

单视影像下的人脸快速三维重建

袁 理^{1,2} 陈庆虎¹ 廖海斌¹ 段柳云¹

(1 武汉大学电子信息学院,武汉市珞喻路 129 号,430079)
(2 武汉职业技术学院电子信息学院,武汉市关山一路 463 号,430074)

摘 要:提出了一种单视影像下的人脸快速三维重建方案,能够对非可控条件下获取的单幅人脸图像进行有效、快速的三维重建。在三维人脸重建阶段,提出了一种特征动态选择策略,针对待重建人脸图像动态选择合适的 2D-3D 特征变换子空间,在保证执行效率的同时能显著提高三维人脸模型的重建精度。实验证明,本方法不仅简洁高效,而且对人脸姿态和光照变化具较好的鲁棒性,在 WHU-3DFaceData 数据库上取得了理想的重建效果。

关键词:三维人脸重建;子集划分;ASMS;动态选择策略

中图法分类号:P237.3

人类视觉系统具备从单张二维图像恢复其三维信息的能力。利用计算机模拟人类这一图像认知过程,研究单幅人脸图像的快速三维重建技术,是当今计算机图形学、计算机视觉、人工智能等领域中极具挑战性的问题之一^[1-3]。

一般而言,在没有任何约束的情况下,基于单张二维图像的三维重建是一个病态问题^[4-6]。Blaiz 和 Vetter 等^[7-9]以人脸形状和纹理统计信息为约束条件,利用三维形变模型(3D morphable model,3DMM)成功地实现了基于单张人脸图像的三维重建。但该模型匹配算法对初值依赖性较强,往往会因为局部最优解而导致建模失败。针对此问题,Romdhani 等^[10]提出采用多重特征对优化目标函数进行约束,在一定程度上缓解了 3DMM 局部最优解的问题。胡永利等^[5]首次建立了面向中国人特征的大规模三维人脸数据库(BJUT-3D Face Database),并以此为基础通过网络重采样等方法对 3DMM 进行改进,实现了面向中国人特征的三维人脸重建。在此基础上,龚勋等^[6]以统计学为工具,利用人脸先验知识对整个三维重建过程进行约束,实现了基于少量特征点的三维人脸重建,取得了较为理想的重建效果。但目前基于 3DMM 的重建技术还存在参数匹配优化较为复杂、迭代次数过多、计算量大等问题。

针对这些问题,Mena 等人^[11]提出了一种较

为高效、可靠的三维人脸重建方案,能够对具有表情的单张人脸图像进行三维重建。本文以 Mena 等人的研究成果为基础,重点探索如何对非可控条件下获取的单幅人脸图像进行快速三维重建,以满足实际应用的需要。

1 数据获取及训练阶段

1.1 三维人脸数据的获取及预处理

三维人脸数据是由 3DCaMega 结构光三维扫描系统来获取的。对于获取的三维人脸数据通过 CLOUDFORM 软件^[12]进行裁剪、校正、补洞和降噪等处理。同时,利用文献[6]中的方法对人脸关键特征点进行自动定位,在提高定位效率的同时能有效避免人为误差的引入。经过定位后的训练样本包含 62 个固定编号的特征点,从而最大限度地保留人脸信息,如图 1 所示。

1.2 二维虚拟图像的获取

利用球面谐波理论^[13]构造光照模型,通过对三维人脸模型进行二次处理获取大量具有不同姿态和光照条件的二维虚拟图像,从而提高样本的完备性。由于 2D-3D 特征点存在对应关系,因此在生成的所有二维虚拟图像中其特征点依然得以保留,生成的部分二维虚拟图像如图 2 所示。

