

人脸的表情与年龄变换和非完整信息的 重构技术 (上)

郑南宁,付 昀,张 婷,卓 峰

(西安交通大学人工智能与机器人研究所,陕西西安 710049)

摘 要: 人脸表情变换、年龄变换和非完整信息重构是计算机视觉与图像处理领域的重要研究内容. 该文研究了基于线性统计学模型的真实感人像变换和处理技术. 应用对人像形状与纹理的矢量化表达,并结合 Lambertian 表面比率图理论,提出了基于稠密特征对应的人脸复杂表情的合成与克隆算法;根据人像平均化最美的生理学理论,提出了基于平均脸的人像年轻化变换算法,以及基于衰老纹理映射和差异性理论的人像衰老化变换算法. 以 MPI 高加索人像和 AIAR 东方人像库数据为测试样本的实验结果,验证了人像表情合成、衰老化与年轻化算法的有效性.

关键词: 人像变换;人像处理;表情变换;年龄变换;视点变换

中图分类号: TN911. 73 **文献标识码:** A **文章编号:** 0372-2112 (2003) 12A-1955-08

Facial Expression Transformation, Aging and Invisible View Reconstruction (1)

ZHENG Nan-ning, FU Yun, ZHANG Ting, ZHUO Feng

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

Abstract: Facial expression synthesis, facial aging and the reconstruction of the invisible facial views are important research contents in the computer vision and computer graphics fields. This paper addresses the statistical model-based techniques of facial image transformation and processing. With the pixel-wise correspondence computing, the facial image is represented as coupled shape-texture vectors. Based on the quotient image of Lambertian surface classes, an algorithm for realistic expression synthesis and clone is first proposed. Then, the new technique for face juvenility transformation from average face, according to the related hypotheses for face attractiveness in Psychology, is presented. Finally, an ageing transformation algorithm is proposed from merging the age template mapping and the distinctiveness caricaturing methods. With the examples of MPI Caucasian Face and AIAR Asian Face databases, realistic and natural experimental results validate the approaches for the facial expression synthesis and facial age transformation.

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

1 引言

人像处理与变换是计算机视觉领域和图像处理领域的重要研究课题,一直受到诸多研究人员的关注. 在刑侦、医疗、娱乐、信息、空间等领域中,人像的处理与变换有着广泛的应用需求. 特别是,公安刑侦部门在刑事侦察过程中需要根据目击者口述信息进行人脸画像、图像拼接,以及嫌犯模糊查询;需要从单一的侧视、俯视以及仰视等由监视系统获取的视点图像恢复出正面人像;需要从化妆人像识别个人身份信息;需要从失踪儿童照片预测多年后的人脸相貌,等等. 医疗领域中的人脸外科美容术,需要对人脸缺损部位进行计算机填补造型,或者对要求部位进行美化处理等,人像的建模与处理技术成为手术成功与否的关键环节. 在影视娱乐的多媒体特技中,真

实人像与卡通人像的夸张表情,瞬间衰老与年轻化处理都需要有计算机的辅助实现. 人脸表情、年龄变换以及非完整信息恢复与重构是该研究领域中重要的研究内容,有着丰富的研究素材与极高的应用价值,是本文的研究重点^[59~68].

本文研究的人像处理与变换的核心内容主要分为四个方面:人脸建模、表情变换、年龄变换以及非完整信息恢复与重构.“人脸建模”要求结合对人脸对象的一般理解,对人类面容及其生物特征形态的共同特性进行数学表达;“表情变换”研究如何对特定人脸进行图像变换处理,合成真实鲜活的表情动画;“年龄变换”需要结合一定的生理学、认知学以及心理学理论对人脸多年后的衰老化,以及多年前的年轻化形态进行预测、重构与仿真;“非完整信息的恢复与重构”主要研究如何利用一定的先验知识,从少量的人脸图像信息中,结合统计学

收稿日期:2003-09-18;修回日期:2003-12-30

基金项目:国家自然科学基金创新研究群体(No. 60024301);国家自然科学基金(No. 60205001);2002-2003年度河南省重大科技基金(No. 0222020400)

的理论方法对缺失信息进行预测并恢复.人脸模型是实现人像处理与变换的基础,模型的选择与理论依据直接影响到表情、年龄变换以及非完整信息恢复与重构的效果.

