

# 基于形变模型的人脸建模及其应用综述

尹宝才, 胡永利, 程世铨, 谷春亮

(北京工业大学 计算机学院多媒体与智能软件技术实验室, 北京 100022)

**摘 要:** 形变模型(morphable model)是近几年提出的人脸建模新方法. 基于若干原型人脸的知识, 该方法实现了真实感三维人脸的自动建模, 并在图像分析、人脸识别、表情动画等方面的应用中取得了良好效果. 回顾了基于形变模型的人脸建模的研究概况、基本原理及应用情况, 并对形变模型的进一步应用和发展进行了展望.

**关键词:** 形变模型; 人脸建模; 人脸识别

**中图分类号:** TP 391.41

**文献标识码:** A

**文章编号:** 0254-0037(2003)03-0364-08

由于人脸不规则的表面, 非刚性的变形, 在不同姿态和光照下的复杂变化, 以及人脸的多样性和差异性, 使得人脸建模和表示一直是计算机图形学和计算机视觉领域极具挑战性的问题. 由于人脸建模和人脸动画在影视娱乐、公安安全、身份验证、虚拟现实、视频分析与编码传输等方面广阔的应用前景, 国内外研究者在这方面做了大量的研究工作, 近 10 年来, 真实感人脸建模迅速成为计算机图形学、计算机视觉、模式识别等领域的研究热点. 传统真实感人脸建模方法使用三维几何模型描述人脸形状, 并用人脸图像合成三维纹理. 该方法自动化程度低, 对于特定人的建模需要专业人员的大量人工交互工作. 随着基于形变模型的人脸建模方法的出现, 使得计算机人脸建模的研究进入了一个新的阶段. 形变模型是一种基于知识的、自动化的真实感人脸建模方法, 由于模型本身的优势以及模型在人脸识别、人脸动画等方面的成功应用, 使得该模型成为人脸建模最具前景的方法.

## 1 人脸建模研究概况

Parke 是第一个用计算机生成人脸图像的人, 他使用一种非常初级的多边形表示人的头部结构, 并在 1974 年实现了第一个参数化的人脸模型<sup>[1]</sup>, 通过模型参数的变化可以产生基本的人脸动画. 人脸参数模型表示复杂, 且缺乏真实感, 如图 1(a)所示. 1981 年 Platt 等人提出了一种基于物理肌肉控制的人脸表情模型<sup>[2]</sup>, 随后 Waters 提出了一种新的层状肌肉模型来模拟人脸面部表情<sup>[3]</sup>. Thalmann 等人也描述了一个抽象的肌肉运动模型<sup>[4]</sup>. 人脸物理肌肉模型可以比较真实地模拟人脸的运动变化, 但模型的计算复杂度很高, 模型产生的人脸同样缺乏真实感, 如图 1(b)所示.

由于参数模型和肌肉模型建模过程都比较复杂, 且缺乏真实感, 20 世纪 90 年代以后真实感人脸建模是人脸建模研究的主要目标, 这一阶段出现了许多有代表性的人脸建模方法. Parke 和 Waters 在文献[5]中对 1996 年以前的人脸建模和动画研究作了详尽的综述. 1995 年 Lee 等人首先实现了真实感的人脸模型, 并产生了人脸的各种表情<sup>[6]</sup>. 由于该模型使用真实人脸图像合成纹理, 因此具有较好的视觉效果, 但模型动画的驱动仍使用肌肉模型, 计算量大. 随后出现了基于一般几何模型变形的人脸建模方法, 该方法使用不同角度的人脸图像提取人脸特征点, 并使用视觉方法恢复这些点的三维位置, 然后修改一般人脸模

收稿日期: 2003-07-01

基金项目: 国家“973”计划基金资助项目(2001CCA03300); 国家“863”计划基金资助项目(2001AA114160);

北京市自然科学基金资助项目(D070601-01); 北京市科委基金资助项目(N070601-01);

北京市教委基金资助项目(P070701-01, P070702-01).

作者简介: 尹宝才(1963-), 男, 教授, 博士生导师.

