

# 基于稀疏形变模型的三维卡通人脸生成

署 光,姚莉秀,杨晓超,左 昕,杨 杰

(上海交通大学图像处理与模式识别研究所,上海 200240)

**摘 要:** 随着数字娱乐产业的发展,由照片生成卡通人脸的技术将取得广泛应用.此前的方法主要集中在平面卡通化的领域,风格较为单一.对于三维人脸,尽管形变模型方法可以由照片合成各种属性的三维人脸,但它计算量较大,不适用于实时应用场合.本文提出了一种基于稀疏形变模型的三维卡通人脸生成方法,提高了计算速度,且只需要单幅正面人脸照片.首先由稀疏形变模型拟合照片人脸获得特定的稀疏人脸模型;然后将一个一般人脸模型变形到特定人脸并合成纹理;最后对三维人脸进行卡通化.实验结果证明本文方法能够快速自动地合成生动的三维卡通人脸.

**关键词:** 卡通人脸; 三维重建; 稀疏形变模型; 肤色模型

**中图分类号:** TP391.41      **文献标识码:** A      **文章编号:** 0372-2112 (2010) 08-1798-05

## 3D Facial Caricaturing Based on Sparse Morphable Model

SHU Guang, YAO Li-xiu, YANG Xiao-chao, ZUO Xin, YANG Jie

(*Institute of Image Processing and Pattern Recognition, Shanghai JiaoTong University, Shanghai 200240, China*)

**Abstract:** Facial caricaturing from images could be widely used in digital entertainment and emotional communication. Previous methods mainly focused on 2D caricature generation. Although the Morphable Model can synthesize 3D faces with various attributes, it is not fit for the real-time application as the high computational cost. We proposed an improved 3D caricaturing method based on Sparse Morphable Model. It is much faster and only requires a single image. Firstly a sparse face model is reconstructed from the image by Sparse Morphable Model. Then a General Face Model is warped and texture is synthesized. Finally the 3D model is caricatured. Experimental results demonstrate that this method can generate impressive 3D caricature from a single frontal image fast and automatically.

**Key words:** facial caricature; 3D reconstruction; sparse morphable model; skin model

### 1 引言

随着网络和移动通讯的发展,图像在通讯中取得了较多应用,它们比文字或语音能更好地传达信息和表达感情.人脸图像在通讯中也有着重要作用,由于卡通人脸比真实人脸更生动有趣且特征鲜明,由照片合成相应的带有丰富表情卡通人脸将有助于人们的情感交流,也会在数字娱乐等领域得到广泛的应用.

过去出现了一些由计算机自动生成卡通人脸的方法<sup>[1,2]</sup>,它们多用于平面的卡通人脸变形,因此效果较为单一. Fujiwara 等<sup>[3]</sup>使用一种距离图像采集仪器获取了特定人脸的 2.5D 距离图像,然后变形生成了相应的三维卡通人脸,它比平面卡通有着更丰富的姿态、光照和表情,不过该方法需要使用特殊仪器对采集距离图像,难以实际应用.

由于目前的数码相机等设备可以随时拍摄人脸照片,则通过照片直接生成三维卡通人脸会利于实时通讯

的应用. Akimoto 等<sup>[4]</sup>由人脸的正侧面照片重建出三维模型,但在实际中也难以同时满足同时获得正侧面两幅照片. Blanz 和 Vetter<sup>[5]</sup>提出了基于形变模型的三维人脸重建方法,只需要单幅照片,通过稠密三维样本的线性组合来拟合对应的三维模型. Jiang 等人<sup>[6]</sup>通过在形变模型匹配过程中使用二维特征点,提高了计算速度.基于形变模型的重建方法自动化程度高,结果真实感好.但是形变模型的构建需要大量稠密点对应的三维人脸样本,尽管 Bootstrapping<sup>[7]</sup>光流法可以计算点对应,但是对于样本间差异较大的样本库则很难得到稳定的结果.另外,形变模型方法也不适用于复杂的光照环境,且大量样本数据的优化匹配过程会消耗很多时间和空间计算成本,不适用于实时环境或者在移动通讯中的应用.