从上个世纪七十年代至今,国内外有关人像的计算机处理研究成果很多^[24],有代表性的研究包括 1974 年 Parke^[11]首先用计算机合成人脸,利用将近 400 个顶点组成的 250 个多边形定义人脸表面皮肤,利用余弦插值法计算连续表情的中间图像,从而实现真人动画,这是人脸模型的早期雏形.1978 年 Paul Ekman^[11]提出的人脸动作编码系统 FACS (Facial action coding system),利用表情单元控制肌肉运动,实现复杂表情的组合与仿真.1982 年 Parke^[8]提出的人脸参数模型创建方法,是一种表情参数模型主框架理论,目前使用最为广泛.1987 年 Waters 建立了肌肉模型,利用独立于皮下组织的表皮特征点矢量产生几何形变,可以根据脸部肌肉的收缩和扩张等动作合成脸部组织的细微变形,脸部皮肤动作和表情动画^[24].1990 年 Demetri 等人^[21]扩展了 Waters 模型,建立了一种人脸物理模型,利用三层变形网格结构模拟人脸组织,这三层结构包括皮肤层,皮下脂肪组织层和肌肉层.1993 年 Takaaki 等人^[2]提出了窄带通信中 3D 人脸模型的自动创建,通过特征自动提取和人脸特征边界的跟踪,对标准人脸模型进行修改,实现对特定人脸的表达,并实现人脸表情的编码和通信.1993 年与 1995 年 Lee 等人^[23,12]提出了一种自动建立人脸物理模型的精确算法,根据眼睛、鼻子和耳朵等人脸特征,寻找网格拓扑中的对应位置,实现模型网格到高分辨率人脸照片的覆盖和对准,根据解剖结构和生物力学特性,在人脸三角形几何网格基础上合成具有非线性的弹性分层组织结构.1996 年 Koch 等人^[17]建立了基于有限元分析的人脸几何模型,对于面部受损的病人进行脸部外形的预测和修复.1997 年 Sannier 等人^[3]提出了一个纹理匹配模型,实现真实人脸向特定模型的克隆.1997 年晏洁^[4]在已知的参考人脸中性模型与参考人脸基本表情模型的基础上,根据特定人脸的不同特征,经过整体与局部两种变换,完成从参考人脸中性模型到特定人脸中性模型的修改.

近几年相关的人像处理技术研究趋势与先进技术还包括:1997 年 O'Toole^[43,44]提出了一种人像漫画生成算法,利用人脸的差异性 (distinctiveness) 与可辨识度 (recognizability) 之间的关系,实现 3D 人像的衰老化与年轻化的变换,变换中既有人脸纹理的皱纹变化也有形状变化,动画效果有效且非常逼真.1998 年 Pighin 等人^[6]开发了一套系统,通过多幅照片实现人脸三维重建,可以生成具有相当真实感的三维人脸模型和人脸表情.1998 年 Jones 等人^[5]提出的多维渐变模型 (Multidimensional morphable models),是一种基于线性对象类 (Linear object class) 的统计模型,算法中直接在高维空间进行模型匹配和参数优化.1998 年 Vetter^[39]结合线性对象类模型与 3D 人像模型研究了从单视点图像到多视点新图像的合成技术,可以对非完整信息进行统计预测,并且利用 3D 模型增加个体特征.1999 年 Zicheng Liu 等人^[7]提出用人脸视频图像作为输入,用户只需在其中两帧图像中各标出 5 个特征点,就可以自动生成三维人脸模型,同时生成人脸动画.1999 年 Blanz 与 Vet-

ter^[35,36]提出的渐变模型 (Morphable model),把人脸样本数据的形状和纹理映射到一个矢量空间,先对人脸图像进行 3D 建模,再产生生动的图像渐变 (morphing) 效果.这种方法能够调节人脸模型变形的自然性和流畅性,输入简单,数据冗余度低,真实反映物体本质特征,但建库较为复杂,且人脸库也只局限于同一年龄层,离线操作的计算量很大.1999 年 Jeffrey 等人^[38]利用支持向量机 (SVM) 建立人脸多视点间的非线性变化模型,利用自然图像的内在因素以及 SVM 分类算法产生多视点检测,通过增加额外姿态信息的支持向量扩展了 SVM,生成姿态估计分配模型.2000 年 Boissieux^[42]研究了人类衰老的皮肤模型与衰老中皱纹的变化形态,对于人脸衰老化的动画仿真研究有较高的借鉴意义.2001 年 Cootes 等人^[37]提出了一种基于统计学习的 AAM 模型 (Active appearance model),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.AAM 模型算法对于可变形对象的跟踪定位快速、精确且具有较好的鲁棒性.2001 年 Liu^[41]等人提出表情比例图 (ERI) 的概念,结合人脸图像的稀疏特征表达,获取表情变化前后的形状变形与纹理变化细节信息,并将这种信息映射到其它人脸图像中实现逼真表情变换.2001 年邹北骥等人^[25]建立了一个人脸弹性体物理模型,并实现了一个交互表情编辑系统 HUFACE.2002 年张强等人^[59~68]提出并实现了一个名为 FACETRANSFER 的基于稠密特征对应的人像属性变换系统.