型,最后进行人脸图像的纹理映射得到特定的三维人脸.该方法简单,真实感好,是比较实用的人脸建模方法,也是到目前为止研究和应用最为广泛的方法.其中典型的有 Guenter 等人的人脸造型技术<sup>[7]</sup>; Pighin 等人的基于图像的人脸造型和动画<sup>[8]</sup>; Zhengyou Zhang 等人的基于视频的快速人脸造型和动画<sup>[9]</sup>; Lee 等人的基于正交图像的三维人脸克隆及动画技术<sup>[10]</sup>,如图 1(c)所示.国内研究者也在这方面做了一定的工作,其中浙江大学实现了基于实拍图像的人脸真实感重建<sup>[11]</sup>.中国科学院计算所与北京工业大学实现了基于图像的三维人脸编辑生成系统<sup>[12]</sup>.虽然基于一般模型的三维人脸建模方法可产生真实感的人脸,但建模过程需要大量的人机交互工作.

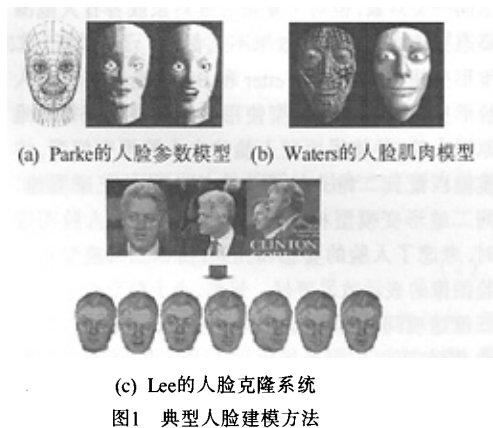


图1 典型人脸建模方法

Vetter 等人 1999 年提出基于形变模型的人脸建模方法,该方法第一次实现了人脸建模的自动化<sup>[13]</sup>,同时可以由一副人脸图像重建三维人脸.基于形变模型的人脸建模使用原型人脸的线性组合建立人脸参数化模型,将模型与特定人脸图像进行优化匹配,从而实现三维人脸的重建.由于该模型是建立在像素级对应的三维原型人脸数据上,且在模型中考虑了人脸姿态、光照等因素,因此该模型可以生成高度真实感的三维人脸.另外,通过分析原型人脸的表情变化,该模型还可以产生人脸表情和动画.

## 2 形变模型基本原理

形变模型的基础是线性组合思想,即一类对象可以用该类对象的基底的线性组合来表示. Ullman 等人<sup>[14]</sup>的研究指出,在正投影和不考虑自遮挡的条件下,利用物体不同的 3 个视角图像的线性组合就能合成任意视角的图像,这是线性组合在图像表示方面最初的尝试.线性组合思想在图像分析和合成方面的深入研究发展出了现在的形变模型,虽然在研究中不断对形变模型进行改进,但模型的基本思想没有变,即使用一类图像中的若干典型图像作为原型,并张成该类图像的一个子空间,使用子空间基底的线性组合表示该类图像的特定实例.

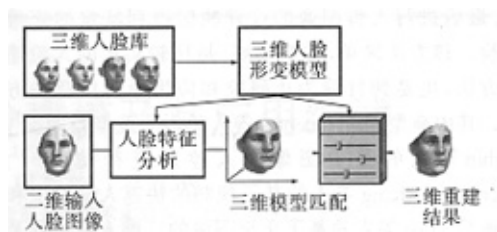
形变模型最初是在二维图像的分析表示方面取得成功的.实际上,形变模型的名称来源于图像处理中典型的图像生成方法——Morphing 方法,即通过有对应关系的两副图像产生中间图像和图像间的平滑过渡. Choi 等人提出的图像表示形式可以将一副图像分解为几何形状矢量和与形状无关的纹理矢量<sup>[15]</sup>,这使得一类图像进行线性运算成为可能. Poggio 等人提出线性对象类的概念(linear object classes)<sup>[16]</sup>,给出了三维对象满足线性对象类的充分条件,并论证了人脸对象在近似条件下是线性对象类.另外,基于光流算法的图像对应技术的发展,使建立图像间的像素对应有了理想的效果,尤其 Vetter 等人提出的光流改进算法——Bootstrapping 算法,可以完全自动化地建立一类图像间的像素级对应<sup>[17]</sup>.有了图像间像素对应,图像间的线性运算就可以通过其对应像素的线性运算实现.在 Jones 等人的实验中<sup>[18]</sup>,形变模型被用来表示人脸、汽车以及手写数字等对象,并取得了很好的效果.其中人脸模型中使用了 62 副正面照片作为原型人脸,使用光流算法建立像素级对应后,建立人脸图像的形变模型,并使用模型表示特定的人脸图像.图 2 是部分原型人脸和特定人脸表示的结果.二维图像形变模型除了在图像分析表示方面的应用外,还可用于图像压缩编码、对象识别(如人脸识别),以及三维对象恢复估计等方面.

