

人脸的表情与年龄变换和非完整信息的 重构技术 (下)

付 昀, 郑南宁, 张 婷

(西安交通大学人工智能与机器人研究所, 陕西西安 710049)

摘 要: 非完整视点信息的重构技术可以利用少量的对象视点信息, 结合统计学的理论方法对缺失视点信息进行预测和恢复. 该文探讨了从单幅 2D 人像恢复其它视点图像并重构 3D 形态效果的技术, 提出了非完整人像视点连续变的技术框架. 通过大视点数据库的离线创建, 单视点输入的模式匹配, 以及大视点空间映射与连续视点重建三个模块, 从单视点合成连续的多视点图像. 该技术以尽量少的人像 2D 信息获得高真实度的 3D 效果, 绕过了 3D 模型重构, 可恢复出未知人脸表面的纹理信息. 最后结合 AIAR 人像库的实际建库经验, 讨论了人像数据的采集技术和分类方法, 提出了在有限拍摄条件下的人像库建立方法与建库实例.

关键词: 人像处理; 表情变换; 年龄变换; 视点变换; 人脸图像库

中图分类号: TN911. 73 **文献标识码:** A **文章编号:** 0372-2112 (2003) 12A-1963-08

Facial Expression Transformation, Aging and Invisible View Reconstruction (2)

FU Yun, ZHENG Nan-ning, ZHANG Ting

(Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China)

Abstract: The reconstruction technique of the invisible views of human head belongs to the statistical technique for predicting unknown information. This paper researches on the sequential novel view synthesis or the 3D model computing from a single facial image. An advanced framework for the invisible sequential view reconstruction, which includes discrete view database creating, single view input and model matching, view space projection and view morphing, is presented. The techniques can generate realistic face appearances and invisible views with limited prior information bypassing the building of 3D models. The practical experiences for the AIAR face database creation under limited shooting angles and conditions are also presented at last to help implement the view reconstruction techniques.

Key words: facial image processing; expression synthesis; facial aging; view transformation; facial image database

1 引言

本文研究的人像处理与变换主要包括四个方面: 人脸建模, 表情变换, 年龄变换, 非完整信息的恢复与重构. 在上篇^[1]中主要介绍了近年来人像表情与年龄变换的基本方法, 并对基于稠密特征对应的人脸表情合成与克隆算法, 基于平均脸的人脸图像年轻化变换算法, 以及基于衰老纹理映射和差异性的人脸图像衰老化变换算法进行了理论分析与实验论证. 本篇中要讨论的“非完整信息的恢复与重构”主要研究如何从少量的人脸图像信息中, 结合统计学的理论方法对缺失信息进行预测并恢复. 最基本的研究任务是仅从单幅 2D 人脸图像恢复出其它视点下的图像. 这种对于非完整视点信息的恢

复与重构需要依据大量样本统计信息的经验知识实现.

与此相关的代表性研究工作包括 1996 年 Seize 和 Dyer^[10]在图像渐变技术^[13,14,17]的基础上, 提出了一种利用 2D 图像信息恢复 3D 视图角度渐变的 View Morphing 技术. 该技术首先对两幅不同视点下的人脸图像进行自动的预变换, 产生可以进行非失真 Morphing 操作的平行视点图像; 然后对两幅平行视点图像进行常规的 Morphing 处理获得多幅中间图像; 最后在少量的用户定义操作下, 对这些图像进行后变换处理产生无畸变的新视点人脸图像. 这种方法完全不需要 3D 深度信息, 可以实现逼真的三维立体 Morphing 效果, 具有手工操作量少, 视觉效果自然的优点. 1999 年 Blanz 与 Vetter^[7,8]提出的 3D 渐变模型 (Morphable model), 可以通过人像库中多幅图

像线性组合自动建立.人像的任意视点信息在建立 3D 模型后都可以有效计算.优点是输入简单,数据冗余度低,真实反映人像本质特征.缺点是建立人脸库比较复杂,而且人脸库也只局限于同一年龄层,离线操作的计算复杂度相当高.1998 年 Thomas Vetter^[9]提出了大视点的人像变换,该方法基于线性对象类模型,并且结合了 3D 模型,进行人脸图像单视点至多视点的变换处理,而且视点计算结果具有人像的个性特征(如斑点,皱纹等).该视点变换技术依赖于人脸视点图像数据库,如果所需要合成的视点图像没有相对应的视点图像库,则无法实现相应的变换.小视点范围内的人像视点估计以及小范围的视点变换可以通过 AAM 和 ASM 模型^[26~27]加以实现,这是对文献^[9]模型的有效补充.

本篇的第 2 节首先介绍了人像视点变换的原理与技术;第 3 节在基于稠密特征对应的人像统计学模型基本理论的基础上,讨论了基于 2D 单输入的非完整人像视点信息的恢复与重构技术,然后提出了一种融合多种新方法的视点续变技术框架.该技术可以从一幅输入的人脸 2D 图像,绕过摄像机内外参数,投影矩阵和深度等 3D 信息的获取与利用,借助视点参数空间映射与离散视点的插值技术,实现对输入人像的连续多视点的合成与 3D 的视觉效果仿真^[16~22,25];第 4 节中结合实际经验讨论了该领域人像库的建库技术、分类方法,并从方法、重要性以及关键问题等几个角度提出了在有限拍摄条件下的人像库的建立方法与一些建库实例;第 5 节中给出了人像视点变换的实验结果,证明了算法的有效性;最后,第 6 节对全文进行总结,并对下一步的工作做出展望.