本文分为上下两篇,主要内容包括:建立了一种基于稠密特征对应的形状与纹理分离的人像线性统计模型,提出了相应的人脸图像矢量化表达方法,以及具有强鲁棒性的模型匹配算法,该模型结合人脸的共性与特性信息,为具有真实感的人像变换技术提供了理论基础.在上篇中提出了基于稠密特征对应的人脸表情合成与克隆算法,既考虑人像表情的形状变化,又考虑到纹理的相应变化,对于基本表情以及复杂表情都有效地实现了真实感的克隆与合成;将人像年龄变换分解成年年轻化变换和衰老化变换两部分,在定义人脸年龄图像空间的基础上,结合心理学中对人脸认知的相关假说,提出了基于平均脸的人脸图像年轻化变换算法;利用衰老纹理比率图投影技术,以及形状和纹理的差异性渐变技术,提出了基于衰老纹理映射的人脸图像衰老化变换算法.在一定的模板细腻程度以及人脸样本图像的数量要求的前提下,算法可以得到有效的主客观评价结果.在非完整信息恢复与重构的研究中,本文在下篇^[68]中主要侧重于非完整人像的视点信息的重构与变换问题,提出了一种视点续变技术框架,同时讨论了人像库的建立技术,从多方面对现有人像库进行分类,并在实际经验的基础上从方法、重要性以及关键问题等几个角度探讨人像库的建立过程.

2 人像处理的原理与技术

2.1 人脸建模

人脸具有复杂的物理和生理结构,对人脸属性的分析和理解涉及包括心理学和行为学等多领域知识.人脸图像描述

了人脸的多种属性包括人脸形状、皮肤纹理、五官位置、毛发分布等复杂个体形态特征的差异,以及人物性别、人脸表情、衰老程度、骨骼化、肉感化等生物形态特征的差异.因此人脸图像的建模就是对人类面容及其生物形态共同特性的数学表达.人脸模型的创建方法通常包括如下几点^[13,24]:

(1) 直接通过激光扫描或光栅法获得三维人脸模型.该方法的设备比较昂贵,细节效果也不够好,但是模型的控制能力与可操作性强,图像信息与个性信息完整;

(2) 利用一组平行的二维切层图像(CT图像、MRI图像等)进行边界跟踪,重构三维人类脸部模型.二维切层图像越细致,模型重构的效果越佳;

(3) 利用两张或两张以上不同视点(如正面与侧面90度或两帧序列图像)的二维平行投影照片定义人脸表面参考点的三维坐标,实现三维人像重构和纹理贴图.难点在于如何获取准确的三维深度信息,以及如何设计自动化的投影矩阵计算方法;

(4) 从一个固定的具有人脸共性拓扑结构的初始模型变形获得,交互式的调整模型中的特征点位置,逐渐逼近特定人脸的形状与纹理特征,适用于人脸动画领域,真实性较低;

(5) 利用插值技术从两个已知的人脸模型合成一个新模型.根据特定人脸照片对新模型进行局部修改与逼近,实现不同表情的组合与动画;

(6) 建立人脸二维图像样本数据库,利用线性对象类的统计学理论对人脸空间进行数学描述,总结出人脸对象的一般规律,对特定人像进行模型匹配与表达.该方法可以结合不同属性特征的人像数据库实现不同的人像处理效果.

人脸模型^[24]从表达效果的角度通常分为三维模型与二维模型两大类;从人脸建模的原理可以分为参数模型^[8-10]、肌肉模型^[11-15]、几何模型^[16-20]、物理模型^[12,21-25]、生理模型^[26-29]和统计模型^[30-38]等六大类.

(1) 参数模型.在人脸图像中定义一个点特征的特征集合,通过选择适当的参数组合,改变参数值实现脸部变形和人脸面部表情动画;

(2) 肌肉模型.充分强调了肌肉层的弹性特征,以及肌肉动力对脸部表情的控制,提供了充分的人脸解剖信息,克服了参数模型中缺少的人脸动画机制的理论依据;

(3) 几何模型.侧重于对几何曲面的光滑性以及拓扑结构的研究,将人脸表示为几何多边形(如三角形)或多形体(如椭球体),对不同的几何网格区域分别进行处理;

(4) 物理模型.是几何模型的拓展,并在某种程度上考虑到肌肉模型的机理.该模型主要采用多层的弹性网格结构,将皮肤划分为皮肤层、皮下组织层和肌肉层,利用约束性有限元网格表达皮肤的物理运动特性;

(5) 生理模型.在传统物理几何模型的基础上,充分考虑到人脸的生理特征和解剖结构,利用人体测量学理论、神经组织学理论、解剖学理论、心理学理论、生理学理论和统计学理论等,扩展了参数模型、肌肉模型、物理模型和几何模型的理论基础;

(6) 统计模型.基本思想是认为一幅人脸图像库中的样本

图像或者一幅库外的新图像,都可以用库中的样本图像的线性组合来表示.这种表示是建立在一个人脸空间上的,该人脸空间中的基底可以用原型样本图像表示,也可以由其他理论获得的抽象基底表示(如PCA主元等).