虽然使用二维图像形变模型可以比较好地表



图2 形变模型用于图像分析和合成

示同一类对象,但对于复杂三维对象或者有大范围姿态变化的人脸表示效果不是很好.于是,基于二维形变模型的思想,Vetter和Blanz提出了三维人脸形变模型<sup>[13]</sup>,该模型使用激光扫描仪获得三维原型人脸,并使用原型人脸建立人脸形变模型,将模型匹配到二维人脸图像实现人脸的三维重建.同二维形变模型相比,三维模型在表示人脸图像时,考虑了人脸的姿态和光照,所以三维模型对人脸图像的表示效果更好.另外,由于模型将人脸的三维建模问题看作模型和给定人脸图像的误差优化,因此可以使用最优化的方法解决,从而实现基于图像的特定人脸自动建模.图3是三维形变模型用于人脸的建模过程和建模的典型结果.



(a) Vetter提出的三维人脸建模过程<sup>[13]</sup>



(b) 左为原始人脸图像,右为使用模型恢复的人脸图像

图3 三维人脸形变模型

### 3 形变模型的建立

形变模型的建立包括原始数据获取、原型人脸数据量化、人脸图像像素对应的建立和使用主元分析方法(PCA)建立人脸参数模型等处理步骤.

#### 3.1 三维原型人脸数据库

文献[13]中采集了男女青年各100个共200人的三维人脸数据作为原型人脸,建立三维原型人脸库.三维人脸数据使用CyberWare激光扫描仪获取,该激光扫描仪可以扫描得到人头部的几何信息和彩色纹理信息.扫描仪使用柱面坐标来记录几何信息,扫描精度为圆周方向(用 $\phi$ 表示)512个采样点,轴方向(用 $h$ 表示)512个采样点(300 mm),扫描半径(用 $r$ 表示)260~340 mm.每一个几何采样点对应一个24位彩色(用R、G、B表示)纹理像素点,并以512×512大小的纹理图像存储.由于该扫描仪对于头发等结构复杂或黑色部位的扫描效果比较差,而这里感兴趣的只是人面部的信息,因此被扫描者都戴了浴帽,没有化妆,没有眼镜和饰物.获取原始数据后对其进行预处理,去除浴帽、耳朵后面、颈部以下等部分,保留面部数据.预处理后,每一个人脸大约由7万个采样点组成.

#### 3.2 人脸数据像素对应

建立形变模型的基础是原型数据间要满足线性运算,对于这里的三维人脸数据,即要基于面部特征建立不同人脸间点对点的稠密对应关系.由于人脸的几何差异,要在三维数据上建立基于特征的点对点的对应非常困难,而原型人脸的实际表示是柱面坐标,因此可以通过柱面坐标将三维数据转化为二维图像并借助图像间的光流算法<sup>[19]</sup>进行对应计算,即将三维人脸沿柱面轴展开形成二维纹理图像和几何形状图像,然后在二维图像上计算点与点的对应关系.

光流算法一般用于物体的运动估计,即使用视频序列相邻两帧图像中物体连续平滑变化的约束来计算点的位移.由于人脸图像的相似性,可以将不同人脸图像近似地看作视频中相邻两帧图像,并使用光流计算对应点位置变化.但是当人脸图像的差异比较大时就不能满足光流算法的平滑性约束,因此计算的对应效果不理想.为此,Vetter等人提出了Bootstrapping算法<sup>[17]</sup>,使用中间临时形变模型改进光流算法,很好地解决了人脸图像间的像素对应问题.

