

基于 NURBS 变形和单视角图片的人脸表情生成

孙 思^a, 葛卫民^a, 冯志勇^b, 徐 超^b, 彭伟龙^a

(天津大学 a. 计算机科学与技术学院; b. 软件学院, 天津 300350)

摘 要: 传统人脸表情重建方法大多通过线性组合的方式, 难以生成精细的人脸表情, 并且需要使用三维激光扫描仪得到图片, 不利于推广。为此, 提出一种基于非均匀有理 B 样条建立三维人脸表情模型的方法。根据人脸形状上与表情相关的特征点调节曲面的控制点, 引入几何约束和平滑约束保证整体和局部平滑。实验结果表明, 与线性方法相比, 该方法鲁棒性较好, 可以方便地从标准图片和无约束图片中生成较为精细的三维人脸表情。

关键词: B 样条; 形式化建模; 特征点检测; 单视角; 合成图片

中文引用格式: 孙 思, 葛卫民, 冯志勇, 等. 基于 NURBS 变形和单视角图片的人脸表情生成[J]. 计算机工程, 2017, 43(11): 227-233, 238.

英文引用格式: SUN Si, GE Weimin, FENG Zhiyong, et al. Facial Expression Generation Based on NURBS Warping and Single Perspective Images[J]. Computer Engineering, 2017, 43(11): 227-233, 238.

Facial Expression Generation Based on NURBS Warping and Single Perspective Images

SUN Si^a, GE Weimin^a, FENG Zhiyong^b, XU Chao^b, PENG Weilong^a

(a. School of Computer Science and Technology; b. School of Computer Software, Tianjin University, Tianjin 300350, China)

[Abstract] In facial expression generation, most of traditional methods are difficult to produce fine facial expression by linear combination, and the images need to be obtained with 3D laser scanner, which is not conducive to the promotion of the method. In order to solve this problem, a method Non-uniform Rational B-spline (NURBS) building 3D face expression model is proposed. In this method, the control points of the surface are adjusted according to the feature points related to facial expression, and geometric constraint and smoothness constrain are introduced to guarantee global and local smoothness. Experimental results show that compared with the linear method, this method is more robust, can easily generate fine 3D facial expression from standard image and unconstrained image.

[Key words] B-spline; formal modeling; feature point detection; single perspective; synthetic image

DOI: 10.3969/j.issn.1000-3428.2017.11.037

0 概述

在计算机视觉和计算机图形学领域中, 从二维人脸图片中重建三维人脸表情有着广泛的现实应用, 如人脸动画^[1-2]、三维人脸识别^[3-4]、三维表情识别^[5-6]和人脸表情建模^[7-8]等, 因此, 它成为当下一个热门的研究方向。

针对带表情的人脸生成, 研究者们提出了很多重建方法。文献[5]基于文献[9]中的三维通用弹性模型(3D Generic Elastic Model, 3D-GEM)提出了一个三维人脸表情通用弹性模型(3D Facial

Expression Generic Elastic Model, 3DFE-GEM), 从一张正脸人脸图片来重建一个带表情的三维人脸。该方法依据惊讶平均模型(Surprised Mean Model, SMM)、中性平均模型(Neutral Mean Model, NMM)和高兴平均模型(Happy Mean Model, HMM)的带权线性组合得到新模型, 由于其权重取决于距离比例因子, 因此, 重建结果受到很大的局限。文献[10]提出一种自适应的人脸规范模型技术, 利用 B 样条曲面来拟合三维激光扫描的人脸数据。文献[11]提出使用基于张量模型的多元统计方法的统计形状模型, 该模型根据三维扫描仪获得的训练人脸数据集

基金项目: 国家自然科学基金(61304262)。

作者简介: 孙 思(1989—), 男, 硕士研究生, 主研方向为计算机图形学、模式识别; 葛卫民, 副教授、博士; 冯志勇, 教授、博士; 徐 超, 副教授、博士; 彭伟龙, 博士。

收稿日期: 2016-10-17 **修回日期:** 2016-12-06 **E-mail:** sunsi@tju.edu.cn

在人脸不同部位进行形状统计并将其分为若干类,并建立点与点之间的对应关系。由于以上 2 种方法的输入图片都是通过三维激光扫描仪得到的,因此,这 2 种方法很难得到推广。

以上述研究为基础,本文提出使用非均匀有理 B 样条 (Non-uniform Rational B-spline, NURBS)^[12] 来表示人脸的方法。通过人脸特征点进行建模和引入约束,从图片中生成三维人脸表情。

1 本文方法

本文方法为一张带表情的输入图片生成与其表情相对应的三维人脸表情。首先通过同一个人的一些人脸图片基于 NURBS 曲面生成一个中性人脸作为三维人脸参考模型,然后对每一张带表情的人脸图片,通过构建其和参考模型之间的稀疏对应关系,并在几何约束和平滑约束下优化 NURBS 曲面参考模型的控制点,最后由优化后的控制点生成带表情的三维人脸形状。