2.2 人像表情变换

人脸表情是人类对自身脸部行为姿态的一种主观定义.人脸表情千差万别,据心理学家研究表明,人脸能够产生大约55,000种不同的表情,其中仅仅30多种具有代表性的人脸表情可以通过人类的自然语言进行定义和描述,大部分人脸表情是无法用语言表达的^[59].目前人脸表情的计算机仿真技术主要集中在四个方面:

(1) 表情参数方法.在人像上建立虚拟的肌肉(伪肌肉)、皮肤、骨骼等人脸组织结构,定义特定的人脸动画控制点,利用特征参数点的位置变化,控制肌肉运动,产生不同的人像表情^[45-52].这种方法的人脸表情表达能力有限,依赖于动画操作者对于人脸表情的理解与认知.表情的仿真效果僵硬,不适用于真实感人像处理,而且应用中需要大量繁琐的人工操作,计算量大;

(2) 表情插值方法.通过对同一对象,两种不同的表情图像进行图像的插值与融合(morphing),生成中间的表情图像^[1,67,53,40].这种方法的实现需要给出同一个人的两幅不同的表情图像,这就大大限制了其应用范围,只能作为一种辅助的表情合成方法;

(3) 表情拼合方法.根据人脸表情的生理机制,定义出多个人脸表情的基本单元^[11],人脸的特定表情是由这些表情单元的不同动作组合而成,表情动画过程是通过对这些表情基元的运动控制实现的.这种方法计算开销比较大,而且主观因素较多,表情合成的准确性低,不能合成人脸表情的细微效果;

(4) 表情克隆方法.是一种表情映射技术,即将其它人脸对象的表情映射到特定人脸上^[41,60].为了对表情细节进行刻画,需要同时考虑表情变换过程中人脸特征点的位置改变,以及诸如皱纹等纹理细节的改变.表情克隆方法可以应用于真实感人像表情处理中.

总之,人脸模型越复杂,人脸表情合成越有效,但计算复杂度也越高;反之,人脸模型越简单,人脸表情合成效果越有限,计算复杂度也越低.人脸表情变换方法应该同时考虑计算复杂度以及表情合成质量两个方面的要求,尽量在降低算法复杂度与提高表情动画效果间寻找合适的折中点,同时应该考虑实际应用中不同场合对人脸表情变换的需求.

2.3 人像年龄变换

虽然人类年龄是一个与时间相关的客观概念,但是即使相同年龄的人所呈现的外貌特征也会因个人生存环境、生活习惯等各种因素的影响存在很大的差异,因此我们很难从一幅人脸图像判断出一个人的准确年龄.对于计算机而言,辨别年龄则更是一个困难的问题.而另一方面,对于两幅人脸图像,我们却可以容易地辨认出谁更年轻或谁更年老,而且可以描述出人类年轻与衰老时的基本特征,例如皮肤的光滑程度,皱纹与斑点的多少,骨骼的形状,肌肉的松弛程度等.利用图像处理技术实现人脸图像的衰老化与年轻化变换是人像处理

领域的前沿课题,这种技术极大地依赖于生理学、认知学、心理学等学科间的理论支持与相互融合。下面主要介绍本文提到的人像衰老变换实现的一些理论依据^[59,61,62]:

(1) 平均脸与年轻化. 平均脸是人脸共性特征的集合,在人脸空间中通常作为坐标原点,在图像表达中是五官信息的平均分布.“平均即是美”(Averageness is attractiveness)理论^[54]说明了一类人群中所谓最美的人脸图像的基本特征,即平均化.德国的研究人员通过实验支持了这一假说,即从一组人脸图像样本中得到平均脸是最美的.求取平均脸型的操作,可以消除特定人脸的非对称与不规则脸型;求取平均纹理的操作,可以平滑特定人脸的皱纹与疤痕信息,使得整个脸部的皮肤看起来有更年轻的感觉^[59,61];

(2) 人脸的差异性、典型性与可辨识性.人脸差异性(distinctiveness)描述的是特定人脸与平均脸的区别程度.人脸典型性(Typical)描述的是特定人脸与平均脸的接近程度.人脸可辨识性(recognizability)描述的是人类观察者对于特定人脸的可识别程度.差异性越大的人脸越容易被人所识别,而越典型的人脸则越容易在识别时被混淆^[55~57].文献^[43],^[44]表明,人脸的差异性增加的同时会使人像产生衰老的效果;