有了二维柱面展开图的像素对应,返回到三维空间即可得到三维人脸的点对点的稠密对应.由于不同三维人脸的几何点数量不一定相同,因此可以使用光流建立的稠密点对应通过插值等办法将不同人脸的几何点数规格化成相同数量,从而原型人脸可用下面统一的向量形式来表示

$$S_i = (X_{i1}, Y_{i1}, Z_{i1}, X_{i2}, \dots, X_{in}, Y_{in}, Z_{in})^T$$

$$T_i = (R_{i1}, G_{i1}, B_{i1}, R_{i2}, \dots, R_{in}, G_{in}, B_{in})^T \quad 1 \leq i \leq N \quad (1)$$

其中:  $S_i$  是第  $i$  个人脸的点的三维坐标组成的几何形状向量;  $T_i$  是对应于形状向量中点的 RGB 值组成的纹理向量;  $N$  是原型人脸个数;  $n$  是规格化后人脸的点数。

### 3.3 模型建立

有了上面规格化后的人脸向量, 就可以建立人脸形变模型。使用原型人脸的线性组合产生新人脸的直观想法可用下面形式表示

$$S_{\text{new}} = \sum_i^N a_i S_i \quad T_{\text{new}} = \sum_i^N b_i T_i, \quad \text{其中} \quad \sum_i^N a_i = \sum_i^N b_i = 1 \quad (2)$$

因为在上面的人脸线性组合中, 原型人脸数量比较大 ( $N=200$ ), 且人脸原始数据间有一定相关性, 所以使用主元分析的方法 (PCA) 对模型数据进行处理, 即对形状和纹理分量进行 PCA 变换。通过 PCA 变换可以压缩数据量, 且由于变换的正交性可以消除数据相关性。

先计算  $N$  个形状向量和纹理向量的均值  $\bar{S}$ 、 $\bar{T}$ , 然后计算标准化后的形状向量  $\{S_i - \bar{S}\}$  和纹理向量  $\{T_i - \bar{T}\}$  的自相关阵  $C_S$ 、 $C_T$ , 对  $C_S$ 、 $C_T$  分别求特征值和特征向量, 根据特征值的贡献率确定前  $M$  个最大的特征值  $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_M)$   $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_M)$  及相应的特征向量  $(s_1, s_2, \dots, s_M)$ ,  $(t_1, t_2, \dots, t_M)$ 。这时式 (2) 的模型可表示为:

$$S_{\text{model}} = \bar{S} + \sum_i^{m-1} \alpha_i s_i \quad T_{\text{model}} = \bar{T} + \sum_i^{m-1} \beta_i t_i \quad (3)$$

可假设人脸数据满足正态分布, 经过正交变换后形变参数  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{m-1})$ ,  $\beta = (\beta_1, \beta_2, \dots, \beta_{m-1})$ , 有  $P(\alpha) \sim \exp[-\frac{1}{2}(\sum_{i=1}^{m-1} (\alpha_i^2 / \sigma_i))]$ ,  $P(\beta) \sim \exp[-\frac{1}{2}(\sum_{i=1}^{m-1} (\beta_i^2 / \lambda_i))]$ 。式 (3) 即为三维人脸模型的最终表示形式, 给定形变参数  $\alpha, \beta$  即可产生新的三维人脸。

## 4 形变模型的优化

形变模型的优化过程实际上是对特定人脸的三维建模过程, 即将模型与特定人脸图像进行匹配得到模型参数。

### 4.1 模型误差

有了三维人脸形变模型, 对于给定的人脸图像要进行三维建模, 就是要调节模型参数, 使模型产生的三维人脸在相同视点看到的图像与给定人脸图像的误差最小。如果使用图像对应像素点的灰度差的平方和作为两图像的误差, 则要求得模型参数使得下式最小:

$$E_I = \sum_{x,y} \|I_{\text{input}}(x,y) - I_{\text{model}}(x,y)\|^2 \quad (4)$$

其中:  $I_{\text{input}}$  是给定人脸图像;  $I_{\text{model}}$  是模型产生的人脸图像。由于  $I_{\text{model}}$  是由三维模型产生的模型图像, 所以在上面误差中除了模型参数  $\alpha, \beta$  外, 还涉及摄像机的水平和俯仰角度、尺度变换因子、像平面上的旋转和平移、环境光和方向性光源等与投影和光照有关的参数, 使用  $\rho$  来表示这些参数。实际计算时, 使用了 Phong 光照模型和透视投影来计算  $I_{\text{model}}$ 。