1.1 NURBS 人脸建模

三维人脸表情形状的结构是复杂的,鼻尖部分是凸出的,嘴角部分是凹陷的,嘴唇的轮廓是凹凸不平的,从而造成了人脸表情形状表示和优化的困难。最近,文献[13]利用 B 样条曲面来重建三维中性人脸形状。但是,B 样条节点之间的间隔是均匀的,对人脸表情局部性的表达还不尽人意。那么,如何更好地表达人脸表情局部形状复杂度高的地方和表达人脸表情的多样性,是研究的重要方向。

考虑到 NURBS 可以定义非均匀的节点间隔,对于人脸局部形状复杂度较高的地方,让间隔小一点,而对于复杂度较低的地方,让间隔大一点,从而对人脸表情具有更精细结构表达的特性;同时,NURBS 控制点具有灵活性,可以表示人脸表情的多样性。因此,相比于文献[13]使用 B 样条曲面,本文提出使用 NURBS 曲面来表示三维人脸表情形状以获得更好的表示效果。

记人脸参考模型矩阵 S 是 $p \times q$ 的 NURBS 曲面,控制点记作矩阵 $B \{b_{ij}\}_{m \times n}$,节点向量记作 $U = \{u_i\}_{i=1}^{m+p}$, $V = \{v_j\}_{j=1}^{n+q}$,节点向量在 u 和 v 方向分割 uv 参数平面。所以,NURBS 曲面有如下等式:

$$S(u, v) = \sum_{i=0}^n \sum_{j=0}^m R_{i,j}(u, v) b_{i,j} \quad (1)$$

$$R_{i,j}(u, v) = \frac{N_{i,p}(u) N_{j,q}(v) w_{i,j}}{\sum_{i=0}^n \sum_{j=0}^m N_{i,p}(u) N_{j,q}(v) w_{i,j}} \quad (2)$$

样条函数如下:

$$\begin{cases} N_{i,1}(u) = \begin{cases} 1, u_i \leq u \leq u_{i+1} \\ 0, \text{otherwise} \end{cases} \\ N_{i,p}(u) = \frac{u - u_i}{u_{i+p-1} - u_i} N_{i,p-1}(u) + \frac{u_{i+p} - u}{u_{i+p} - u_{i+1}} N_{i+1,p-1}(u) \end{cases} \quad (3)$$

NURBS 节点间隔的不均匀性已经能够体现不同部位形状的不同精细程度,因此,将权重设为 $w = 1$ 。这样,给定控制点和节点向量,就确定了 NURBS 人脸曲面。

NURBS 曲面控制网格的每一个点都能够影响人脸的局部形状,因此,表达三维人脸表情形状有很大的灵活性。三维参考模型上的一个点 $S(u, v)$ 可以用如下的等式来计算得到:

$$S(u, v) = c_{u,v} \cdot [B_x | B_y | B_z] \quad (4)$$

其中,向量 $c_{u,v}$ 是由式(1)中 $b_{i,j}$ 的系数所得; B_x 、 B_y 、 B_z 分别是控制点矩阵 B 的列向量。特别地,向量 $c_{u,v}$ 是 $1 \times mn$ 维的向量, B_x 、 B_y 、 B_z 分别是 $mn \times 1$ 维的向量。

本文方法使用 NURBS 曲面替代文献[13]所提方法中的 B 样条,从二维人脸图片集中生成中性人脸作为参考模型,而对于每一张带表情的二维人脸图片通过 NURBS 人脸表情优化介绍可以生成与其表情相应的三维人脸形状。基于 NURBS 曲面生成的参考模型如图 1 所示。

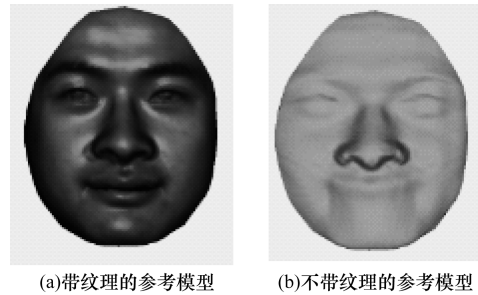


图 1 参考模型示意图

1.2 NURBS 人脸表情优化

给定参考模型,对于每一张带表情的二维人脸输入图片希望生成与其表情相对应的三维人脸表情形状。为此,利用二维人脸输入图片上的特征点重新估计三维人脸上的特征点,通过三维特征点的变化来调节参考模型上的控制点,并且提出了整体上的约束(几何约束)和局部上的约束(平滑约束),从而生成更平滑的带表情的三维人脸形状。

1.2.1 形式化建模

对于不同的人脸表情,其特征点位置也不尽相同,希望通过特征点的变化来刻画不同的表情,特征点位置如图2所示,其中实线所连接的特征点是用来刻画嘴巴的轮廓。通过二维人脸图片上的特征点和参考模型上的特征点构建它们之间的稀疏对应关系,得益于较为成熟的人脸特征点检测算法^[14-15],从人脸图片中可以提取较为准确的人脸特征点,从而可以建立这一稀疏对应关系。