(3) 口述记忆画像^[69]技术中人像的年龄变化修整,以及人脸衰老与年轻化图像细节规律.少年时期,人脸脸型特征表现为小脸、短圆、小下巴、皮肤光滑有弹性;青年时期,人脸脸型特征表现为竖向拉长,下巴稍尖,柔和感消失,边沿变硬,皮肤较光滑,伴随青春痘等皮肤纹理变化;中年时期,人脸脸型特征为边沿棱角更清晰,颧骨突出两颊有下凹,肌肉出现下垂,下颌稍前凸,皮肤光泽度降低,颜色变深,有皱纹出现;老年时期,人脸脸型特征表现为两颊下凹明显,面部肌肉下垂加重,胖人双下巴及下颌前凸明显,皮肤光泽度更低,皱纹多,伴随老年斑等皮肤纹理变化.图1示出了公安刑侦技术在目击证人口述画像过程中总结出的人类衰老后面部特征的变化规律.

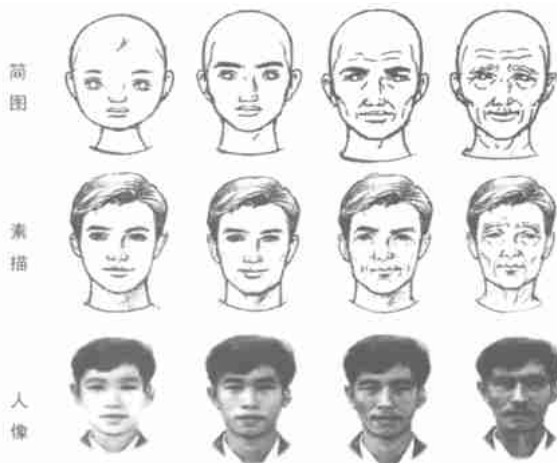


图1 人脸衰老过程中的特征变化描述和画像效果^[69]

3 基于稠密特征对应的表情与年龄变换技术

人像处理所面对的首要问题是如何对人脸图像进行合适的重新表达.只有将原始人脸图像经过适当的方法重新表达

后,才有利于后续的人像处理.本文从模式识别的角度出发,同时比较了常见的人脸图像表达方法,采用了基于稠密特征对应的人脸图像表达方法,先利用图像表达完成特征提取,将人脸数据空间转换到一个特征空间,然后结合具体的特征处理方法完成不同的处理任务,如表情及年龄变换等.

由图2可以看出,人脸图像的两个主要特征就是形状(shape)和纹理(texture),形状包括人脸的轮廓大小和五官的位置等外观信息,纹理则表示基于该形状特征的图像灰度信息.基于稠密特征对应的图像表达方法就是通过人脸图像的对准计算将人脸图像分解为两个主要的特征矢量:形状矢量和纹理矢量^[33,34,64,66].



图2 人脸图像的形状与纹理

3.1 人像预变换、对准与表达

在建立精确的像素级人像对准前,需要先进行基于主要点特征的预变换^[59,64,66],实现特定人像在方位和整体尺度上与参考人像或模型逼近,消除图像在空间上的旋转现象.这些主要的点特征通常包括两个瞳孔中心、鼻尖和两个嘴角等.

对预处理之后的人脸图像进行人像对准计算,实现人脸图像的矢量化表达.主要包括对特征点或特征线段的手动或自动标定,以及根据特定人像与参考人像或参考模型中对应特征的映射关系,根据人像中每一个像素点位置与各个标定特征的位置关系,建立像素点的位移矢量场,并进行新的像素位置排布,从而产生像素级人像对准效果.

基本技术的实现算法可分为如下4步:

(1) 利用边缘检测、分割、统计直方图以及滤波等图像处理方法分离出人脸照片中的人像区域;

(2) 对特定人像和参考人像或参考模型中主要特征点的手动或者自动标定,选择两个瞳孔中心、鼻尖或两个嘴角中的有效特征点为参考,实现特定人像中人脸对象的旋转,平移和尺度变换;

(3) 在预对准后的特定人脸图像中,确定人脸对象的6个部分:脸庞、眼睛、眉毛、鼻子、嘴巴和耳朵.利用特征线段(或网格)将这几部分的位置进行标定;利用相同的方法,在参考图像或参考模型中标定相同特征数目的参考特征集合,确定特征的映射关系;

(4) 逐一经过参考图像或参考模型的每一个像素点,在预对准和形状标定后的特定人像中确定对应像素位置.可以定义不同的像素位置映射规则并采用不同的映射技术,如多线段 Warping 技术^[53].