2 人像视点变换的原理与技术

非完整信息恢复与重构技术中的一个主要研究内容就是关于对象任意视点的变换.视点图像隐含着研究对象的三维深度信息,随着视点的改变,原来一些可见的部分逐渐被遮挡或出现孔洞,而另外一些原来不可见的区域逐渐显现出来.传统的视点图像合成方法需要采用 2 幅或 2 幅以上不同视点的人脸图像进行一定的图像处理与变换(例如建立 3D 模型),恢复人脸在其它视点下的信息.如何从单视点输入,利用一定的先验知识重建研究对象的多视点信息是当前视点变换技术的重要研究内容.常见的人像视点变换技术主要归结为如下 5 类:

(1) 3D 人像建模^[7,28],利用 2 幅以上的不同视点人像建立 3D 人像模型,通过模型在水平与垂直方向的旋转,分别获得人像左右旋转以及俯视和仰视的 2D 图像.这种方法可以对人像的个性化信息进行有效的刻画,例如人像形状的不对称,纹理的个性特征(胎记、痣、斑纹等),甚至脸部残缺等;

(2) 光流映射^[9,15],计算两幅在不同视点下的样本图像之间的光流场,然后将光流叠加到新的人脸图像上实现视点变换.由于光流计算方法的局限性,对于大视点的人像变换,该技术无法应用,一般视点变化在 15 度以内基本上有效;

(3) 视点 Morphing^[10~12,20],输入相同对象的两幅不同视点图像,利用图像的预 Warp 操作将两幅图像变为平行视点图像;然后进行常规的 Morphing 处理获得多幅中间图像;最后对

这些图像进行相应的后 Warp 操作产生新视点下的人脸图像.该方法绕开了投影矩阵的计算和 3D 信息的获取,对于未知对象或场景仅需给出多幅 2D 图像信息,可以实现逼真的 3D 效果,具有操作量少,视觉效果自然的特点;

(4) 2D 线性对象类^[3~5,9,15]模型,是建立在线性统计模型基础上的人像建模技术.该理论认为一幅特定人脸图像总可以由一组样本人像通过线性组合来线性表示,而这一类图像表达的对象类就称为线性对象类.通过建立不同视点下的样本库,利用不同样本空间的视点投影关系,就可以将新的人脸图像在其它视点下的图像合成出来.这种方法对于人像库内的样本数目以及视点角度等要求比较高,此外合成的新视点图像无法表现特定人脸的个性纹理信息;

(5) 主动形状(ASM)和主动形态模型(AAM)^[26,27,29],AAM 和 ASM 模型都基于特征点集合分布模型.利用包含强结构信息的点特征集合表示人脸形状.特别是 AAM 模型对人脸的形状和纹理进行综合的统计学习,以对象形状差异作为约束条件,定义模型参数来控制模型形状与模型的纹理变化,经过反复迭代缩小参数误差,运算收敛后可以获得准确的模型匹配.这类模型对于人脸对象不同姿态的跟踪定位快速且精确,因此可以确定人脸视点角度值,并通过变换技术产生小视点范围内的人像姿态变化.

其中(2),(4)和(5)均为单图像输入下非完整信息恢复与重构技术.而(1)和(3)的方法需要多幅输入才可以应用.

3 基于稠密特征对应的非完整人像视点信息重构技术

3.1 基于稠密特征对应的统计学人脸模型

线性对象类(Linear object class)模型^[2~9]是一种广泛用于对象表达(representation)的统计学模型,基本思想是假设一类对象集中的任一个对象都可以由该类中其它对象或对象元通过线性组合的方式表达.对象元可以是主分量分析(PCA)主元,或者是未经过数据处理的样本(prototype)等.在人脸对象类的研究中,经过预变换、对准与表达后的人脸图像可以通过线性对象类模型进行有效表达.下面讨论线性对象类模型的建库方法:

假设 I^{eigen} 为一个对象元图像,并表示为一个映射: $I^{\text{eigen}}: \mathbf{R}^2 \rightarrow I$,其中 $I^{\text{eigen}}(x, y)$ 表示对象图像中一个像素的点的灰度值.计算对象集中每一个对象元到参考对象元图像 I_r 的稠密对准 $M(x, y)$,对准后的对象元图像表示为 $\tilde{T}_1^{\text{eigen}}, \tilde{T}_2^{\text{eigen}}, \dots, \tilde{T}_n^{\text{eigen}}$,其中 n 是集中对象的个数.通过对象元图像的线性组合建立的对象类模型可由式(1)表示.

$$I_{\text{model}} = I_r + \sum_{i=1}^n p_i \cdot \tilde{T}_i^{\text{eigen}}(x, y) \quad (1)$$

这里 $p = [p_0, p_1, \dots, p_n]$ 是线性模型的线性系数.结合对象形状的稠密对准,线性组合 $\sum_{i=1}^n p_i \cdot \tilde{T}_i^{\text{eigen}}(x, y)$ 描述了一个由模型表达的对象元的纹理信息.对象集合外的同类新对象图像可以用已建立的对象类模型匹配表达,即

$$I_{\text{novel}} \approx I_{\text{model}} \quad (2)$$

这里取 I^{eigen} 为原型样本 I , 则经过对准后的图像纹理为 \tilde{I} . 则

$$I_{\text{model}} = \sum_{i=0}^n p_i \cdot \tilde{I}_i(x, y), \text{ 其中 } I_0 = I.$$

为了获得最佳的模型匹配和对象表达效果, 这里定义新图像和模型重构图像间的误差能量函数(3)

$$E(p_0, p_1, \dots, p_n) = \frac{1}{2} \int_{x,y} [I_{\text{novel}}(x, y) \cdot M(x, y) - I_{\text{model}}(x, y)]^2 dx, y$$

$$E(p_0, p_1, \dots, p_n) = \frac{1}{2} \int_{x,y} [I_{\text{novel}}(x, y) \cdot M(x, y) - \sum_{i=0}^n p_i \tilde{I}_i(x, y)]^2 dx, y$$

式(2)中最佳的建模匹配求解是一个最优化问题. 配合我们提出的新的模型匹配技术(基于相关扰动的随机梯度下降算法^[18]、学习率自适应^[25]、非统一抽样和动态高斯金字塔分析^[30]、伪逆运算^[16,51])求解误差能量函数 $E(p)$ 达到全局近似最优解时的线性系数 p , 就可以利用有限的样本集合描述最佳的模型匹配效果. 这里计算出的局部最优线性系数解表示为 $p^* = (p_1^*, p_2^*, \dots, p_n^*) \in \mathbb{R}^n$, 使误差 $E(p^*)$ 满足最小值.