因为由二维图像重建三维信息的问题本身是一个病态的问题, 所以这里要对参数  $\alpha, \beta, \rho$  使用先验经验进行合理性约束。  $\alpha, \beta$  由前面给定的分布进行约束,  $\rho$  的先验分布可用正态分布进行估计。给定  $\alpha, \beta, \rho$  时, 图像  $I_{\text{input}}$  的观察受标准差为  $\sigma_N$  的高斯噪音的影响,  $P(I_{\text{input}} | \alpha, \beta, \rho) \sim \exp(-E_I / (2\sigma_N^2))$ 。从而问题转化为给定特定人脸图像的最大后验估计问题。由 Bayes 估计理论, 后验概率的最大化等价于下式最小化:

$$E = \frac{E_I}{\sigma_N^2} + \sum_{j=1}^{m-1} \frac{\alpha_j^2}{\sigma_{\alpha,j}^2} + \sum_{j=1}^{m-1} \frac{\beta_j^2}{\sigma_{\beta,j}^2} + \sum_j \frac{(\rho_j - \bar{\rho}_j)^2}{\sigma_{\rho,j}^2} \quad (5)$$

从而模型对于特定人脸图像的匹配问题转化为上式关于参数  $\alpha, \beta, \rho$  的最小优化问题。

## 4.2 随机优化

对式(5)的优化使用梯度下降法来解决. 由于式(5)中的  $E_i$  由式(4)给定, 需要计算三维模型投影产生的图像, 而三维模型的点数  $n$  约为 7 万, 要进行所有点的光照和投影计算, 显然开销很大, 因此文献[13]中采取了随机梯度下降法<sup>[20]</sup>来解决式(5)的优化. 随机梯度下降法就是在计算模型人脸图像和给定人脸图像的误差时, 随机在三维人脸模型上选择一定数量的点(实际上使用三角面的重心点)来计算光照和投影点, 而不是在所有模型点上计算, 即用式(6)代替  $E_i$

$$E_K \approx \sum_{k \in K} \|I_{\text{input}}(\bar{x}_k, \bar{y}_k) - I_{\text{model}}(\bar{x}_k, \bar{y}_k)\|^2 \quad (6)$$

其中:  $K$  是随机点集;  $(\bar{x}_k, \bar{y}_k)$  是随机点在图像平面上的投影点. 使用随机梯度下降法不但可以大大加速最小优化的效率, 而且可以避免局部最小对优化的影响.

## 4.3 模型匹配的进一步改进

由于形变模型的原型人脸数据量非常大 ( $N = 200$ ), 一个原型人脸由维数约为 21 万的形状和纹理向量表示, 且模型匹配过程是多参数的非线性优化问题, 因此形变模型建模的运算和内存开销都非常大. 文献[13]中, 模型匹配一副人脸在 SGI R10000 上需要约 50 min 的运算. 所以如何加快模型匹配速度, 提高模型匹配的真实感效果, 避免模型收敛到局部极小点, 是模型进一步应用的关键. 为此采用了由粗到细、逐步求精的策略改进形变模型. 主要的改进体现在以下方面:

1) 使用多分辨率模型进行匹配, 以避免局部极小点的干扰. 建立多分辨率形变模型, 并建立输入图像的图像金字塔, 参数匹配首先在低分辨率的图像和模型间求出, 并将结果用于高一级分辨率的匹配, 逐步提高分辨率, 直到最后对原始尺寸进行匹配.

2) 保持光照和姿态参数不变, 先对模型主要特征向量对应的参数进行迭代, 再逐步增加特征向量的个数进行迭代. 这样可以使人脸主要特征优先收敛到目标, 而减少次要特征对模型匹配产生主导影响.

3) 在模型匹配开始使用较大的图像噪声  $\sigma_N$ , 使初始解最大限度地包含合理解, 然后逐步缩小  $\sigma_N$  使匹配更精确, 顾及到图像的细节.

4) 为了更进一步表现人脸器官的细节, 使用分片策略. 将原始人脸整体模型分割为眼睛、鼻子、嘴巴和脸颊 4 个区域, 并分别建立各个区域的形变模型. 在匹配时, 先用整体形变模型进行匹配, 然后使用分片模型分别对各个区域分别进行匹配, 即细化人脸器官的各个细节, 从而提高模型效果. 分片策略实际上增加了模型参数集的自由度, 这样可以改善模型的细节, 但带来的一个问题就是最终建模时, 要对各分片区域进行缝合处理.