本文采用的人像表达方法是同时考虑形状与纹理的矢量化图像表达.矢量化的表达可以认为是将特征点集(如图像的像素点)依次排列成一个有序的矢量,通过相对于一个确定的参考图像来完成特征对准,从而使每个矢量的相应维都代表同一个特征点.不同人像间的形状与纹理都存在着差异性,形状差异性的矢量化表达描述了特定人像与参考人像特征点位

置的差异值,即每个像素点存在一个 (x, y) 位置偏移量. 纹理矢量描述的则是消除了人脸形状差异后的纹理信息,即经过对准后的特定人脸纹理. 图 3 描述了人像对准与表达效果. 其中表达人脸形状的特征线段的位置采用手动标定,对准后的特定人脸与参考人脸形状一致,位移场显示特定人脸图像经过对准后各个像素的位移矢量,也即人脸表达后的形状矢量.

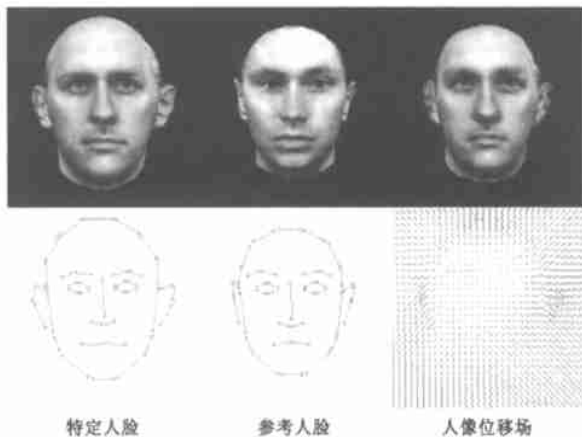


图 3 人脸图像对准与人像表达效果

3.2 基于稠密特征对应的人脸表情合成与克隆

为了实现具有真实感的人脸表情变换,并且可以对无法用语言表达的复杂表情进行仿真,我们提出了基于稠密特征对应的人脸表情合成与克隆算法. 在处理中同时考虑表情变化前后人脸形状与纹理信息的变换,并假设人脸类属于 Lambertian^[58,41]表面类,借助 Liu 等人提出的表情比例图 (Expression ratio image)^[41]的概念克隆皱纹、阴影、遮挡等细节变化,并充分利用了稠密特征对应人像模型的矢量化机制.

假定已知某特定人脸的中性(自然)表情图像为 A , 特定表情图像为 A' , 待变换人脸的中性(自然)表情图像为 B , 我们的目标是合成具有类似于 A' 表情的人脸图像 B' . 算法中首先对人脸图像进行预处理,包括基本几何形状与位置的仿射变换 P 合成表情图像后,再经过一个相应的几何逆变换 P^{-1} 来得到最终的结果.

基于稠密特征对应的人脸图像的表情合成算法可以如下描述^[59,60]:

- (1) 根据参考图像的脸型对 A 、 A' 和 B 进行图像分解,得到各自对应的图像表达 $(shpA, texA)$ 、 $(shpA', texA')$ 和 $(shpB, texB)$, 其中 shp 表示形状矢量, tex 表示纹理矢量;
- (2) 计算表情变换后的新的形状矢量: $shpB' = shpB + Shp$, 其中 $Shp = shpA' - shpA$;
- (3) 计算表情变换后的新的纹理矢量: $texB' = texB \cdot R$, 其中 $R = texA' / texA$;
- (4) 根据 $(shpB', texB')$, 利用前向 Warp, 重构得到图像 B' .

我们可以将不同的表情样本图像制成模板,即在离线状态下计算各种表情变化信息: $Shp = shpA' - shpA$ 和 $R = texA' / texA$. 这样可以加快实时的处理速度.

Liu^[41]提出的算法中,是直接对图像数据进行处理. 为了完成特征对应,在处理阶段需要完成三次 Warp 操作,并且对表情比例图进行了滤波操作. 与此相比,我们的算法采用了稠密特征的图像表达方法,将图像分解成形状和纹理两部分矢量,因而在处理阶段无须对 ERI,权重图以及 B 图像进行 Warp 操作,只需在最后重构图像时使用一步前向 Warp 的操作. 在这种图像表达下,处理表情变换不但和人们对表情的普遍认识相符,而且在线处理的数据量也大大降低,极大的提高了实时处理的速度. 并且不需要对表情比例图进行滤波操作.