图2 72个特征点的位置

根据NURBS曲面的仿射不变性,参考模型上特征点的位置变化可以用来指导NURBS曲面控制点的变化。因此,根据上面所描述的二维人脸图片和三维参考模型之间的稀疏对应关系,可以得到如下等式:

$$\min_{B_{\text{new}}} (\|A_p \times B_{\text{new}} - Q_{\text{new}}\| + \lambda \times T_g + \theta \times T_s) \quad (5)$$

其中, A_p 是 $72 \times mn$ 维的控制点系数矩阵,其每一行对应用于生成一个人脸特征点的NURBS曲面控制点系数,称 A_p 为特征点配置矩阵, B_{new} 是要优化的NURBS曲面控制点矩阵,其维度是 $mn \times 3$, 矩阵 Q_{new} 是输入图片对应的三维特征点集,假设二维人脸图片上的特征点集 q , 三维参考模型上的特征点集存在如下关系 $q = s_1 R Q + t$, s_1 , R 和 t 分别为尺度、旋转和平移参数。给定 q 和 Q , 依据文献[16]中所描述的方法对三维特征点位置进行重新估计,得到 Q_{new} ; 矩阵 T_g 和 T_s 是引入的约束项。

只从式(5)中第1项求解 B_{new} 时,可以发现方程组中方程的个数远小于未知数的个数,造成在非特征点处存在噪声,因此,增加第2项 T_g (二阶偏导约束项)来消除噪声,保证在调整过程中生成的三维人脸表情形状可以平稳变化。同时,由于所使用的特征点是很稀疏的,调节特征点和其周围区域时出现不均衡现象,造成在生成的三维人脸表情形状上的特征点处变

成奇异点,因此,增加第3项 T_s (带高斯权重的局部控制点平滑约束)来避免特征点处变成奇异点,达到在特征点周围区域局部更加平滑的效果。

1.2.2 几何约束

基于NURBS曲面的连续高阶的特性提出一个几何约束——二阶偏导约束项,让人脸表情整体在调整的过程中保持平稳变化。增加的几何约束如下:

$$T_g = \|A_2 \times B_{\text{new}} - A_2 \times B_{\text{old1}}\| \quad (6)$$

其中,矩阵 A_2 是 z 对 x, y 2个方向偏导数的和,其维度是 $t \times mn$, t 是NURBS曲面上点的个数,矩阵 B_{old1} 是前一次迭代结果的控制点。该方法通过增加一个几何约束,在整体上保证其要生成的三维人脸表情形状在调整过程中不会偏离参考模型太远,避免非特征点出现噪声,从而在整体上使得所生成的三维人脸表情形状更加平滑。

1.2.3 平滑约束

本文提出一个带高斯权重的局部控制点平滑约束,使特征点周围区域变得平滑。增加的平滑约束如下:

$$T_s = \|GM \times B_{\text{new}} - GM \times B_{\text{old2}}\| \quad (7)$$

其中, B_{old2} 是初始的三维参考模型的控制点矩阵, GM 是 $t \times mn$ 维的矩阵,其每一行存储是 mn 个 $\{c_{u,v}\}$ 乘以相应的权重 $G(u, v)$ 。 $G(u, v)$ 为:

$$G(u, v) = \begin{cases} \frac{1}{2\pi\sigma^2} e^{-\frac{(u-u_f)^2 + (v-v_f)^2}{2\sigma^2}}, & (u, v) \in [u_f - \frac{W}{2}, u_f + \frac{W}{2}) \\ & \times [v_f - \frac{W}{2}, v_f + \frac{W}{2}) \\ 1, & \text{otherwise} \end{cases} \quad (8)$$

其中, (u_f, v_f) 是特征点对应的 uv 参数, W 是窗口大小。该方法通过使用带高斯权重的局部控制点平滑约束调节特征点和其周围区域更加均衡,使得特征点和其周围区域变得平滑,避免特征点成为奇异点,从而从局部上保证所生成的三维人脸表情形状更加平滑。

因此,式(5)既能保证生成的三维人脸表情形状在迭代过程中平稳变化,又能保证在生成的三维人脸表情形状上特征点周围区域的平滑,分别从整体和局部保证了良好的三维人脸表情形状。如上所述,可以通过式(5)生成和输入的二维人脸图片中的表情相对应的三维人脸表情形状。

1.2.4 嘴部轮廓的拟合

对于不同的表情,嘴部的轮廓也不同,造成嘴部的轮廓是多变的。因此,基于 NURBS 曲线局部可调节的特性在迭代过程中使用 NURBS 曲线来拟合嘴部的轮廓,将其形式化为:

$$\min_{B_{\text{new_mouth}}} \| A_{\text{p_mouth}} \times B_{\text{new_mouth}} - Q_{\text{new_mouth}} \| \quad (9)$$