# 5 形变模型应用

由于基于形变模型的三维人脸建模可以自动化地、真实感地恢复二维人脸图像的三维结构, 并且能很好地估计人脸的姿态和光照信息, 因此该方法在人脸识别、表情动画、性别识别、特征提取以及模型基编码等方面有很好的应用前景, 并解决了以往基于图像的方法无法解决的一些难点问题. 这里主要介绍形变模型在人脸识别和表情动画方面的应用.

## 5.1 人脸识别

人脸识别的研究已经有 30 多年的历史了, 虽然现有的人脸识别系统已经具有很好的效果<sup>[21, 22]</sup>, 但对于多姿态、不同光照的人脸识别问题依然没有得到很好解决. 主要原因是, 基于图像分析的人脸识别方法本身就缺失人脸三维结构和纹理变化信息, 虽然通过视觉补偿和一般三维人脸模型可以对人脸姿态和光照做粗略的估计, 但本质上不能解决特定人脸的光照和姿态变化问题. 由于三维人脸形变模型在模型匹配时, 对光照和人脸姿态进行参数估计, 与人脸形状和纹理参数同时进行优化, 因此将三维形变模型用于

人脸识别可以很好地解决人脸识别中姿态和光照问题。

最先 Heisele 等人将三维人脸形变模型用于支持向量机(SVM)人脸分类模型的训练<sup>[23]</sup>,把使用三维人脸模型投影产生的不同姿态、不同光照的人脸图像作为多姿态、多光照人脸识别模型的训练集。文献[24]中Leopold 等人从神经心理学角度对三维人脸形变模型中的人脸变形规律以及人脸的识别特征问题作了分析研究。Romdhani 等人<sup>[25]</sup>将三维人脸形变模型的匹配参数用于人脸识别,实现了多姿态、多光照人脸识别。该方法使用三维人脸形变模型重建二维人脸图像的三维模型,并使用模型的形状和纹理参数作为人脸特征来进行人脸识别,图4是该方法的识别过程。

## 5.2 人脸表情和动画

三维人脸肌肉模型可以通过模拟人脸肌肉的运动产生效果不错的人脸动画,但生成动画的计算过程非常复杂,计算量很大<sup>[3]</sup>。基于一般模型的三维人脸形变模型可以通过表情特征点的变形对模型进行插值产生动画,但这种动画很难特定化,产生人脸动画时需要大量的人机交互标定人脸特征点<sup>[8]</sup>。三维人脸形变模型通过表情人脸的学习将人脸表情等人脸属性参数化,通过将这些参数向量赋予特定三维人脸产生人脸表情。

文献[13]中,人脸的基本表情,如微笑、沮丧等可以使用有表情人脸与中性表情人脸的差来参数化,即计算  $\Delta S = S_{\text{expression}} - S_{\text{neutral}}$  和  $\Delta T = T_{\text{expression}} - T_{\text{neutral}}$ ,并将  $(\Delta S, \Delta T)$  加权与特定三维人脸向量  $(S^*, T^*)$  相加减产生表情人脸  $(S^* \pm \mu \Delta S, T^* \pm \mu \Delta T)$ 。对于人脸的其他属性,如人脸的胖瘦、人脸的性别差别等属性可通过人为选择一定数量的人脸,根据其特征的明显程度给定权值来进行特征学习,即计算

$$\Delta S = \sum_{i \in Q} \mu_i (S_i - \bar{S}), \quad \Delta T = \sum_{i \in Q} \mu_i (T_i - \bar{T}),$$

其中  $Q$  为选定的特征人脸集,权值  $\mu_i$  可以通过回归的方法进行估计。同样,将  $(\Delta S, \Delta T)$  加权与特定三维人脸向量相加减即可产生表情人脸。有了表情人脸就可以通过简单插值等方法产生人脸动画。图5所示为文献[13]中给出的人脸形变和表情。