3.3 基于平均脸的人像年轻化与衰老化

根据前面提到的认知学与心理学的相关理论,我们提出了基于平均脸的人像年龄变换方法. 核心思想如下^[59,61,62]:

为特定人脸对象建立一个人脸年龄属性空间 F , 空间中的每个点表示一幅人脸图像. 将平均脸作为属性空间的原点 O (当然也可以用待处理人像一致的年轻化或衰老化图像作为原点,这里采用的是一种泛化操作),而空间中的其它点所表示的图像是与原点图像有差异的人像. 离原点的距离越远表示该人脸图像的差异性越大,反之越小. 定义给定的特定人脸图像在该空间中固定的坐标位置点为 Q , 该点与坐标原点连线为 OQ . 图 4 描绘了人像年龄属性空间的示意图.

定义 1 Q 顺着 OQ 向原点方向移动,得到人脸年轻化图像;向远离原点方向移动,得到人脸衰老化图像.

本文提出的人像年龄变换包括人像的年轻化变换与衰老化变换两个部分. 从图像处理的角度讨论,两部分的算法思想基本一致. 记平均脸图像的形状矢量和纹理矢量为 $(S_{Average}, T_{Average})$.

- (1) 在参考图像下,对新的人脸图像进行分解,得到对应的形状和纹理矢量,记为 (shp, tex) ;
- (2) 计算新的人脸图像在形状和纹理方面与平均特征之间的差异,记为 (shp', tex') ;
- (3) 利用参数 α 和 β 来调整新的人脸图像与平均脸之间的形状特征及纹理特征的差异程度. $shp'' = \alpha \cdot shp + \beta \cdot shp'$, $tex'' = \alpha \cdot tex + \beta \cdot tex'$, 其中 $0 < \alpha, \beta < 1$;
- (4) 计算差异程度变化后,新的人脸图像的形状矢量及纹理矢量 (shp'', tex'') , 其中 $shp'' = \alpha \cdot shp + \beta \cdot shp'$, $tex'' = \alpha \cdot tex + \beta \cdot tex'$. 根据 (shp'', tex'') , 通过前向 Warp 重构出变换后的人脸年轻化或衰老化图像.

应用上面的方法,在人脸年轻化变换过程中,纹理变换有整体平滑的效果,这正好符合年轻时人脸的皮肤特征,光滑细腻;但是在人脸衰老化变换过程中,脸部会出现皱纹、斑点等细节变化,利用上面的算法对这些细节信息的勾画能力有限.

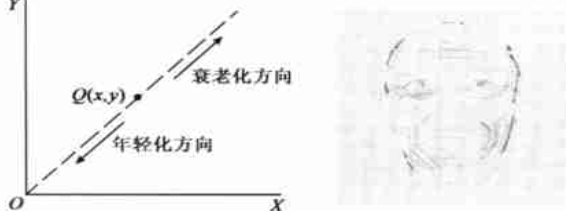


图 4 人像年龄属性空间 图 5 衰老纹理模板^[59]

因此我们增加了一个衰老纹理的模板来配合如上算法,图 5 示出了一个衰老纹理模板的可视化示意图.模板的制作方法如下所述.

定义 2 给定一个老年正面样本人像纹理矢量 O_{tex} ,对于其中眼睛、鼻子与嘴巴以外的部分进行低通滤波,得到一个平滑的纹理矢量 Y_{tex} 即为人像年轻化纹理矢量.

定义 3 假设衰老纹理模板定义为 W ,则 $W = O_{tex} / Y_{tex}$.

这样在步骤(4)中的纹理变换公式变为: $tex = G(\cdot, W) \cdot tex$,其中 $G(\cdot, W)$ 是一个与 W 相关的比例函数,用于调节衰老纹理的变换效果.其具体函数式可以根据研究需求自定义,这里的研究需求包括年龄与人像表达效果的对应性,以及观察者对于人像衰老化的主观评价等.

4 实验结果与分析

实验数据是本文下篇中介绍的 MPI 和 AIAR V1.0 (Expression), AIAR V2.0 (Aging), AIAR V3.0 (View) 4 个人像库中的样本图像.实验分别用于验证本文所提出的基于稠密特征对应的人脸表情合成与克隆算法和基于平均脸的人脸图像年轻化与衰老化变换技术.

图 6 示出了利用基于稠密特征对应的人脸表情合成与克隆技术合成的人脸表情图像.图 6(a) 显示了 8 幅 AIAR V1.0 (Expression) 表情人像库中的表情图像,以此为克隆模板,图 6(b) 分别显示了对 MPI 人像和 AIAR 人像的表情变换结果.其中第一行左边第一幅人像为 AIAR 人像库中的东方人脸图像,右边 4 幅为经过表情克隆所合成的人脸图像.第二行左边第一幅人像为 MPI 人像库中的高加索人脸图像,右边 4 幅同样为经过表情克隆所合成的人脸图像.从实验结果可以看出,这种方法可以有效的捕捉和刻画表情的纹理细节变化,合成的表情逼真、自然.该算法不依赖于人种的限制,且对自然语言无法定义的人脸复杂表情(如最后一个撇嘴动作)同样可以有效地合成.