其中, $A_{\text{p_mouth}}$ 是 $16 \times l$ 维的 NURBS 曲线控制点系数矩阵,其每一行对应用于生成一个嘴部特征点的 NURBS 曲线控制点系数; l 是 NURBS 曲线控制点的个数;在参考模型上已计算出三维嘴部特征点,进而可以确定 $A_{\text{p_mouth}}$; $B_{\text{new_mouth}}$ 是要优化的 NURBS 曲线控制点矩阵,其维度是 $l \times 3$; 矩阵 $Q_{\text{new_mouth}}$ 是输入图片对应的三维嘴部特征点集,从 1.2.1 节中求解出的 Q_{new} 中抽取出来得到。图 3 中实线表示嘴部的轮廓,带圆点的实线上的点表示控制点,带圆点的实线将控制点连接起来表示控制线。

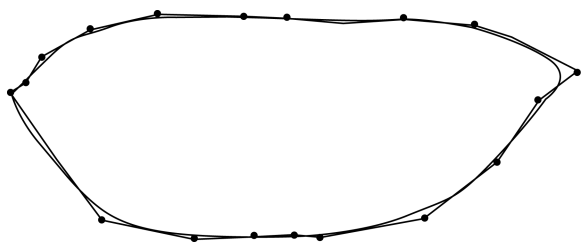


图 3 NURBS 曲线刻画嘴部的示意图

1.3 人脸表情生成方法

给定输入二维人脸图片集和迭代次数,本文的整个方法流程如下:1)对二维人脸图片集中的图片进行预处理,从每一张图片中提取二维人脸特征点集;2)基于文献[13]中所提方法使用 NURBS 曲面从二维人脸图片集中生成中性人脸作为参考模型;3)进行迭代,首先重新估计三维人脸特征点的位置,矩阵 Q_{new} 和矩阵 $Q_{\text{new_mouth}}$,然后计算式(5)和式(9),将优化出来的矩阵 B_{new} 赋值给矩阵 B_{old} ,迭代这一过程到迭代次数停止;4)根据上面的矩阵 B_{new} 计算新的人脸形状,并且对嘴部的轮廓进行拟合;最后,输出和输入图片中的表情相对应的三维人脸表情形状。

方法 1 基于 NURBS 曲面和单视角图片的三维人脸表情生成方法

输入 人脸图片集,迭代次数 $IterNum$

输出 带表情的三维人脸形状

步骤 1 预处理,在人脸图片上提取二维人脸特征点集。

步骤 2 生成参考模型,基于人脸图片集生成 NURBS 中性人脸曲面,并计算三维人脸特征点集矩阵 Q 。

步骤 3 重新估计三维特征点集位置,对于单张图片重新估计矩阵 Q_{new} ,矩阵 $Q_{\text{new_mouth}}$

步骤 4 计算式(5),其解是:

$$B_{\text{new}} = pinv(A_p^T \times A_p + \lambda \times A_2^T \times A_2 + \theta \times GM^T \times GM) \times (A_p^T \times Q_{\text{new}} + \lambda \times A_2^T \times A_2 \times B_{\text{old}} + \theta \times GM^T \times GM \times B_{\text{old}})$$

计算式(9),其解是:

$$B_{\text{new_mouth}} = pinv(A_{\text{p_mouth}}^T \times A_{\text{p_mouth}}) \times (A_{\text{p_mouth}}^T \times Q_{\text{new_mouth}})$$

重复步骤 3 和步骤 4,使得赋值 $B_{\text{old}} = B_{\text{new}}$ 达到迭代次数 $IterNum$ 。

步骤 5 由矩阵 B_{new} 计算新的人脸形状,并对嘴部进行拟合。

2 实验结果与分析

本文设计了 2 个实验来验证所提方法。为了验证 NURBS 曲面形变可以表示带表情的三维人脸表情形状,设计了第 1 个实验,从合成图片中生成三维人脸表情;为了验证本文所提基于 NURBS 形变和单视角图片生成三维人脸表情的方法具有泛化性,设计了第 2 个实验,从 LFW (Labeled Faces in the Wild) 数据集里的无约束图片中生成三维人脸表情。

2.1 基于合成图片的表情生成

本文实验评价所提方法从合成图片中生成三维人脸表情形状的效果,从而验证在多角度单一表情数据上的稳定性和在标准数据上的有效性。实验结果如图 4、图 5 所示。本文第 1 组实验是从不同姿态和光照条件下的合成图片中生成三维人脸表情,第 2 组是从不同表情下的合成图片中生成三维人脸表情。该实验所使用的图片是合成图片,截取自文献[13]中的人脸建模结果。在不同姿态和光照条件下的图片数据集作为该实验第 1 组的输入图片,如图 4(a)所示,前 4 个表示不同的姿态,后 2 个表示不同的光照。在不同表情下的图片数据集作为该实验第 2 组的输入图片,如图 5(a)所示,表示不同的 6 种表情,即恶心、高兴、惊讶、生气、悲伤和恐惧。利用人脸特征点检测算法,在每张图片上提取特征点,通过迭代式(5)得到实验 1 结果。在迭代过程中,参数 λ, θ 设置为 $\lambda = 0.000\ 001, \theta = 1$, 高斯函数方差是 $\delta = 2.022\ 973$,窗口的大小是 $W = 7$ 。