其中,第一行是在光线方位角由 $[-45^{\circ}, 45^{\circ}]$

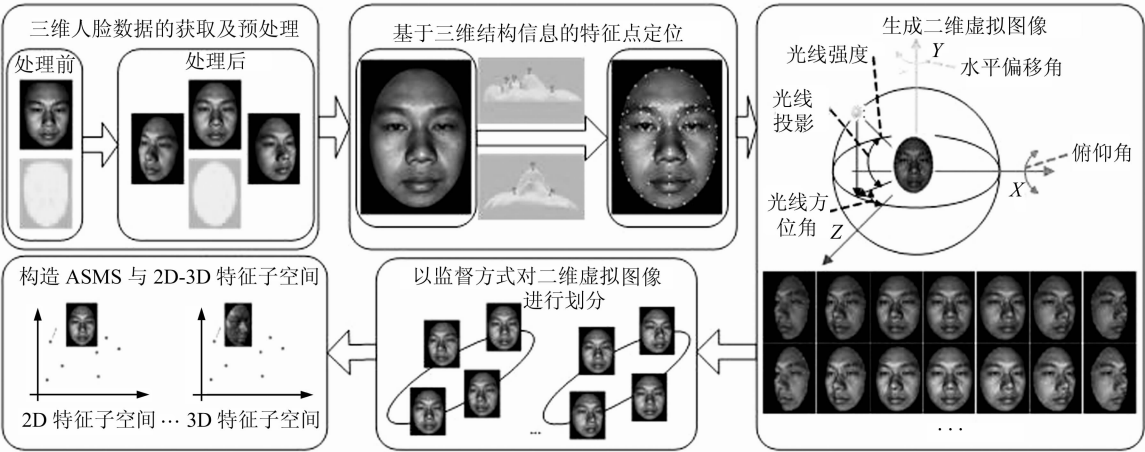


图 1 数据获取及训练阶段

Fig. 1 Data Acquisition and Training Phase

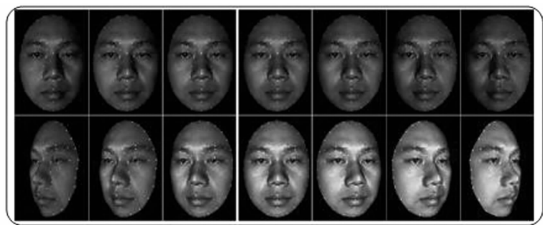


图 2 部分二维虚拟图像

Fig. 2 Part of 2D Virtual Images

变化下生成的二维虚拟图像集,角度变化增量 $\Delta\alpha=15^\circ$,光线强度为 1.0,而其他参量保持不变;第二行则是在水平偏移角和光照强度两个参量共同控制下生成的二维虚拟图像集,其中三维人脸模型的水平偏移角变化范围为 $[-45^\circ, 45^\circ]$,角度变化增量 $\Delta\alpha=15^\circ$;光照强度变化范围为 $[1.25, 2.75]$,强度增量为 0.25,而其他参量保持不变。

1.3 以监督方式对二维虚拟图像进行子集划分

当人脸姿态和光照条件发生变化时,人脸的有效特征很难在线性空间中准确表示,而非线性特征的提取过程又相对复杂。不同于设计复杂的聚类算法,本文提出了根据虚拟图像的控制参量,将获得的二维虚拟图像划分为与姿态和光照有关的 M 个子集,通过提高训练样本的聚类性从而有效地缓解人脸特征提取过程中的非线性问题。某一特定对象的子集数量 M 可以表示为:

$$M = \left\lceil \frac{nP_0}{d} \right\rceil \times \left\lceil \frac{nP_v}{d} \right\rceil \times \left\lceil \frac{nL_0}{d} \right\rceil \times \left\lceil \frac{nL_v}{d} \right\rceil \times \left\lceil \frac{nE_0}{d} \right\rceil \tag{1}$$

式中,参数 nP_0 、 nP_v 、 nL_0 、 nL_v 和 nE_0 分别表示人脸姿态和光照等控制参量的变化范围;而参数 d 则表示在每一个子集中各个控制参量的变化数量。因此,二维虚拟图像构成的子集可记作 S_p^i ,其中 $p=1,2,\cdots,M$;而每一个子集 S_p^i 中包含的

二维虚拟图像数 $N=d^5 \times R$ 幅,其中 R 表示训练集中的对象数。

1.4 建立子集主动形状模型 ASMS

本文利用 ASMS 对非可控条件下获取的人脸图像进行特征点定位。由于在建立模型的过程中提高了训练样本的完备性与聚类性,因此,模型定位能力显著提高,对姿态和光照具有理想的鲁棒性。由子集划分规则可知,每个子集 S_p^i 中的训练样本具有类似的姿态、光照条件以及点分布模型(point distribution model, PDM),本文采用 Procrustes 分析对子集中的训练样本进行对齐处理,通过最小化 $\sum_{i=1}^N |x_i - \bar{x}|^2$ 让训练数据处于同一可比坐标系中,其中 \bar{x} 为子集中所有训练样本的平均形状。然后分别获取训练样本的形状模型与灰度模型并最终建立子集主动形状模型 $ASMS_p$,其中 $p=1,2,\cdots,M$ 。