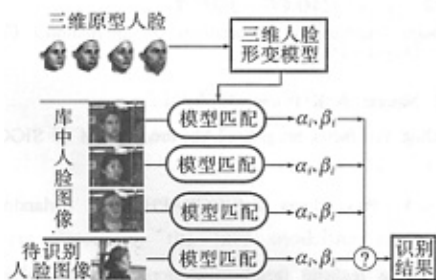


图4 人脸形变模型用于人脸识别



图5 人脸形变和表情

## 6 进一步的研究和展望

作为一种新的人脸建模方法,基于形变模型的人脸建模从提出到现在还不到10年时间,但已经获得了良好的应用效果。作者认为在跟踪最新研究成果的同时,应主要在以下方面开展研究工作。

第1,建立面向中国人特征的三维人脸数据库。由于形变模型是一种基于知识的人脸建模方法,建模效果直接依赖于原型人脸数据库的多样性和丰富性。国外的研究基于面向欧洲人种的MPI人脸库,该库的数据不适合中国人的特征,且库的容量比较小(200人)。因此,建立中国人的三维人脸库是开展该方面研究的基础,可以考虑建立不同性别、不同年龄、不同表情的大容量的三维人脸数据库,为国内的研究提供良好的数据平台。

第2,提高原型人脸的对应效果。建立形变模型的基础是三维人脸数据点对点的稠密对应,这种对应

建立的好坏直接影响模型匹配的效果。虽然现有的光流对齐算法可以实现原型人脸数据的自动化对应,且有较好的效果,但光流是一种基于全局约束的算法,在计算时没有考虑人脸的局部特征,且原型人脸数据的空间点数并不一样,因此现有的对应不是严格意义上的基于特征的点对点对应。考虑到原型人脸的对应是离线进行的,为了取得更好对应效果,可以借助图形学重采样等技术,基于人脸特征将原型人脸数据进行人工分割,并重采样建立拓扑结构完全相同的人脸数据,从而实现真正意义上的基于特征的点对点的对应。

第3,进一步加速模型的匹配优化。虽然现有的模型采用了随机优化、多分辨率等提高效率的方法,但模型的优化效率对于要求实时性的应用(人脸识别)还是不够的。因此可以考虑使用多种方法加速模型的优化。直观地,可以使用牛顿法等更好的正面解决算法来提高算法效率;从图形学和人眼视觉感受来看,描绘一个真实感的三维人脸,一般只需要几千个空间点,从而形变模型中实际上存在大量的数据冗余,因此可以考虑将模型原型数据进行简化来降低运算量,但这样会对模型基于稠密点对点的许多思想产生影响;另外,如果不正面解决优化问题,是否可以考虑用机器学习的思想,通过已有的匹配实例来不断学习人脸图像与模型参数间的关系,最终使用学习规律代替优化过程。

第4,给模型加入头发模型。由于原始数据获取的原因,现有的模型没有考虑头发等区域。要产生完整的三维头部模型,可以考虑将图形学的头发建模技术同形变模型结合起来,产生视觉上更好的模型。

最后,对模型进行更为广泛的应用研究。除了在人脸建模、人脸识别、人脸动画方面的应用,可以进一步探讨和研究该模型在性别验证、人物老化模拟、虚拟现实以及与语音技术结合产生活者人脸等多个应用领域展开研究。