对于人脸对象而言, 我们同时考虑形状与纹理的线性对象建模. 先在离线状态下对样本图像进行人像表达, 得到对应的形状和纹理矢量然后分别建立形状与纹理的模型, 通过模型匹配技术仿真出新的人脸图像, 实现一系列变换处理. 图 1 描述了基于稠密特征对应的线性人脸对象模型的建模框架. 图中第一行对应纹理模型, 第二行对应形状模型.

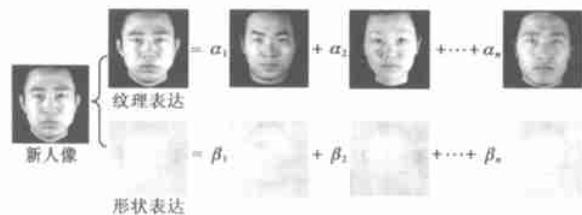


图 1 基于稠密特征对应的线性人脸对象模型

3.2 基于 2D 单输入的非完整人像视点信息的恢复与重构

在 2D 单视点图像的非完整信息输入下, 恢复与重构其它视点信息是人像处理的一个研究课题^[9,10,20,21,25,30]. 对于非完整信息的预测、恢复与重构, 其关键问题主要包括建模与视点空间映射两方面. 我们建立了多视点图像库, 对每一个样本对象分别获取固定视点下的模拟光图像(如左右 15 度, 左右 30 度, 左右 45 度, 左右 60 度, 左右 75 度, 左右 90 度, 正面等), 如图 2 所示. 在各视点库中, 每个对象保持固定的顺序存储. 然后建立基于稠密特征对应的线性对象类模型, 通过某视点空间的模型匹配训练出一组模型参数, 再利用视点空间的投影变换在其它视点空间中利用对应模型参数重建视点图像.



图 2 人脸视点图像

假设所建立的线性对象是 3D 人像 X , 则 $X = \sum_i x_i \cdot X_i$, 假设经过线性的 3D 旋转变换操作 R 后, $X^r = R \cdot X$. 对线性对

象类的所有样本进行相同的旋转变换操作, 则 $X^r = \sum_i x_i \cdot X_i^r$.

这意味着 3D 人像旋转变换操作前后的线性组合形式一致, 系数与组合对象样本顺序都不变. 将 3D 人像在不同视点的 2D 平面上投影获得 I , 其投影变换为 P , 则有 $I = PX$ 与 $I^r = PX^r$. 由此推出 $I = \sum_i x_i \cdot I_i$ 与 $I^r = \sum_i x_i \cdot I_i^r$, 即 2D 人像的在三维意义下的旋转变换前后的线性组合形式一致, 系数与组合对象样本顺序都不变^[9,16]. 在这样的理论前提下, 对应 2D 人像的视点处理就可以完全绕开 3D 人像信息. 当然这种操作的前提是选择原型样本作为人像空间基底. 如图 3 显示了选择原型样本作为人像空间基底时, 2D 人像视点变换前后的人像空间投影, 其中参数 α_k^* , β_k^* 分别表示不同视点空间下样本线性组合的系数, 采用原型样本基底时, $\alpha_k^* = \beta_k^* = \delta_{ik}$.

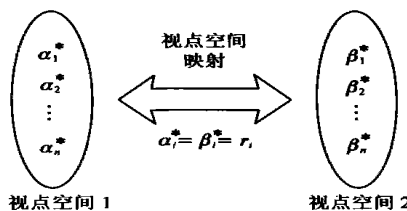


图 3 原型样本人像视点空间投影

如果采用其它空间基底(如 PCA 主元)建立人脸模型, 这种空间映射关系就更加复杂. 假设选择 n 个 PCA 主元作为空间基底, 表示为 (e_1, e_2, \dots, e_n) , 则 $I_i = \sum_k x_k \cdot e_k$; 旋转后的空间基底表示为 $(e_1^r, e_2^r, \dots, e_n^r)$, 则 $I_i^r = \sum_k x_k \cdot e_k^r$. 由如上的结论可知 $I = \sum_k x_k \cdot e_k$ 与 $I^r = \sum_k x_k \cdot e_k^r$, 其中两个视点空间的基底数目一致, 都从大到小排列. 经过推倒可知: $x_k = \sum_i x_i \cdot \delta_{ik}$, $x_k = \sum_i x_i \cdot \delta_{ik}$. 模型匹配算法可以采用伪逆运算, 即首先将 $I = \sum_i x_i \cdot I_i$ 写成矩阵形式: $I_{\text{novel}} = I \cdot p$, 其中 $I = [I_1, I_2, \dots, I_n]$, 则 $p = I^+ \cdot I_{\text{novel}}$, 这里 $I^+ = (I^T \cdot I)^{-1} \cdot I^T$ 是矩阵 I 的伪逆.