1.5 建立 2D-3D 特征子空间

定义 1 LFA 的核是由所有元素(对于二维纹理数据而言,就是指每个像素;相对于三维结构数据而言,就是指每个顶点)的子核组合得到的:

$$K = (k_1, \cdots, k_j, \cdots, k_n) \in R^{(dn) \times (dn)} \tag{2}$$

式中, n 是元素数目, d 是训练样本的维数,对于二维纹理数据 $d=2$,而三维结构数据 $d=3$, k_j ($1 \leq j \leq n$) 是第 j 元素的核,定义为特征矩阵 Q 中第 j 个元素所对应的 d 行在 Q 上的投影:

$$K_j = Q\Omega(Q_{dj(l)})^T \in R^{(dn) \times d} \tag{3}$$

式中, $\Omega = \text{diag}(1/\sigma_i) \in R^{(dn) \times d}$ 是归一化因子, $1 \leq i \leq m'$; $\sigma = (\sigma_1^2, \cdots, \sigma_{m'}^2)$ 由特征值构成的对角矩阵; $Q_{dj(l)} \in R^{l \times m'}$ 表示从特征矩阵 Q 中第 dj 行开始连续取 l 行($l=2,3$),构成第 j 元素对应的特征矩阵。同时,为了用于三维重建,第 j 元素的 LFA 逆核定义为:

$$K_j^{-1} = Q\Omega^{-1} (Q_{dj(l)})^T \in R^{(dn) \times d} \quad (4)$$

由 LFA 核 K 的定义可以看出, K_j 能够保留顶点 j 及其周围点的信息。从滤波的观点来看, LFA 为每个特征点均提供一个特定的滤波器。各滤波器具备自己独特的形状, 且形状与相邻特征之间的相关性相适应。

利用 LFA 对不同子集中的 2D 纹理数据与 3D 结构数据进行特征提取, 并利用 PCA 对其进行降维。为了获得相同维数的特征空间, 本文分别选择累积贡献率为 95% 的前 k 个主成分建立 2D-3D 特征子空间。

2 三维人脸重建阶段

根据子集划分原则可知, 由不同子集所建立的特征子空间对同一对象的表示能力是不同的。因此, 为了提高三维重建的精度, 本文通过设计特征动态选择策略, 为不同的待重建人脸图像选择合适的 2D-3D 特征变换子空间, 具体步骤如下。

1) 利用 ASMS_p 依次搜索图像中每个特征点的邻域, 在邻域内找到与该点灰度模型最佳匹配点的位置并得到新的形状; 根据特征点的位移计算 PDM 模型中的姿态和形状参数的变化; 更新姿态参数和形状参数, 重新搜索特征点直至收敛。

2) 计算人脸特征点平均误差 E , 平均误差越小说明匹配结果越理想:

$$E(p) = \frac{1}{N} \sum_{i=1}^N [\frac{1}{n} \sum_{j=1}^n \text{dis}(D(p)_{ij}, D(p)'_{ij})] \quad (5)$$

式中, N 是每一个子集中训练样本的总数; n 是 PDM 中的点数; $D(p)_{ij}$ 是子集中第 i 个训练样本的第 j 个特征点; $D(p)'_{ij}$ 是利用 ASMS_p 定位得到的特征点; $\text{dis}(D(p)_{ij}, D(p)'_{ij})$ 是两点之间的欧氏距离。

3) 通过计算 $\arg \min_{p=1,2,\dots,M} E(p)$ 确定子集索引号 p , 并以此为依据为待重建人脸图像确定合适的 2D-3D 特征子空间。

基于 2D-3D 特征子空间变换的人脸重建过程如图 3 所示。

3 实验及结果分析

3.1 数据库描述

本文采用自建 (WHU-3DFaceData) 三维人脸数据库, 该数据库包含 20 位对象的三维人脸数据。通过改变控制参量, 可以得到一定数量的二

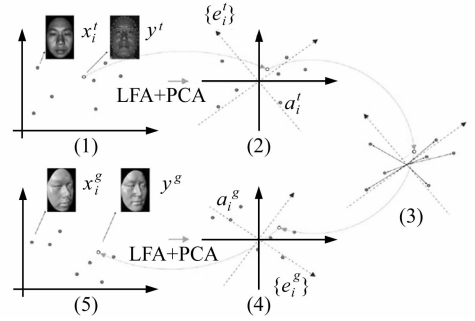


图 3 基于 2D-3D 特征子空间变换的三维人脸重建
Fig. 3 3D Facereconstruction Based on 2D-3D Subspace Transformation