受上述工作的启发,我们提出了一种快速的三维卡通人脸生成方法,即使用稀疏的形变模型代替完整模型以提高计算速度和避免不稳定的稠密点对应过程;同时使用一个一般人脸模型变形代替大量样本组合以节

省空间.该方法只需要单幅正面照片作为输入,可以自动生成相应的三维卡通人脸,算法流程如图 1 所示:首先,对人脸进行检测并提取输入照片的 60 个特征点,利用稀疏形变模型重建照片人脸对应的稀疏三维人脸;然后将一般模型弹性变形得到特定模型并进行纹理映射;最后使用一种改进的三维卡通化方法得到卡通三维人脸.

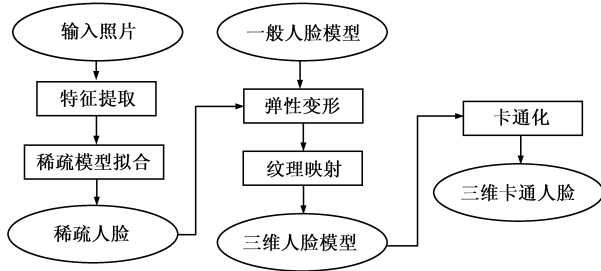
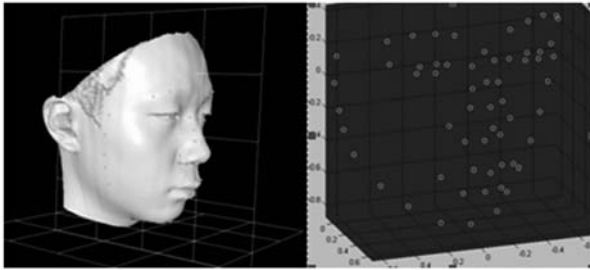


图1 三维卡通人脸生成的流程图

## 2 三维人脸重建

### 2.1 稀疏人脸模型

稀疏形变模型(Sparse Morphable Model, SMM)是稀疏人脸模型的线性组合.每一个稀疏人脸模型由 60 个三维特征点组成,包括脸、眉毛、眼睛、鼻子和嘴的轮廓,如图 2.选用这些点作为稀疏人脸是因为:它们不仅代表了用于卡通夸张时的人脸主要特征,同时也可以使用主动形状模型算法(Active Shape Model, ASM)自动提取.



(a) 完整人脸模型 (b) 60点组成的稀疏模型  
图2 完整模型与稀疏模型

本文使用中国三维人脸数据库“The BJUT-3D Large-Scale Chinese Face Database”创建 SMM.此数据库包含 200 个激光扫描的三维人脸模型;每个模型由约 60,000 个点和 120,000 三角形组成.通过手工交互在每一模型上选取一一对应的 60 个三维特征点作为一个稀疏人脸模型,表示为  $S = (X_1, Y_1, Z_1, \dots, X_{60}, Y_{60}, Z_{60})^T \in R^{3 \times 60}$ ,其中  $X, Y, Z$  是点坐标.通过这些稀疏对应的模型即可建立 SMM,并可合成任意的稀疏人脸  $S_{new}$ ,如式(1),其中  $a$  为权重系数.

$$S_{new} = Sa \quad (1)$$

为获取特定的稀疏人脸,需要获取照片人脸的特征点.

输入带有均衡光照和中性表情的单幅正面人脸照片,通过 AdaBoost 算法<sup>[8]</sup>检测出人脸区域,然后使用扩展 ASM 算法<sup>[9]</sup>即可自动提取出 60 个二维特征点,记为  $S_0 = (x_1, y_1, \dots, x_{60}, y_{60})^T$ ,如图 3 所示.