3.3 基于 2D 单输入的非完整人像视点续变技术框架

不同角度的水平视点变换会产生诸多空洞, 遮挡等现象. 传统的方法需要采用 2 幅以上不同视点的人脸图像建立 3D 模型, 从而恢复人脸的不同视点信息. 这里我们提出了一种有效的单输入人像的非完整视点续变技术框架. 如图 4 所示, 该技术框架主要包括三大模块: 大视点数据库的离线创建模块, 单视点输入与模型匹配模块, 以及大视点空间映射与连续视点重建模块. 首先在离线学习阶段分别建立研究对象多个离散视点下的对象数据库以及线性对象类模型; 然后利用近似全局最优化方法以及匹配提升技术(基于线性相关性扰动的随机梯度下降算法, 学习率自适应, 结合不等概抽样和整群抽样技术动态调整每级高斯金字塔图像抽样分布的动态高斯金字塔分析^[18,22,25,30])求解模型参数; 利用视点空间映射技术对研究对象进行多视点下的模型匹配表达^[9], 最后利用 4 点或 8 点渐变算法^[10-12]实现从离散视点到连续视点的重建. 该技术可以应用于对象立体视点变换的视频分析与合成.

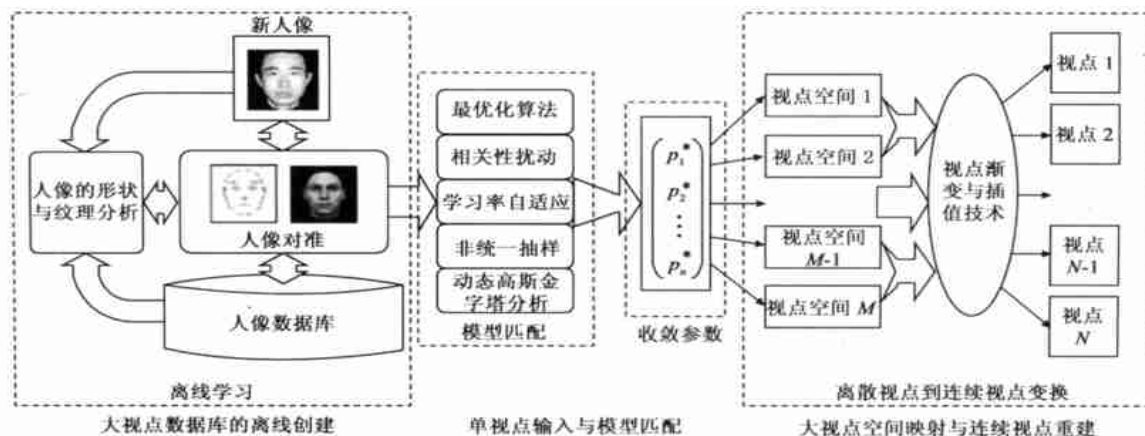


图4 基于2D单输入的非完整人像视点连续变换框架结构

4 人像库的建立方法与关键技术

人像处理与变换是在人脸图像高层信息理解的基础上实现的计算机仿真。这种对人脸对象的姿态、属性、表象的理解需要基于大量的样本素材。人脸图像库的建立^[19]就是在这种工程需求的前提下提出来的。目前人像库的种类很多,根据不同的应用分类,通常包括人脸识别的测试图像库,人脸高层属性变换的测试图像库,人脸表情图像库,人脸姿态图像库等等;根据图像库的图像采集方法与图像性质分类,通常包括结构光人像库与模拟光人像库。

本文所提出的人脸表情、年龄和非完整信息恢复与重构的处理技术,表达效果的好坏直接受到所采用的人脸图像库的限制。在基于稠密特征对应的人脸表情合成与克隆技术中,可合成的人脸表情的数目与种类依赖于人脸表情图像库中已知的样本人脸表情图像;在基于平均脸的人像年轻化与衰老化技术中,平均脸图像会随着人像库中样本数目,样本人像年龄分布,人像性别比例,人种的统一性,人脸表情姿态的统一性以及光照条件的变化而呈现不同的形状和纹理,也因此人像的年龄变换会随之呈现不同的变换结果;在基于2D单输入的非完整人像视点信息的恢复与重构技术中,人像不同视点库中人脸对象的一致性,姿态视点的统一性,表情的一致性,以及人像数目都直接影响非完整信息恢复的准确性与真实性;模型匹配中,新人像的光照条件需要通过技术手段统一与人像库的光照条件一致,否则难以保证人像的准确表达。

4.1 结构光人像库

结构光人像库是通过激光扫描设备创建的^[7,31]。在整个过程中,拍摄对象保持不动,扫描设备以对象为轴心,从对象的一边旋转至另一边进行数据采集,由此得到一组数字化的点,每一个点用 (x, y, z) 坐标值记录它在空间中的位置信息,同时使用24-bit RGB值记录该点的颜色信息。这种人像库中的样本数据真实的记录了人像的3D头部结构信息,但这种信息为非可视化数据,只能通过投影技术将数据投影到特定视点角度下得到对应的2D图像。

MPI人像库^[7]属于结构光图像库,是由德国MPI(Max-planck institute for biological cybernetics)实验室用激光扫描器

(CyberwareTM)^[31]创建的,通过扫描获取的人像结构光柱面表达信息,并存储在 R, G, B 三个相同分辨率的8 bit色彩信道中。目前该库只提供科学研究用途,通过签署授权协议可以获得该人像库的使用权。

我们通过授权得到的MPI人像库中包含1400张 256×256 大小的RGB人脸图像,其中包含了200个高加索地区的中青年人脸图像,其中男性女性各100人。每个人均有正面、左右30度、左右60度和左右90度7张视图。所有的人脸图像都消除了眼镜、装饰、发式,通过预处理,消除颈部以下,两耳以外的区域。