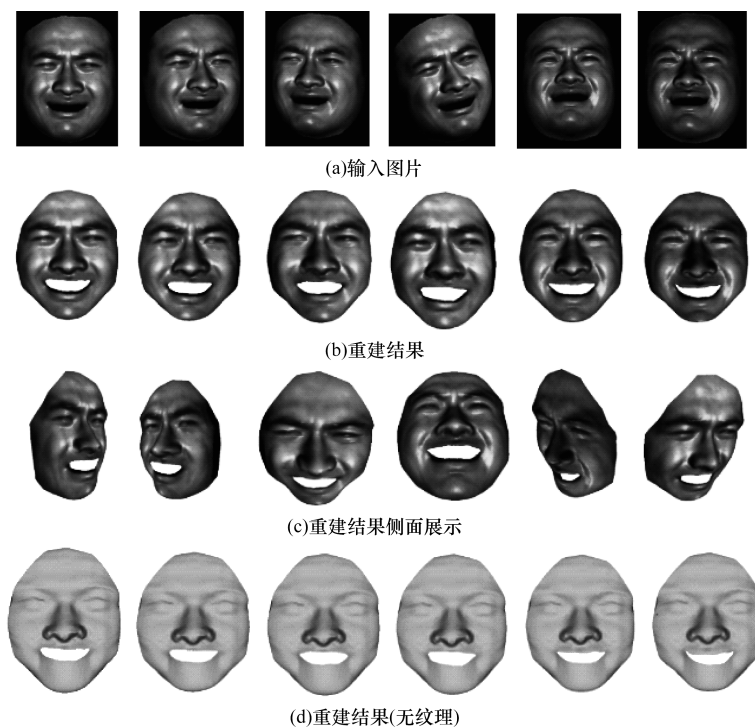


图 4 一个表情在不同姿态和光照条件下生成的结果示意图

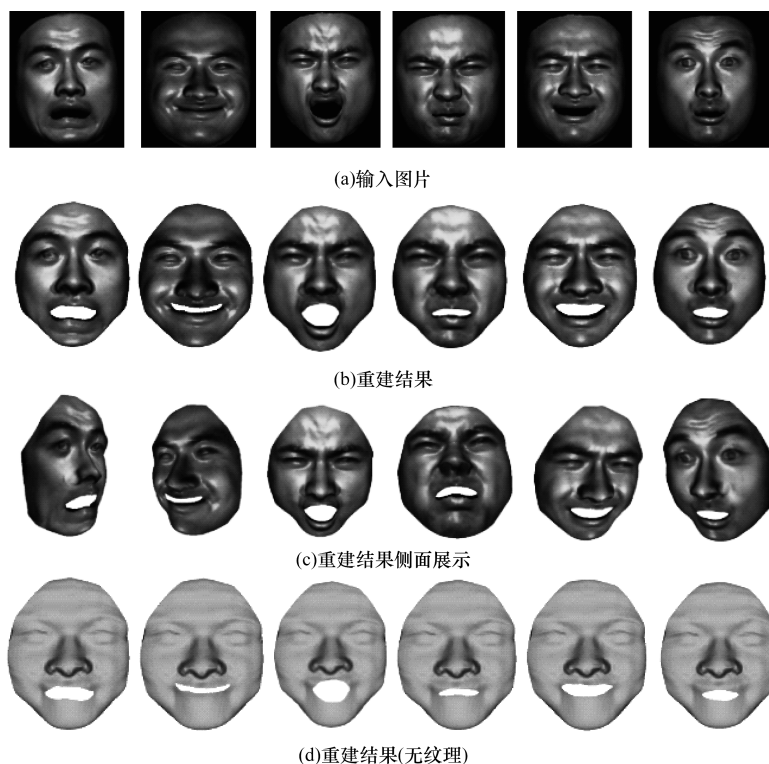


图 5 6 种表情生成的示意图

在第 1 组实验中,针对不同的姿态同一表情重建的结果几乎一样,证明了本方法对单一表情重建的稳定性。从图 4 中的第 6 列可以发现,其嘴角部分重建的不是很好,然而这并不影响人脸表情效果的表达。在第 2 组实验中,针对 6 种表情图片生成的结

果分别能良好地表达相应的表情形状,证明了本文方法在标准数据上的有效性。从图 5(c)发现,重建的结果在眼睛周围并没有重建出很细致的变化,这是由于特征点检测算法对带有表情眼部检测不准确,因此在眼部的重建结果无差异。在图 5(b)和