维虚拟图像。本文规定三维人脸模型的水平偏移角和俯仰角的变化范围分别为 $[-45^\circ, 45^\circ]$ 和 $[-15^\circ, 15^\circ]$, 角度变化增量 $\Delta a = 15^\circ$; 光线投影角与方位角的变化范围分别为 $[0^\circ, 60^\circ]$ 和 $[-45^\circ, 45^\circ]$, 角度变化增量均为 $\Delta a = 15^\circ$; 将光线强度设定为 1.5, 每一个三维人脸模型最终可以获得 288 幅二维虚拟图像。

同时, 通过 USB 摄像头获取 10 位对象的 10 幅人脸图像用于三维重建实验。在拍摄过程中, 要求对象不断改变头部姿态, 并始终保持中性表情。

3.2 三维人脸重建系统参数设定

在参数 $d=1, 2, 3, 4$ 时, 同一幅二维人脸图像的三维重建结果如图 4 所示。

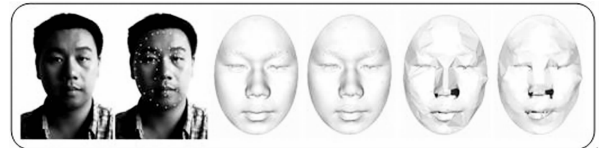


图 4 不同参数 d 下的三维重建结果
Fig. 4 Result of 3D Facereconstruction According to Different Parameter d

通过图 4 可以看出, 三维重建效果随着参数 d 的增加反而下降。因为当参数 d 较小时 ($d=3$ 或 $d=2$), 每个子集中的样本具有“类似”的姿态和光照条件; 但是随着参数 d 的增加, 每个子集中将会包含更多的异构图像, 这会导致以下两方面的问题: ① ASMS 会由于大量异构图像而失去模型的专属性, 从而导致特征点定位精度的降低, 进而影响到人脸特征的提取; ② 大量异构图像的有效特征很难在线性空间中进行提取, 建立的 2D-3D 特征子空间无法对待重建人脸图像进行准确表示, 所以会导致三维重建效果的下降。

但值得注意的是, 参数 d 同样会影响到三维人脸重建系统的执行效率, 因为它控制着子集的

数量和特征点迭代搜索的次数。本文参数 d 与子集数量 M 的对应关系如表 1 所示。

表 1 参数 d 与子集数量 M 的对应关系
Tab. 1 Correspondence Between d and M

子集数量	d			
	1	2	3	4
M	288	18	4	1

为了同时兼顾重建效果和执行效率,本文将参数 d 设定为 2,即在每个子集中水平偏移角和俯仰角的变化范围为 15° ;光线方位角与投影角变化范围同样为 15° 。由于采用统一划分标准,因此,训练样本的子集数量 M 为 18。

3.3 三维人脸重建系统性能分析

为了充分验证本文方法的有效性,在参考文献[6, 7]后设计了两组实验:① 将 WHU-3DFaceData 数据库中的全部样本用于训练,待重建对象属于训练样本集;② 将 WHU-3DFaceData 中的样本进行编号,奇数编号用于训练,偶数编号的对象用于重建实验,然后进行交换。两组实验中的部分重建结果如图 5 所示。



图 5 部分三维重建结果
Fig. 5 Parts of 3D Face Reconstruction Results

通过图 5 可以看出,本文所提出的重建方案能够对非可控条件下获取的单幅二维人脸图像进行有效的三维重建,重建的三维人脸模型具有较强的真实感,能够达到视觉应用的要求。

同时,为了对重建性能进行定量分析,本文利用文献[14]中的方法提取三维人脸模型的轮廓曲线特征进行人脸识别实验,如图 6 所示。在 WHU-3DFaceData 数据库上对 10 位对象进行三维人脸识别实验,并取得了 93% 的识别率。可见,本文所提出的三维人脸重建方案具有理想的重建精度,能够满足实际应用的需要。

3.4 重建系统运行时间分析

对 § 3.3 中两组重建实验的运行时间进行统计并与文献[6]中的方法进行对比,如表 2 所示。可以看出,在重建过程中由于对特征子空间进行动态选择,会降低系统的执行效率;但是本文方法能够对非可控条件下的人脸图像进行有效重建,且总体重建时间与文献[6]中的方法接近。