图5显示了MPI人像库中一名女性与一名男性在5个不同视点下的8bit灰度图像。在本文的研究中,大部分算法测试实验都是基于MPI人像库进行的,特别是视点变换的研究需要依赖于该库的7个人像视点子库。



图5 MPI人像库中不同视点下的人脸图像

4.2 模拟光人像库

激光扫描器是一种比较昂贵的设备,在实际建库中难于提供。我们根据所在研究机构(XJTU, AIAR)的基本科研条件与实际环境,提出了一种模拟光人像库建库方法,并且在此基础上建立了4个可满足不同应用需求的人像库。该人像库具有自主版权,归属于西安交通大学人工智能与机器人研究所(©AIAR, XJTU)。四个人像库依次命名为:AIAR V1.0 (Expression), AIAR V2.0 (Aging), AIAR V3.0 (View), AIAR V4.0 (Illumination)。分别应用于人像表情变换,人像年龄变换,人像视点变换与人像光照效果变换。四个人像库中的样本图像数目随着研究的进程还在不断的扩充。

模拟光人像库在创建时采用的仪器主要包括:OLYMPUS

C-5050ZOOM 数码照相机(包括闪光灯),三角架,三个点光源灯,一个多点光源灯,圆形坐标图,旋转座椅,铅垂,背景幕布等.拍摄环境为暗房或微弱散射光环境.

我们首先在地面绘制圆形坐标图,并在座标图上标定旋转角度(主要用于视点人像库的建立);之后固定旋转座椅位置,使旋转轴垂直于坐标图平面并相交于坐标原点;同时将铅垂固定于座椅前沿,并指向地面圆形坐标图;摄像机固定于三角架上,与座椅距离 1.5m 左右.

在拍摄过程中,要求拍摄对象坐于旋转座椅上,保持身体正直,使颈椎的延长线与座椅旋转轴重合;保持自然静止状态,眼睛凝视正前方,眼睑与虹膜正切,上下嘴唇接触嘴巴闭合,上下牙齿接触,脸部自然放松.要去掉头部附属物以及头发对脸部信息的遮挡.

由于 4 个人像库有不同的应用目的,所以对库内人像的要求也不尽相同,拍摄人像的基本操作也不同,下面根据不同建库要求进行分别介绍.

4.2.1 AIAR V1.0 (Expression) 表情人像库 AIAR V1.0 (Expression) 表情人像库中有 10 个人脸对象,共 300 幅 640 x 480 的人脸表情 JPG 正面图像.包括标准表情,如喜怒哀乐等,以及部分难以用语言描述的复杂表情.在拍摄时,保持光照条件一致,并让拍摄对象保持身体姿态与头部位置静止不动,正对照相机.脸部自由做出喜怒哀乐等各种基本表情,并可以自由发挥做出复杂的、奇怪的夸张动作.为了对表情纹理进行细节刻画,要求拍摄对象尽量夸大动作幅度,使得纹理变换信息更加突现出来.图 6 示出了 2 个人脸对象的 10 幅人脸表情图像.这些人像为表情合成与表情克隆技术的应用提供了大量素材.

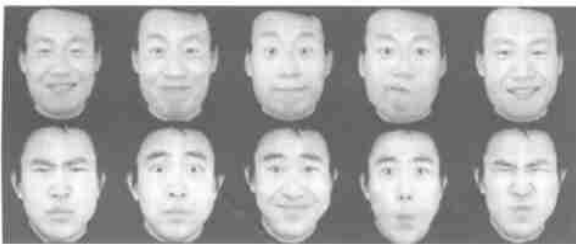


图 6 AIAR V1.0 (Expression) 表情人像库

4.2.2 AIAR V2.0 (Aging) 衰老人像库 AIAR V2.0 (Aging) 衰老人像库由 17 个人组成,每人若干张正面人脸组成.分为年老者与年轻者两组,年老者主要由中老年人像组成,年轻者主要由青少年人像组成.拍摄时的光照条件一致.要求拍摄对象尽量保持自然的中性表情状态.每幅图像均为 256 x256 的



图 7 AIAR V2.0 (Aging) 衰老人像库

8-bit 灰度图像.经过预处理,消除了头发以及背景的影响.图 7 示出了 10 幅该人像库中的人脸图像,第一行显示了 5 幅中老年人像,第二行显示了 5 幅青年人像.人脸衰老模板可以通过前文的方法,从库中的年老人像获取.

4.2.3 AIAR V3.0 (View) 视点人像库 AIAR V3.0 (View) 视点人像库目前由 80 个青年人组成,男性人数略多于女性人数.库中包括每个人在 11 个不同视点角度(15°为一个单位)下拍摄的自然表情图像.每幅图像均为 256 x256 的 8-bit 灰度图像.经过预处理,消除了头发以及背景的影响.目前库中分为 11 个视点子库,共计 880 幅图像.

在该库的建立过程中,采用相机固定而对象旋转的方法.拍摄对象坐于旋转座椅上,保持身体正直,使颈椎的延长线与座椅旋转轴重合,脸部自然放松.操作时,由技术人员控制旋转座椅的旋转,利用铅垂指向坐标图的角度标线估计旋转的角度.该方法的难度在于如何准确的进行旋转角度的测定,目前由技术人员目测决定角度存在一定误差.图 8 示出了该人像库中两个人脸对象的不同视点图像,其中包括一个女性人像与一个男性人像.