图 5(d)中发现不管是从正面还是从不同角度来看,每一个表情形状都有不同的嘴型,这表明本文所提的方法可以用 NURBS 曲线刻画不同表情的嘴部轮廓。

从图 1(b)和图 4(d)和图 5(d)对比可以发现,本文所使用的 NURBS 在表情生成过程中可以较好地表达三维人脸模型。该实验证明了本文所提的方法可以从不同的姿态、光照和表情条件下的合成图片中生成三维人脸表情形状。

2.2 比较与定量评价

由于可视结果重点突出嘴部的变化,但不能明显体现脸部其他部分重建的精准度。因此,给出了本文方法和线性方法^[17-18]的量化对比。对于文献[17]的方法,使用 3DMM 拟合中性人脸,然后通过单视角图片上的二维人脸特征点来调节中性人脸;对于文献[18]方法,利用视频序列中某一表情的前后若干帧来恢复三维人脸特征点,为了公平性,用其恢复出的三维人脸特征点来调节本文中的中性人脸。文献[18]方法所使用的视频序列是根据文献[13]的生成结果制作的视频序列,包含了图 5 中的 6 个表情。图 6 中展示了对悲伤表情重建的例子。

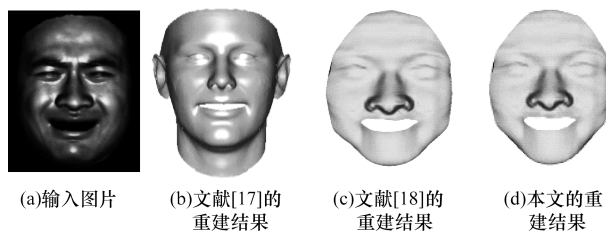


图 6 对比结果示意图

表 1 中的表情(1,2,...,6)对应着图 5 中的 6 个的表情,图片编号(1,2,...,12)对应着图 4 和图 5 的输入图片。对于定量评价,给出 2 个评价指标:平均欧式距离和均方根距离。首先对生成的人脸表情形状上的每一个点计算出到 ground truth 上最近的点的距离,然后对每一个距离除以两眼之间的距离进行标准化,最后给出生成结果和 ground truth 之间的平均欧式距离和均方根距离,如表 1 所示。从表 1 中发现,本文所提方法的平均欧式距离的平均值为 0.089 1 低于文献[17]的平均值 0.184 9 和文献[18]的平均值 0.095 9;均方根距离的平均值为 0.124 4 低于文献[17]的平均值 0.224 0 和文献[18]的平均值 0.132 8。

表 1 生成结果和 ground truth 之间的距离

表情	图片	平均欧式距离			均方根距离		
		文献[17] 方法	文献[18] 方法	本文方法	文献[17] 方法	文献[18] 方法	本文方法
表情 5	图片 1	0.182 2	0.095 7	0.089 3	0.220 6	0.122 5	0.124 3
	图片 2	0.176 1	0.088 1	0.087 9	0.216 6	0.142 3	0.122 2
	图片 3	0.181 4	0.093 8	0.089 0	0.221 8	0.125 5	0.124 0
	图片 4	0.170 9	0.092 7	0.088 6	0.209 4	0.128 1	0.123 8
	图片 5	0.184 9	0.097 6	0.088 9	0.225 5	0.128 7	0.124 5
	图片 6	0.188 3	0.088 3	0.089 1	0.232 7	0.118 5	0.126 2
表情 1	图片 7	0.193 9	0.094 0	0.089 7	0.229 8	0.122 8	0.125 8
表情 2	图片 8	0.187 3	0.109 2	0.087 0	0.228 7	0.166 2	0.121 4
表情 3	图片 9	0.193 2	0.097 6	0.092 4	0.227 4	0.127 0	0.127 5
表情 4	图片 10	0.181 6	0.110 3	0.089 0	0.219 6	0.166 4	0.124 5
表情 5	图片 11	0.182 2	0.095 7	0.089 3	0.220 6	0.122 5	0.124 3
表情 6	图片 12	0.196 8	0.088 1	0.088 6	0.235 8	0.122 6	0.123 8
平均值		0.184 9	0.095 9	0.089 1	0.224 0	0.132 8	0.124 4

文献[17]使用的方法是基于统计模型的,重建出的测试样本不一定在其形状基和表情基构成的空间中;文献[18]方法和本文方法结果近似,但是如果视频序列中的表情剧烈变化导致生成的表情不是那一时刻的表情从而造成其稳定性较差;提出的方法通过二维特征点条件 NURBS 的变化,也具有较好的稳定性。对比结果表明,本文方法生成的三维人脸