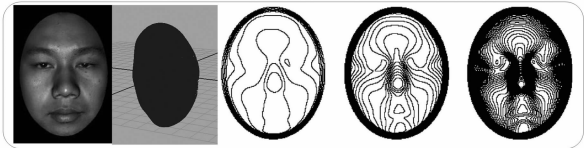


图 6 三维人脸模型不同级别的轮廓曲线
Fig. 6 Level Curves of Depth Function for Several Levels

表 2 不同重建方案时间对比
Tab. 2 Time Comparison of Different Programs

方法	运行环境	时间
基于特征点方法 ^[6]	CPU:3.0 GHz,P4	3 s
本文方法	CPU:3.0 GHz,P4	3.2 s

实验结果表明,与传统方法相比,本文提出的改进方案在保证三维重建效果和执行效率的同时,对光照和姿态变换具有鲁棒性,并在 WHU-3D 数据库上取得了理想的重建效果。探索有表情变化的二维人脸图像三维重建方法将极具挑战性 with 吸引力。另外,由于训练样本数量的有限,如何在特征提取过程中有效避免不相关信息的引入或丢失有用信息,降低噪声的干扰同样值得关注。

参 考 文 献

[1] Massimo T. Dynamic Face Recognition; From Human to Machine Vision [J]. Image and Vision Computing, 2009,27: 222-232

[2] 张祖勋. 数字摄影测量与计算机视觉[J]. 武汉大学学报·信息科学版,2004,29(12): 1 035-1 039

[3] 毛晓波,陈铁军. 基于人类视觉特性的机器视觉系统[J]. 仪器仪表学报, 2010,31(4): 832-836

[4] 廖海斌,陈庆虎,鄢煜尘. 面向形变模型的三维人脸建模研究及其改进[J]. 武汉大学学报·信息科学版,2011,36(2): 649-655

[5] 胡永利,尹宝才,谷春亮. 基于形变模型的三维人脸重建方法及其改进[J]. 计算机学报,2005,28(10): 1 672-1 678

[6] 龚勋,王国胤. 基于特征点的三维人脸形变模型[J]. 软件学报,2009,20(3):724-733

[7] Blanz V, Vetter T. Face Recognition Based on Fitting a 3D Morphable Model [J]. IEEE Trans on Pattern Analysis Machine Intelligence, 2003, 25 (9):1 063-1 074

[8] Blanz V, Vetter T. A Statistical Method for Robust 3D Surface Reconstruction from Sparse Data[C]. The 2nd International Symposium on 3D Data Processing, Visualization and Transmission, Thessaloniki, Greece, 2004

[9] Blanz V, Vetter T. Face Recognition Based on

Frontal Views Generated from Non-frontal Images [C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, USA, 2005

[10] Romdhani S, Vetter T. Face Image Analysis Using a Multiple Feature Fitting Strategy[D]. Basel: University of Basel, 2005

[11] Mena J P, Roberto M, Cesar J. 3D Face Computational Photography Using PCA Spaces[J]. Visual Computer, 2009(25): 899-909

[12] 北京博维恒信科技发展有限公司. 3DCaMega 光学三维扫描系统[OL]. <http://www.3dcamera.com>, 2010

[13] 毕田力,杨旭波,肖双九. 基于谐波图片的重光照技术研究[D]. 上海:上海交通大学,2010

[14] Nagamine T, Uemura T, Masuda I. 3D Facial Image Analysis for Human Identification[C]. International Conference on Pattern Recognition, Hague, Netherlands,1992

第一作者简介:袁理,博士生,研究方向为模式识别与图像处理。
E-mail:principle_net@sohu.com

3D Face Reconstruction for Single Visual Images

YUAN Li^{1,2} CHEN Qinghu¹ LIAO Haibin¹ DUAN Liuyun¹

(1 School of Electronic Information, Wuhan University, 129 Luyoyu Road, Wuhan 430079, China)
(2 School of Electronic Information, Wuhan Polytechnic, 463 Guanshan Road, Wuhan 430074, China)

Abstract: This paper proposes a fast and efficient method of 3D face reconstruction for single visual images which are obtained under uncontrolled conditions. First of all, face feature points are located through the ASMS (active shape model of subset). During this locating processing, ASMS locating accuracy is improved by increasing the completeness and the clustering performance of the training samples, which also sets up the fundamental basis for further face feature extraction. Secondly, 2D-3D feature subspaces are built up according to different subsets respectively using method LFA (local feature analysis), which efficiently fuses the global and local features. Thirdly, a dynamic selection strategy for 2D-3D features subspace selection is designed, which can significantly improve the reconstruction accuracy and keep a content processing efficiency at the same time. What is more, Experimental results show that the proposed method is not only simple and efficient, but also robust to pose and illumination variations, and achieves desired reconstruction results on WHU-3DFaceData database.

Key words: 3D face reconstruction; subsets definition; ASMS; dynamic selection strategy

About the first author: YUAN Li, Ph.D candidate, majors in pattern recognition and image processing.
E-mail: principle_net@sohu.com