图3 特征点自动检测

由于 SMM 为三维模型,而  $S_0$  为二维点集,为了匹配二者,将三维点集  $S$  投影到  $S_0$  所在平面,得到二维样本点集  $S' = (X_1, Y_1, \dots, X_{60}, Y_{60})^T \in R^{2 \times 60}$ ,相应地获得二维的 SMM 为

$$S'_{new} = S'a \quad (2)$$

使用二维 SMM 拟合照片特征点,即找到使  $S'_{new}$  和  $S_0$  的距离  $\epsilon$  最小化所对应的  $a$ ,即

$$\epsilon = \min_a \|S_0 - S'a\| \quad (3)$$

考虑到  $S'_{new}$  与  $S_0$  的坐标系需要对齐,有

$$S'_{new} = cS_0 + d \quad (4)$$

其中  $c$  为尺度因子,  $d$  为位移因子,可由对数据预处理获得.于是问题转化为求式(5)中  $a$  的最小二乘解.

$$S'a = cS_0 + d \quad (5)$$

由式(6)可得到  $a$  的最小二乘解  $a_0$

$$a_0 = (S'^T S')^{-1} S'^T (cS_0 + d) \quad (6)$$

然后将求得的权重系数  $a_0$  代回式(2)求得一个特定的稀疏人脸  $S_p$ .

### 2.2 对一般人脸模型的变形

求得  $S_p$  后,可根据对应关系将一个稠密的一般人脸模型(General Face Model, GFM)变形为特定人脸.GFM 是一个削弱个人特征的平均人脸,它不仅用作变形的初始人脸,也在卡通夸张中作为对比的标准模板.GFM 可由人脸库中的大量样本的平均来创建,但是由于完整人脸之间没有建立起稠密对应,无法直接进行平均处理.本文利用稀疏人脸的对应关系来获得稀疏的平均特征点集,再使用一个与之最相似的完整人脸样本变形得到 GFM.步骤如下:

(1)由式(7)求出  $N$  个稀疏样本的平均稀疏人脸  $S_a$ ,本文中  $N = 200$ .

$$S_a = \frac{1}{N} \sum_{i=1}^N S_i \quad (7)$$

(2)逐一比较此  $N$  个稀疏人脸样本与  $S_a$  的距离,

将最接近  $S_a$  的稀疏人脸作为参考脸。

(3) 将参考脸的完整模型弹性变形为  $S_a$  对应的人脸, 即 GMF。

其中弹性变形算法采用薄板样条插值算法 (TPS)<sup>[10]</sup>。TPS 是一种径向基函数 (RBF) 插值, 可用来产生平滑的变形表面。生成完整的 GMF 后, 即可再次利用 TPS 算法将其变形为特定稀疏人脸  $S_p$  所对应的完整人脸, 如图 4。

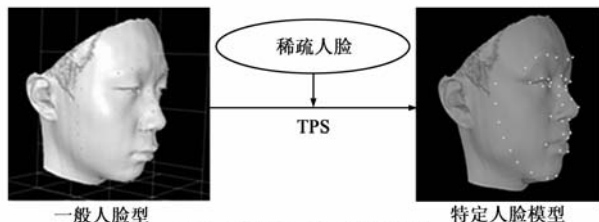


图4 基于TPS的弹性变形

### 2.3 纹理合成

合成准确的肤色纹理可以提高三维人脸的真实感, 但使用单幅正面照片将不可避免造成侧面人脸信息的缺失。在文[5]中通过建立纹理样本模型来合成虚拟的纹理, 计算代价较高, 难以用于实时场合。在文[6]中, 将照片的正面纹理直接映射到三维模型上, 对侧面缺失部分插值, 从而提高了处理速度, 但由于部分照片的脸部被头发遮挡或混入背景颜色, 插值后会造成明显的错误。为解决此问题, 同时提高模型的真实感, 本文使用基于肤色模型的改进插值方法。