表情结果更接近 ground truth。

2.3 基于无约束图片的表情生成

为了评价所提方法从无约束图片中生成三维人脸表情的效果,也为了证明所提方法在无约束数据上的泛化性,因此,本文实验在 LFW 数据集上测试本文所提的方法。

在这个实验中,基于 NURBS 改进文献[13]方

法为 LFW 数据集生成 NURBS 中性人脸作为参考模型,用本文方法为带有表情的二维人脸图片生成带表情的三维人脸形状。

生成结果如图 7 所示。从图 7 中发现,对应每

一个输入二维图片都生成了一个三维人脸表情形状,同时,嘴部的轮廓也得到了刻画。从而,本文方法也可以从无约束图片中生成带表情的三维人脸形状,验证了本文方法的泛化性。

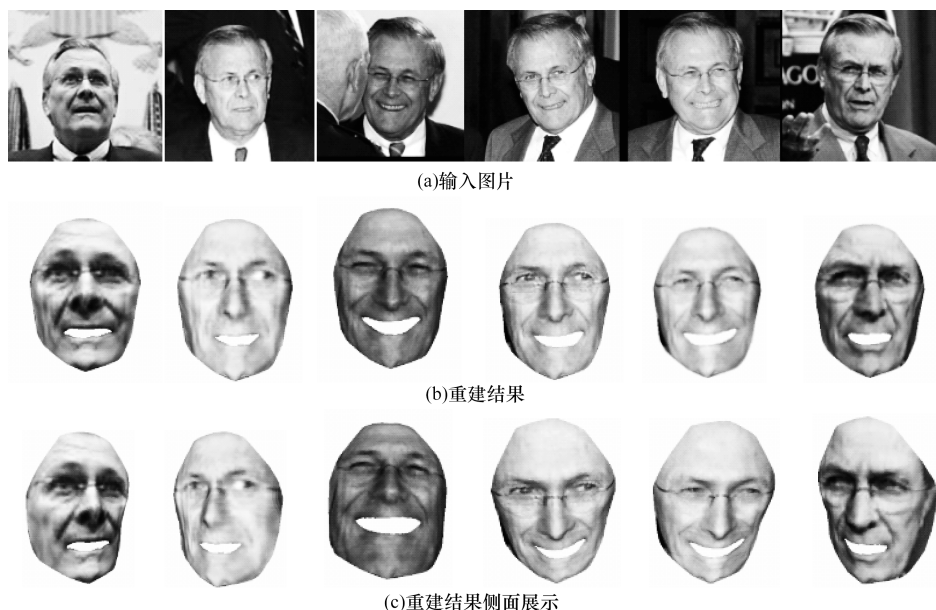


图 7 从无约束图片生成人脸表情的示意图

3 结束语

本文提出一种基于 NURBS 变形的人脸表情生成方法,为单张带表情的人脸图片生成三维人脸表情形状。根据几何约束和平滑约束从整体和局部上保证生成的 NURBS 人脸表情趋于真实。实验结果表明,该方法在多角度单一表情数据上具有较好的稳定性,同时证明了其在标准数据上的有效性和在无约束数据上的泛化性。下一步将利用 NURBS 人脸建模对三维人脸辅助识别、三维人脸表情辅助识别进行研究。

参考文献

- [1] MOIZA G, TAL A, SHIMSHONI I, et al. Image-based Animation of Facial Expressions[J]. Visual Computer, 2002, 18(7): 445-467.
- [2] LUO Changwei, JIANG Chen, YU Jun, et al. Expressive Facial Animation from Videos [C]//Proceedings of IEEE International Conference on Image Processing. Washington D. C. , USA: IEEE Press, 2014: 4617-4621.
- [3] KITTLER J, HILTON A, HAMOUZ M, et al. 3D Assisted Face Recognition: A Survey of 3D Imaging, Modelling and Recognition Approaches [C]//Proceedings of IEEE Computer Society Conference on Computer Vision & Pattern Recognition-workshops. Washington D. C. , USA: IEEE Press, 2005: 114-114.
- [4] BLANZ V, VETTER T. Face Recognition Based on Fitting a 3D Morphable Model[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2003, 25(9): 1063-1074.
- [5] MOEINI A, MOEINI H, FAEZ K. Pose-Invariant Facial Expression Recognition Based on 3D Face Reconstruction and Synthesis from a Single 2D Image[C]//Proceedings of International Conference on Pattern Recognition. Washington D. C. , USA: IEEE Computer Society, 2014: 1746-1751.
- [6] LI Huibin, DING Huaxiong, HUANG Di, et al. An Efficient Multimodal 2D + 3D Feature-based Approach to Automatic Facial Expression Recognition [J]. Computer Vision & Image Understanding, 2015, 140(C): 83-92.
- [7] CHI Jing, TU Changhe, ZHANG Caiming. Dynamic 3D Facial Expression Modeling Using Laplacian Smooth and Multi-scale Mesh Matching[J]. Visual Computer, 2014, 30(6): 649-659.
- [8] DAHMANE M, MEUNIER J. Prototype-based Modeling for Facial Expression Analysis[J]. IEEE Transactions on Multimedia, 2014, 16(6): 1574-1584.
- [9] HEO J, SAVVIDES M. 3-D Generic Elastic Models for Fast and Texture Preserving 2-D Novel Pose Synthesis[J]. IEEE Transactions on Information Forensics & Security, 2012, 7(2): 563-576.
- [10] HOCH M, FLEISCHMANN G, GIROD B. Modeling and Animation of Facial Expressions Based on B-Splines[J]. The Visual Computer, 1994, 11(2): 87-95.
- [11] MINOJ J L, JUPIT A J R, GILLIES D F, et al. Facial Expressions Reconstruction of 3D Faces Based on Real Human Data [C]//Proceedings of IEEE International Conference on Computational Intelligence and Cybernetics. Washington D. C. , USA: IEEE Press, 2012: 185-189.