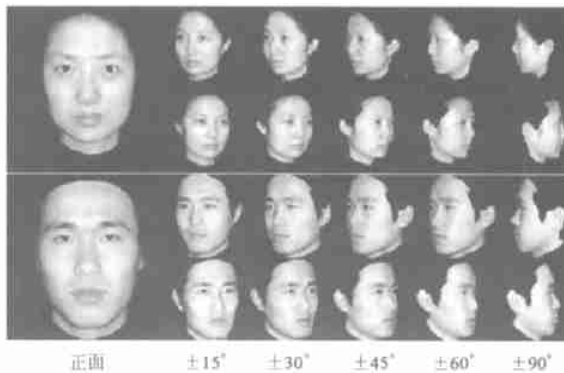


图 8 AIAR V3.0 (View) 视点人像库

4.2.4 AIAR V4.0 (Illumination) 光照人像库 AIAR V4.0 (Illumination) 光照人像库目前由 13 个正面人脸在 8 种不同光照条件下的拍摄得到的图像组成,其中包括 5 位女性和 8 位男性.在该库的建立过程中,要求拍摄对象保持静止不动,表情自然,消除头发,眼镜,饰物的影响;在暗房设置 3 个空间独立的固定点光源和一个远景的固定多点光源,分别拍摄 3 个独立点光源组合开闭情况下的人脸图像,以及远景固定多点光源单独点亮环境下的人脸图像.图 9 示出该人像库中一个人脸对象在不同光照条件下的图像.



图 9 AIAR V4.0 (Illumination) 光照人像库

该人像库的主要用途在于对已知和未知人像光照条件的估计和调节,以满足光照条件的一致性要求.基本方法是利用 Amnon Shashua^[24]提出的比率图理论(Quotient image)结合 AIAR V4.0 人像数据集实现对人脸 Lambertian 表面反射率比例值的获取,从而计算出有效的光照参数;利用光照参数调节不一致光照条件下的人脸图像实现光照统一.

4.3 结构光与模拟光人像库的比较

结构光人像库与模拟光人像库都可以满足不同的人像处理需要.但是由于建库仪器与方式的不同,使得两种人像库的性质不尽相同.结构光人像库采用的设备复杂、昂贵,而模拟光建库的设备简单、低廉、可操作性强;从数据格式角度考虑,结构光人像库中的样本数据是非可视化的数据结构,而模拟光人像库中的样本数据为 2D 的图像数据;从普通的人像处理角度考虑,结构光人像库存在数据冗余,如果不需要 3D 的个性化建模,完全可以通过模拟光的人像库实现;从精确度方面考虑,结构光人像库的人像视点角度精确度较高,而模拟光人像库略低,但是从图像处理技术的角度,这种视点角度的差别可以通过有效的算法进行衡量与矫正.

5 实验

以如上介绍的 MPI 人像库和 4 个 AIAR 人像子库为测试样本进行如下实验,验证本文所提出基于 2D 单输入的非完整人像视点续变框架结构的技术性能.图 10 示出了利用 MPI 人像库为视点样本空间的视点变换结果,其中选择 150 幅人像为库内人像,其它 50 幅人像为库外新图像进行实验.图 10(a)中的输入图像为库内图像,图 10(b)中的输入图像为库外新图像.首先进行人像建模后,获得正面人像视点空间中样本线性组合的系数向量,经过空间投影后重构出其它视点下的人脸图像.图 10(a)和图 10(b)中的第一行图像为已知的原视点图像,第二行为合成的视点图像.通过主观比较可以看出,库内人脸的视点图像合成效果较为理想,而库外人像合成效果逼真程度较低,但是已经具有一定的真实性.第三行图像为合成人像与输入人像的差图像,用于客观评价模型匹配和视点续变效果.可以看出库内和库外人像重构结果都存在一定的误差,特别是在五官和脸部轮廓位置,主要原因在于人像库的样本数目太少,可以通过扩大视点空间维数来提高真实的