首先由式(8)将照片的正面人脸纹理映射到三维

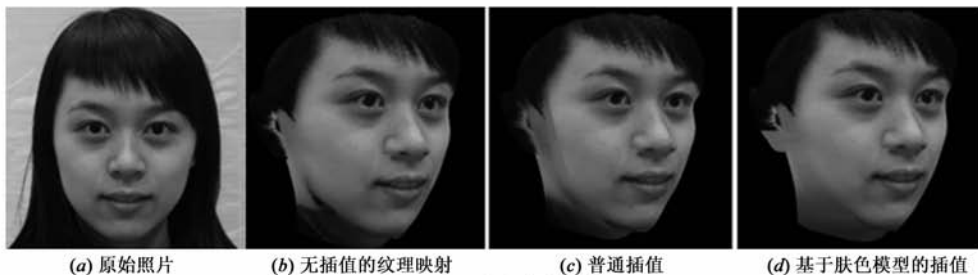


图5 纹理合成方法比较

### 3 三维人脸卡通化

人脸卡通是对人脸特征的夸张。人脸特征主要分为两类: 五官的全局结构和单独器官的局部形状。以眼睛为例, 增大眼睛轮廓属于局部形状变形而增加两眼间的距离则属于全局结构变形。这两种特征在人们对人脸辨别时都起到作用<sup>[16]</sup>, 艺术家创作的卡通人脸, 即是对这两种特征进行了不同程度的夸张, 从而使同一个人脸也能形成多种不同风格。已往的人脸自动卡通化方法中并没有区分这两种特征, 而是将它们视为一种特征同时处理, 尽管处理比较简便, 但是风格比较单

形状模型上。其中  $T_i$  是照片纹理;  $c$  和  $d$  分别是尺度因子和位移因子, 与式(4)中相同;  $T_m$  是映射生成的模型纹理。

$$T_m = c * T_i + d \quad (8)$$

然后使用基于肤色模型的插值对侧面的非肤色区域进行修正。公式(9)、(10)给出的肤色模型所用的颜色空间是修正的  $YC_bC_r$  空间, 详见参考文献[11]原文。若在  $YC_bC_r$  空间中的点通过式(9)处理后落入式(10)表示的椭圆区域中, 即判定为肤色点。

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} C_b - c_x \\ C_r - c_y \end{bmatrix} \quad (9)$$

$$\frac{(x - ec_x)^2}{a^2} + \frac{(y - ec_y)^2}{b^2} = 1 \quad (10)$$

逐点扫描人脸模型侧面, 若某像素  $P_e$  非肤色, 则用临域的肤色像素对其进行高斯滤波插值, 如式(11)、(12), 其中  $P_i$  是最接近  $P_e$  的  $N$  个像素之一;  $d_i$  是  $P_i$  与  $P_e$  之间的距离。

$$P_e(R, G, B) = \frac{\sum_{i=1}^N \omega_i P_i(R, G, B)}{\sum_{i=1}^N \omega_i} \quad (11)$$

$$\omega_i = \begin{cases} \frac{1}{2\pi} \exp(-\frac{d_i^2}{2}), & P_i \in \text{skin} \\ 0, & P_i \notin \text{skin} \end{cases} \quad (12)$$

基于肤色模型插值的效果与及其与普通插值的比较如图5, 可见其改善了侧面纹理映射错误的情况。

一. 本文提出了一种结合五官的全局结构和局部形状的自动夸张方法, 通过对全局和局部进行不同程度的夸张, 可由同一人脸生成多种卡通风格。由于前面步骤中获得的特定稀疏人脸模型包含了大部分卡通夸张所需的特征信息, 因此只需要对此稀疏模型进行处理以提高速度, 再使用完整模型变形即可。

对五官全局结构的夸张方法为, 计算  $S_p$  与  $S_a$  中各对应点的距离并将其放大  $b$  倍, 其中  $b$  为夸张系数。同时使用阈值处理以强调夸张效果, 即将对应点中的最大距离的差距的 20% 作为阈值  $T_x, T_y, T_z$ , 并忽略小于阈值的差距。其结果  $(X_g, Y_g, Z_g)$  可由式(13)~(15)获

得,其中 $(x_p, y_p, z_p) \in \mathcal{S}_p, (x_a, y_a, z_a) \in \mathcal{S}_a$ .

$$x_g = \begin{cases} x_a + b(x_p - x_a), & x_p - x_a > T_x \\ x_p, & x_p - x_a \leq T_x \end{cases} \quad (13)$$

$$y_g = \begin{cases} y_a + b(y_p - y_a), & y_p - y_a > T_y \\ y_p, & y_p - y_a \leq T_y \end{cases} \quad (14)$$

$$z_g = \begin{cases} z_a + b(z_p - z_a), & (z_p - z_a) > T_z \\ z_p, & (z_p - z_a) \leq T_z \end{cases} \quad (15)$$

对单独器官局部形状的夸张需要分别处理眉毛、眼睛、鼻子和嘴.仍以眼睛为例,先分别计算 $\mathcal{S}_a$ 和 $\mathcal{S}_p$ 中眼睛的几何中心,然后平移 $\mathcal{S}_a$ 的特征点使两个中心对齐,如图6.再由式(13)~(15)进行夸张处理并获得结果 $(X_l, Y_l, Z_l)$ ,其他器官的局部夸张与之类似.