#### 参考文献:

- [1] PARKE F I. A Parametric Model of Human Faces[D]. Salt Lake City: University of Utah, 1974.
- [2] PLATT S, BADLER N. Animating facial expression[J]. Computer Graphics, 1981, 15(3): 245-252.
- [3] WATERS K. A muscle model for animating three-dimensional facial expression[J]. Computer Graphics, 1987, 21(4): 17-24.
- [4] THALMANN N, THALMANN D. New trends in computer graphics[A]. Proceedings of International Computer Graphics '88[C]. Tokyo: Springer Verlag, 1988. 403-415.
- [5] PARKE F I, WATERS K. Computer Facial Animation[M]. Natick: A K Peters, 1996.
- [6] LEE Y, TERZOPOULOS D, WATERS K. Realistic modeling for facial animation[A]. Proceedings of SIGGRAPH'95[C]. Los Angeles: ACM Press, 1995. 55-62.
- [7] GUENTER B, GRIMM C, WOOD D, et al. Making faces[A]. Proceedings of SIGGRAPH'98[C]. Orlando: ACM Press, 1998. 55-66.
- [8] PIGHIN F, HECKER J, LISCHINSKI D, et al. Synthesizing realistic facial expressions from photographs[A]. Proceedings of SIGGRAPH'98[C]. Orlando: ACM Press, 1998. 75-84.
- [9] LIU Zicheng, ZHANG Zhengyou, JACOBS C, et al. Rapid modeling of animated faces from video images[J]. Journal of Visualization and Compute Animation, 2001, 12(4): 227-240.
- [10] LEE W S, THALMANN N M. Fast head modeling for animation[J]. Journal Image and Vision Computing, 2000, 18(4): 355-364.
- [11] 梅丽, 鲍虎军. 基于实拍图像的人脸真实感重建[J]. 计算机学报, 2000, 23(9): 996-1002.
- [12] 晏洁, 高文, 尹宝才. 具有真实感的三维虚拟特定人脸生成方法[J]. 计算机学报, 1999, 22(2): 147-153.
- [13] BLANZ V, VETTER T. A morphable model for the synthesis of 3D faces[A]. Proceeding of SIGGRAPH'99[C]. Los Angeles: ACM Press, 1999. 187-194.
- [14] ULLMAN S, BASRI R. Recognition by Linear Combinations of Models[R]. Cambridge: MIT AI Lab Publications Office, 1989.
- [15] CHOI C S, OKAZAKI T, HARASHIMA H, et al. A system of analyzing and synthesizing facial images[A].

- International Symposium of Circuit and Syatems (ISCAS91)[C]. Singapore City: IEEE Computer Society Press, 1991. 2665-2668.
- [16] VETTER T, POGGIO T. Linear object classes and image synthesis from a single example image[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 733-742.
- [17] VETTER T, JONES M J, POGGIO T. A bootstrapping algorithm for learning linear models of object classes[A]. IEEE Conference on Computer Vision and Pattern Recognition—CVPR'97[C]. Puerto Rico: IEEE Computer Society Press, 1997. 40-47.
- [18] JONES M J, POGGIO T. Model-Based Matching by Linear Combinations of Prototypes[R]. Cambridge: MIT AI Lab Publications Office, 1996.
- [19] HORN B K P, SCHUNK B G. Determining optical flow[J]. Artificial Intelligence, 1981, 17(1-3): 185-201.
- [20] VIOLA P, WELLS III W M. Alignment by maximization of mutual information[J]. International Journal of Computer Vision, 1997, 24(2): 137-154.
- [21] CHELLAPPA R, WILSON C L, SIROHEY S. Human and machine recognition of faces: A survey[J]. Proceedings of the IEEE, 1995, 83(5): 705-740.
- [22] HIELMAS E. Face detection: A survey[J]. Computer Vision and Image Understanding, 2001, 83(3): 236-274.
- [23] HUANG J, BLANZ V, HEISELE B. Face recognition using component-based SVM classification and morphable models[A]. First International Workshop SVM 2002[C]. Niagara Falls: Springer-Verlag, 2002. 334-341.
- [24] LEOPOLD D A, O'TOOLE A J, VETTER T, et al. Prototype-referenced shape encoding revealed by high-level[J]. Aftereffects Nature Neuroscience, 2001, 4(1): 89-94.
- [25] ROMDHANI S, BLANZ V, VETTER T. Face identification by fitting a 3D morphable model using linear shape and texture error functions[A]. European Conference on Computer Vision 2002[C]. Copenhagen: Springer-Verlag, 2002. 3-19.

## Face Modeling Based on Morphable Model and Its Application

YIN Bao-cai, HU Yong-li, CHENG Shi-quan, GU Chun-liang

( Multimedia and Intelligent Software Technology Laboratory, College of Computer Science,  
Beijing University of Technology, Beijing 100022, China )

**Abstract:** Recently, the face modeling method based on the knowledge of prototype face, the morphable model, is proposed to reconstruct 3D face from facial images automatically. This model uses the linear combination of the 3D prototypes to interpret facial image. And the model succeeds in applications of image analysis, face recognition and expression animation, etc. The authors review the study of morphable model and its' development and applications, and presents the prospects for the future research on the model and its' application.

**Key words:** morphable model; face modeling; face recognition