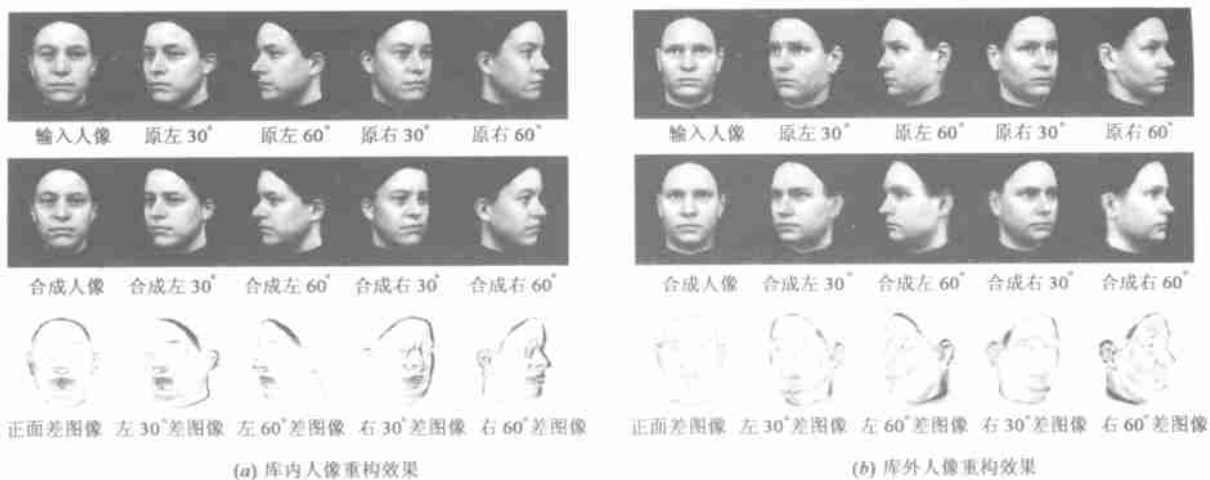


图 10 MPI 人像库中 150 个人像为视点样本空间的视点变换

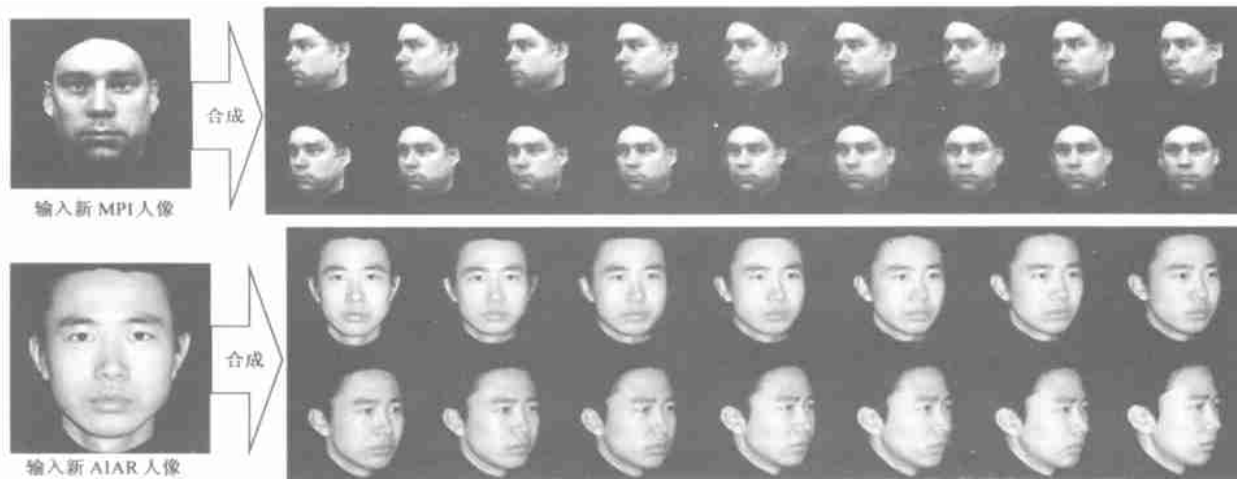


图 11 MPI 人像库和 AIAR 人像库中某男性人像在 2D 单输入条件下的连续视点重构与续变

合成效果;此外,模型匹配近似全局最优解的算法选择也会直接影响视点图像重建效果.图 11 分别示出了在所提出的基于 2D 单输入的非完整人像视点连续变换技术框架基础上,对 MPI 人像库某男性人像和 AIAR 人像库某男性人像在 2D 单输入条件下的连续视点重构与续变效果.首先输入人像的正面(或侧面 30 度,侧面 60 度,侧面 90 度的任意一幅)图像,经过模型匹配模块以及大视点空间映射重构出其它未知的离散视点图像;然后利用视点渐变技术插值出从左侧到右侧连续变化的视点图像.重构与合成结果自然,3D 视觉效果逼真生动.

6 结论与展望

本文在建立基于稠密特征对应的人像线性模型基础上,进行人脸图像形状与纹理分离的矢量化表达,探索了人脸图像处理与变换的相关技术,包括基于稠密特征对应的人脸表情合成与克隆算法,基于平均脸的人脸图像年轻化变换算法,基于衰老纹理映射的人脸图像衰老化变换算法,以及非完整视点信息恢复与重构技术.根据实际应用需求建立了适应不同人像变换要求的人脸图像库,并分析了结构光人像库与模拟光人像库的基本特性与区别,总结了 AIAR 人像库的建库经验.在 MPI 人像库与 AIAR 人像库为样本的基础上进行了有效实验,对本文提出的算法和技术进行了验证.实验结果证实了算法的有效性与鲁棒性.作者建议的进一步工作包括如下 5 个方面:

(1) 多模型融合.采用多种模型相结合的方法,实现多精度的人像处理与合成.例如利用肌肉模型实现人脸表情的个性化微调处理,利用生理模型实现人脸三维模型预测,结合本文算法实现从 2D 人像变换到 3D 人像变换的扩展;基于 Lambertian 类和线性对象类融合模型的人像真实感处理,例如比例图^[16,24,32]的概念可以同时应用于表情,衰老和光照变换等处理中,是否可以统一在一个框架下进行人像处理;

(2) 自动化处理.实现人像稠密特征的自动对准是人像处理中的一个难点问题.常用的技术包括光流算法^[2-9,23]等.但是这种方法算法复杂度高,难以实现实时处理.此外可以考虑在特征线标定模型的基础上,利用边缘检测,图像分割,统计直方图,聚类算法实现人像的自动标线;

(3) 不同人种人像库的融合.本文采用的技术依赖于图像库的人种属性.进一步的研究需要考虑如何建立有效的融合策略,使不同人种的人脸图像可以统一在一个图像库中从而大幅度扩大人像库的容量和信息量;

(4) 表情与年龄模板的扩充.本文提出的表情变换与人像衰老变换中都利用了不同属性的模板,增加模板无疑会增加人像处理的表达效果.同时还可以考虑增加人像形状和纹理的独立变换模板,从而提高人像变换的控制能力;

(5) 人像库的扩展.人像库的重要性在文中已经有所阐述,进一步的工作包括人像库样本数目的扩充,不同属性种类的扩充,不同人种、性别、年龄等类别的扩充,以及特殊人群人像的扩充等.

致谢 感谢德国 Max-Planck 研究所(MPI, Tübingen, Germany) 提

供的人脸图像库;对于在 AIAR 人像库建库过程给予帮助的所有西安交通大学人工智能与机器人研究所成员表示由衷感谢;感谢张强等智能人像研究小组成员在研究工作中做出的贡献.