对全局结构和局部形状分别处理后,再通过权重系数 $\omega$ 将它们合成为 $(X_c, Y_c, Z_c)$ ,即

$$(X_c, Y_c, Z_c) = (1 - \omega)(X_g, Y_g, Z_g) + \omega(X_l, Y_l, Z_l) \quad (16)$$

最后将第二章中得到的完整人脸模型通过TPS变

形,生成最终的三维卡通人脸.

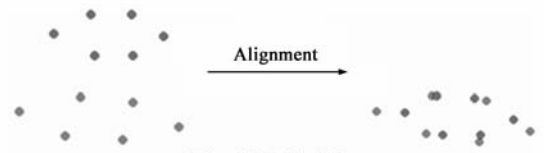


图6 眼部的对齐

## 4 实验结果

我们对各种人脸照片进行三维重建与卡通化实验,其结果如图7.其中第一列为原始输入照片,第二列为使用本文方法进行三维重建得到的三维人脸模型,第三列为强调全局特征夸张的卡通变形结果,第四列为强调局部特征夸张的卡通变形结果.可看到,对于不同性别和相貌的人脸均有较好的效果.由于本方法使用了稀疏人脸模型,因此计算速度较快,使用P4-2GHz处理器,对100张 $512 \times 512$ 大小的照片进行三维重建及卡通化,平均计算时间为3.9s,优于以往方法.

