-10.
- 40 Ig-jae Kim, Hyeong-seok Ko. Intuitive quasi-eigen faces [A]. In: Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia [C], Perth, Australia, 2007:33-38.
- 41 Deng Z, Chiang P Y, Fox P, *et al.* Animating blendshape faces by cross-mapping motion capture data [A]. In: Proceedings of the 2006 Symposium on Interactive 3D Graphics and Games [C], Redwood City, CA, USA, 2006:43-48.
- 42 Williams L. Performance-driven facial animation [J]. Computer Graphics, 1990, **24**(4):235-242.
- 43 Guenter B, Grimm C, Wood D, *et al.* Making faces [A]. In: Proceedings of the ACM Siggraph Conference on Computer Graphics [C], Orlando, FL, USA, 1998:55-65.
- 44 Kass M, Witkin A, Terzopoulos D. Snakes: active contour models [A]. In: Proceedings of First International Conference on Computer Vision [C], London, England, 1987, **53**:259-268.
- 45 Chai J X, Xiao J, Hodgins J. Vision-based control of 3D facial animation [A]. In: Proceedings of the ACM Siggraph/ Eurographics Symposium on Computer Animation [C], San Diego, CA, USA, 2003:193-206.
- 46 Wang Rong-rong, Qiu Xian-jie, Wang Wen-zhong, *et al.* A video-driven approach to 3D human animation synthesis [J]. Journal of System Simulation, 2007, **19**(8):1700-1709. [王榕榕, 邱显杰, 王文中等. 一种视频驱动的三维人体动画合成方法[J]. 系统仿真学报, 2007, **19**(8):1700-1709. ]
- 47 Vlasic D, Brand M, Pfister P, *et al.* Face transfer with multilinear models [J]. ACM Transactions on Graphics, 2005, **24**(3):426-433.
- 48 Duan Hong, Wang Yi-xiao, Cheng Yi-min, *et al.* Drive muscle face model using single video [J]. Acta Simulata Systematica Sinica, 2003, **15**(9):1268-1271. [段鸿, 王以孝, 程义民等. 人脸肌肉模型的单视频驱动[J]. 系统仿真学报, 2003, **15**(9):1268-1271. ]
- 49 Hu Y L, Yin B C, Kong D H. A new facial feature extraction method based on linear combination model [A]. In: Proceedings of IEEE/WIC International Conference on Web Intelligence [C], Halifax, Canada, 2003:520-523.
- 50 Sung J, Kim D. Pose-robust facial expression recognition using view-based 2D + 3D AAM [J]. IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans, 2008, **38**(4):852-866.
- 51 Sucontphunt T, Mo Z, Neumann U, *et al.* Interactive 3D facial expression posing through 2D portrait manipulation [A]. In: Proceedings of Graphics Interface 2008 [C], Windsor, Ontario, Canada, 2008:177-184.
- 52 Peng Antai, Hayes Monson H. Iterative human facial expression modeling [A]. In: Proceedings of the Third International Computer Science Conference on Image Analysis Applications and Computer Graphics [C], Hong Kong, China 1995:479-486.
- 53 Yu Zhong-ji, Li Ji-jun. An efficient algorithm of facial expression generation based on RBF network [J]. Computer Applications, 2005, **25**(7):1611-1615. [余重基, 李际军. 一种基于RBF网络生成人脸表情的算法[J]. 计算机应用, 2005, **25**(7):1611-1615. ]
- 54 Wu Chun-you, Wang Shi-tong. Applying soft computing technique to virtual actor's facial expression modeling [J]. Computer Engineering and Design, 2007, **28**(12):2914-2917. [武春友, 王士同. 软计算在虚拟角色面部表情建模中的应用[J]. 计算机工程与设计, 2007, **28**(12):2914-2917. ]