参考文献:

- [1] 郑南宁,付昀,张婷,卓峰.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.
- [2] S Ullman, R Basri. Recognition by linear combinations of models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13:992-1006.
- [3] 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.
- [4] Michael J Jones, Tomaso Poggio. Model-based matching of line drawings by linear combinations of examples[A]. Proceedings of the International Conference on Computer Vision[C]. Boston Massachusetts, 1995. 532-536.
- [5] David Beymer. Vectorizing face images by interleaving shape and texture computations[R]. A I Memo 1537, Massachusetts:MIT, 1995.
- [6] Michael J Jones, Tomaso Poggio. Model-based matching by linear combinations of examples[R]. A I Memo 1583, Massachusetts:MIT, 1996.
- [7] V Blanz, T Vetter. A morphable model for the synthesis of 3D faces[A]. Computer Graphics Annual Conference Series, SIGGRAPH '99[C]. Los Angeles, California:ACM SIGGRAPH, 1999. 187-194.
- [8] 张辉,许光,谢峰.基于形变模型的物理建模与分析[J].计算机学报,2001,24(6):561-567.
- [9] Vetter T. Synthesis of novel views from a single face image[J]. International Journal of Computer Vision (IJCV), 1998, 28(2):103-116.
- [10] S M Seitz, C R Dyer. View morphing[A]. Computer Graphics Annual Conference Series, SIGGRAPH '96[C]. New Orleans, Louisiana:ACM SIGGRAPH, August, 1996. 21-30.
- [11] S M Seitz, C R Dyer. Physically valid view synthesis by image interpolation[A]. Proceedings of IEEE Workshop on Representations of Visual Scenes[C]. Cambridge Massachusetts:IEEE Computer Society, 1995. 18-25.
- [12] S M Seitz, C R Dyer. Complete structure from four point correspondences[A]. Proceedings of Intl. Conf. on Computer Vision (ICCV)[C]. Cambridge Massachusetts:IEEE Computer Society, 1995. 330-337.
- [13] G Wolberg. Image morphing survey[J]. The Visual Computer, 1999, 14:360-372.
- [14] Beier T, Neely S. Feature-based image metamorphosis[J]. Computer Graphics, 1992, 26(2):35-42.
- [15] Beymer D, Shashua A, Poggio T. Example based image analysis and synthesis[R]. AI Memo 1431, AI Laboratory, MIT, 1993.
- [16] 张强.基于稠密特征对应的人脸图像表达及人脸属性变换[D].西安:西安交通大学人工智能与机器人研究所,2003.
- [17] 付昀.人脸图像渐变算法研究[J].中国图像图形学报(A), 2003,(Special.):174-182.
- [18] 付昀.基于原型样本线性组合的人脸图像 CDSGD 鲁棒对准[R].西安:西安交通大学人工智能与机器人研究所,2002.

- [19] 付昀,张婷.人像库的建立方法与关键技术[R].西安:西安交通大学人工智能与机器人研究所,2003.
- [20] Fu Yun,Zheng Nanning. Sequential view synthesis from a single view [R]. AIAR. XiTU. Xi'an,2003.
- [21] Fu Yun,Zheng Nanning. Non-uniform sampling based dynamic matching for the linear model of human faces [R]. AIAR. XiTU. Xi'an, 2003.
- [22] 付昀.基于非统一抽样的动态高斯金字塔分析[R].西安:西安交通大学人工智能与机器人研究所,2003.
- [23] J R Bergen,R Hingorani. Hierarchical motion based frame rate conversion [R]. Technical report,David Sarnoff Research Center, April, 1990.
- [24] A Shashua,T Riklin-Raviv. The quotient image: class-based re-rendering and recognition with varying illuminations[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2001,23(2):129-139.
- [25] 卓峰,徐维朴,张强,张婷,付昀.基于2D样本的人脸图像视点变换[J].计算机应用,2003,12,(Suppl.)23:74-76.
- [26] Timothy F Cootes,Gareth J Edwards,Christopher J Taylor. Active appearance models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2001,23(6):681-685.
- [27] Lanitis A,Taylor CJ,Cootes TF,Ahmad T. Automatic interpretation of human faces and hand gestures using flexible models[A]. Proceedings of International Workshop on Face and Gesture Recognition [C]. Zurich,Switzerland:IEEE Computer Society,1995.98-103.
- [28] Pighin F,Hecker J,Lischinski D,et al. Synthesizing realistic facial expressions from photographs [A]. Computer Graphics Annual Conference Series, SIGGRAPH '98 [C]. Orlando, Florida: ACM SIGGRAPH, 1998.75-84.
- [29] 陈洪.人脸图像自动分析与绘制的统计学习方法[D].西安:西安交通大学人工智能与机器人研究所,2002.
- [30] 付昀.线性人脸对象类模型的匹配提升技术[R].西安:西安交通大学人工智能与机器人研究所,2003.
- [31] Cyberware Laboratory Inc. The head & face 3D color scanner bundles the 3030RGB scan head with the PS motion platform [EB/OL]. <http://www.cyberware.com/products/psInfo.html>. 2003.
- [32] Liu Z C,Shan Y,Zhang Z Y. Expressive expression mapping with ratio images[A]. Proc. of the 28th Annual Conference on Computer Graphics and Interactive Techniques [C]. Los Angeles, California: ACM Press, 2001.271-276.

作者简介:



付昀男,1979年生于陕西西安,西安交通大学人工智能与机器人研究所硕士研究生,主要研究方向为计算机视觉与图像处理、模式识别与智能系统、人像处理与人脸识别等,发表学术论文10余篇,曾获得美国Rockwell自动化科研奖、美国HP科研银质奖章、美国GE技术创新“爱迪生”奖杯。



郑南宁男,1952年生于江苏南京,中国工程院院士,西安交通大学教授,长期从事计算机视觉与模式识别、数字视频与信号处理等方面的研究工作,发表学术论文100余篇。