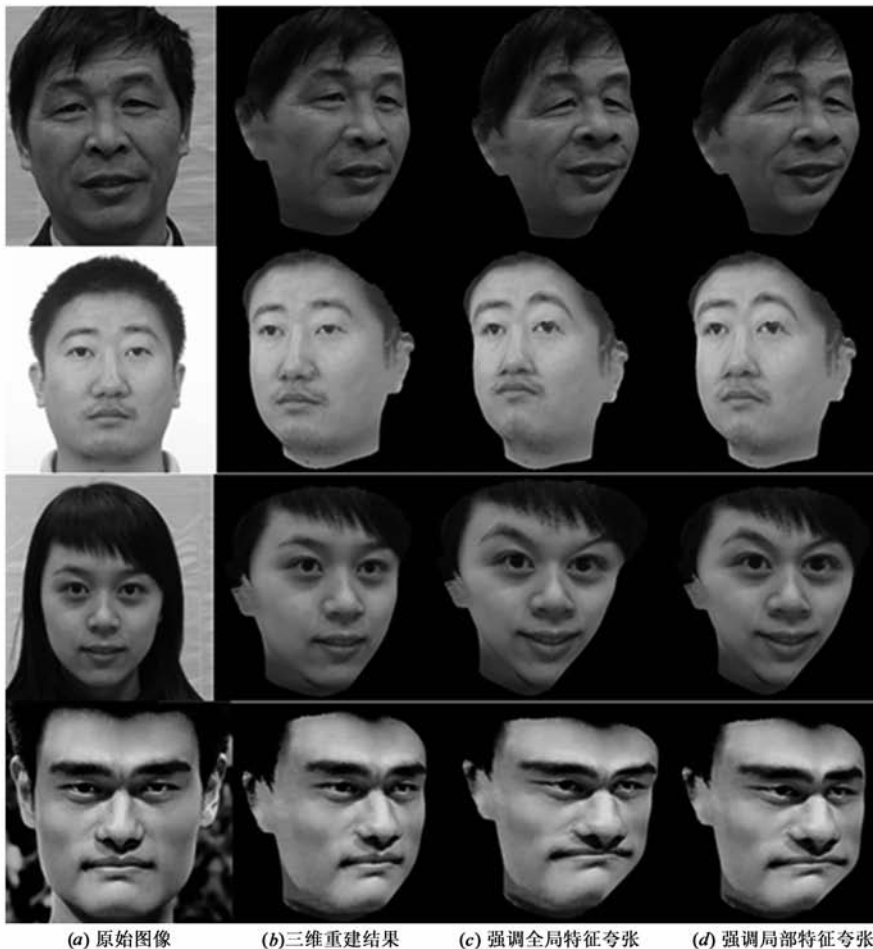


图7 自动卡通化结果

## 5 结论

本文提出了一种基于稀疏形变模型的三维卡通人

脸生成方法.首先使用稀疏形变模型拟合单幅照片获得特定的稀疏人脸;然后变形一个一般人脸模型,并基于肤色模型合成其纹理;最后将三维模型卡通化.

实验结果证实了我们的方法有如下优点:(1)三维人脸卡通比传统二维的效果更加丰富生动;(2)重建算法基于 SMM,有较快的处理速度和较低的储存空间占用,因此更适用于实时应用和移动通讯场合。(3)基于肤色模型的纹理映射提高了模型真实感;(4)对五官的全局和局部特征分别处理使卡通风格灵活多样。

未来的研究将会加强卡通人脸生成的智能化,即能够根据通讯中的语境或情感,自动产生相应风格的卡通人脸。

#### 参考文献:

- [1] Brennan S. Caricature Generator: The dynamic exaggeration of faces by computer[J]. Leonardo, 1985, 18 (3): 170 - 178.
- [2] Chen Hong, Liu Ziqiang, Rose Chunk, Xu Yingqing, Shum Heung-Yeung, Salesin D. Example-based composite sketching of human portraits[A]. Proceedings of the 3rd International Symposium on Non-photorealistic Animation and Rendering[C]. New York: ACM Press, 2004. 95 - 102.
- [3] Fujiwara T, Koshimizu H, Fujimura K, Kihara H, Noguchi Y, Ishikawa N. 3D Modelling system of human face and full 3D facial caricaturing[J]. IEIC Technical Report, 2001, 101(422): 27 - 32.
- [4] Akimoto T, Suenaga Yasuhito, Wallace R. Automatic creation of 3D face models[J]. IEEE Computer Graphics & Applications, 1993, 13(4): 16 - 22.
- [5] Blanz Volker, Vetter Thomas. A morphable model for the synthesis of 3D faces[A]. SIGGRAPH[C]. New York: ACM Press, 1999. 187 - 194.
- [6] Jiang Dalong, Hu Yuxiao, Yan Shuicheng, Zhang Lei, Zhang Hongjiang, Gao Wen. Efficient 3D reconstruction for face recognition[J]. Pattern Recognition, 2005, 38 (6): 787 - 798.
- [7] Vetter T, Jones M, Poggio T. A bootstrapping algorithm for learning linear models of object classes[A]. Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition[C]. New York: Computer Society Press, 1997. 40 - 47.

- [8] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[A]. Proceedings of the 2001 Conference on Computer Vision and Pattern Recognition[C]. New York: Computer Society Press, 2001. 511 - 518.
- [9] Du Chunhua, Yang Jie, Wu Qiang, Zhang Tianhao, Wang Huahua, Chen Lu, Wu Zheng. Extended fitting methods of active shape model for the location of facial feature points[A]. ICVGIP 2006 (Lecture Notes in Computer Science)[C]. Berlin: Springer, 2006. 610 - 618.
- [10] Bookstein F. Principle Warps: Thin-plate splines and the decomposition of deformation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989, 11(6): 567 - 585.
- [11] Hsu Rein-Lien, Abdel-Mottaleb M, Jain A. Face detection in color images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 696 - 706.
- [12] Benson P, Perrett D. Perception and recognition of photographic quality facial caricatures: Implications for the recognition of natural images[J]. European Journal of Cognitive Psychology, 1991, 3 (1): 105 - 135.

#### 作者简介:



署光男, 1983 年生于山东聊城, 现为上海交通大学硕士研究生, 研究方向为图像处理与模式识别。

E-mail: thesg2008@gmail.com



姚莉秀女, 1971 年生于浙江金华, 中科院微系统所博士, 现为上海交通大学副教授, 研究方向为数据挖掘与模式识别。

E-mail: lxyao@sjtu.edu.cn