(下转第 238 页)

了算法的有效性。由于 CNN 特征只使用了图像 RGB 数据,今后将进一步融入动作特征,从而获得更高的性能提升。

参考文献

- [1] WANG Heng, SCHMID C. Action Recognition with Improved Trajectories [C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2013: 3551-3558.
- [2] 程海粟, 李庆武, 仇春春, 等. 基于改进密集轨迹的人体行为识别算法 [J]. 计算机工程, 2016, 42 (8): 199-205.
- [3] SIMONYAN K, ZISSERMAN A. Two-stream Convolutional Networks for Action Recognition in Videos [C]//Proceedings of Advances in Neural Information Processing Systems. Montreal, Canada: MIT Press, 2014: 568-576.
- [4] SHARIF R A, AZIZPOUR H, SULLIVAN J, et al. CNN Features Off-the-shelf: An Astounding Baseline for Recognition [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops. Columbus, USA: IEEE Press, 2014: 806-813.
- [5] 尹宝才, 王文通, 王立春. 深度学习研究综述 [J]. 北京工业大学学报, 2015 (1): 48-59.
- [6] SHARMA S, KIROS R, SALAKHUTDINOV R. Action Recognition Using Visual Attention [EB/OL]. (2016-02-14). <https://arxiv.org/pdf/1511.04119.pdf>.
- [7] ZHA S, LUISIER F, ANDREWS W, et al. Exploiting Image-trained CNN Architectures for Unconstrained Video Classification [C]//Proceedings of BMVC'05. Swansea, UK: BMVA Press, 2015: 1-13.
- [8] SRIVASTAVA N, MANSIMOV E, SALAKHUTDINOV R. Unsupervised Learning of Video Representations Using LSTMs [C]//Proceedings of International Conference on Machine Learning (ICML). Lille, France: Microtome Publishing, 2015: 843-852.
- [9] DENG Jia, DONG Wei, SOCHER R, et al. Imagenet: A Large-scale Hierarchical Image Database [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2009: 248-255.
- [10] KUEHNE H, JHUANG H, GARROTE E, et al. HMDB: A Large Video Database for Human Motion Recognition [C]//Proceedings of 2011 IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2011: 2556-2563.
- [11] SOOMRO K, ZAMIR A R, SHAH M. UCF101: A Dataset of 101 Human Actions Classes from Videos in the Wild [EB/OL]. (2012-12-03). <https://arxiv.org/pdf/1212.0402.pdf>.
- [12] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet Classification with Deep Convolutional Neural Networks [C]//Proceedings of Advances in Neural Information Processing Systems. Montreal, Canada: MIT Press, 2012: 1097-1105.
- [13] SZEGEDY C, LIU W, JIA Y, et al. Going Deeper with Convolutions [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2015: 1-9.
- [14] ZHANG Xiangyu, ZOU Jianhua, HE Kaiming, et al. Accelerating very Deep Convolutional Networks for Classification and Detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38 (10): 1943-1955.
- [15] MAHENDRAN A, VEDALDI A. Understanding Deep Image Representations by Inverting Them [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2015: 5188-5196.
- [16] KEMELMACHER-SHLIZERMAN I, SEITZ S M. Face Reconstruction in the Wild [C]//Proceedings of International Conference on Computer Vision. IEEE Computer Society, 2011: 1746-1753.
- [17] ZHU Xiangyu, LEI Zhen, YAN Junjie, et al. High-fidelity Pose and Expression Normalization for Face Recognition in the Wild [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2015: 787-796.
- [18] VAN Dam C, VELDHUIS R, SPREEUWERS L. Landmark-based model-free 3D Face Shape Reconstruction from Video Sequences [C]//Proceedings of International Conference on Biometrics Special Interest Group (BIOSIG). Washington D. C., USA: IEEE Press, 2013: 1-5.

编辑 刘 冰

编辑 刘 冰

(上接第 233 页)

- [12] PIEGL L, TILLER W. The NURBS Book [M]. 2nd ed. Berlin, Germany: Springer-Verlag, 1997.
- [13] PENG Weilong, XU Chao, FENG Zhiyong. 3D Face Modeling Based on Structure Optimization and Surface Reconstruction with B-Spline [J]. Neurocomputing, 2016, 179 (12): 228-237.
- [14] ASTHANA A, ZAFEIRIOU S, CHENG S, et al. Robust Discriminative Response Map Fitting with Constrained Local Models [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2013: 3444-3451.
- [15] TZIMIROPOULOS G, PANTIC M. Gauss-Newton Deformable Part Models for Face Alignment In-the-Wild [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2014: 1851-1858.