图 7 示出了利用基于平均脸的人像年轻化与衰老化技术合成的人脸图像.第一行中间人像为 AIAR 人脸原始图像,左



(a) AIAR V1.0(Expression)表情人像库中的表情图像



(b) AIAR 人像与 MPI 人像的表情变换与克隆

图 6 基于稠密特征对应的人脸表情合成与克隆

图显示人脸的年轻化效果,右图显示衰老化效果.这里采用的衰老模板,是从 AIAR V2.0 (Aging) 衰老人像库中老年人像中提取的.第二行图像分别显示 MPI 人像衰老化效果和 AIAR 人像年轻化效果.从合成结果可以看出文中提出的方法是有效的,衰老化与年轻化效果都比较明显.我们随机选择 20 位观察者,对人像合成效果进行评价.20 位观察者一致认为人像有明显的年轻变化和衰老变化.这种方法不受限于图像库以及人种等因素的限制,具有较强的鲁棒性.



图 7 基于平均脸的 MPI 和 AIAR 人像年轻化与衰老化

5 结论

本文在建立基于稠密特征对应的人像线性统计模型基础上,进行了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,基于衰老纹理映射和差异性的人脸图像衰老化变换算法,以及下篇中的非完整视点信息恢复与重构算法.对 MPI 人像库与 AIAR 人像库中的样本图像进行的实验,验证了本文提出的算法和技术的合理性,表情细节逼真自然,衰老化与年轻化视觉效果直观,该技术的有效性与其鲁棒性的优点.

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 人像库建库过程中给予帮助的西安交通大学人工智能与机器人研究所所有成员表示由衷感谢;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] Parke F I. Computer generated animation of faces [D]. University of Utah, Salt Lake City, UT, UTEC-CSC-72-120, 1974.
- [2] Takaaki Akimoto, Yasuhito Suenaga, Richard S Wallace. Automatic creation of 3D facial models[J]. IEEE Computer Graphics & Applications, 1993, 13(5): 16 - 22.
- [3] Sannier G, Magnenat-Thalmann N. A flexible texture fitting model for virtual clones [A]. Proceedings of Computer Graphics International [C]. Belgium: IEEE Computer Society, Hasselt-Diepenbeek, 1997. 167 - 176.
- [4] 晏洁. 从一般人脸模型到特定人脸模型的修改[J]. 计算机工程与科学, 1997, 19(2): 21-24.
- [5] MJ Jones, T Poggio. Multidimensional morphable models [A]. The 6th International Conference on Computer Vision [C]. Bombay: IEEE Computer Society, 1998. 683 - 688.

- [6] F Pighin, J Hecker, D Lischinski, R Szeliski, DH Salesin. Synthesizing realistic facial expressions from photographs [A]. Computer Graphics Annual Conference Series, SIGGRAPH '98 [C]. Florida: ACM SIGGRAPH, Orlando, 1998. 75 - 84.
- [7] Liu Z, Zhang Z, Jacobs C, Cohen M. Rapid modeling of animated faces from video [R]. Beijing: Microsoft Research, 1999.
- [8] Parke F I. A parameterized model for facial animation [J]. IEEE Computer Graphics and Application, 1982, 2(9): 61 - 68.
- [9] E Haratsch, J Ostermann. Parameter based animation of arbitrary 3D head models [A]. Proceedings of Picture Coding Symposium 1997 (PCS97) [C]. Berlin, Germany: VDE Verlag, 1997. 81 - 84.
- [10] Y Shan, Z Liu, Z Zhang. Model-based bundle adjustment with application to face modeling [A]. International Conference on Computer Vision [C]. Vancouver, Canada: IEEE Computer Society, 2001. 644 - 651.
- [11] Paul Ekman, Wallace V Friesen. Manual for the Facial Action Coding System [M]. Palo Alto, California: Consulting Psychologist Press, 1978.
- [12] Lee Yuencheng, Demetri Terzopoulos, Keith waters. Realistic modeling for facial animation [A]. Computer Graphics Annual Conference Series, SIGGRAPH '95 [C]. Los Angeles, California: ACM SIGGRAPH, 1995. 55 - 62.
- [13] 詹永照, 宋顺林等. 脸部表情动画建模方法的研究与实现 [J]. 软件学报, 1998, 9(2): 81 - 85.
- [14] Shimoda H, Yang D, Yoshikawa H. Dynamic facial expression generation by using facial muscle model [A]. Proceedings of World Multiconference on Systemics, Cybernetics and Informatics [C]. Orlando, Florida, 2000. 56 - 61.
- [15] Rolf M Koch, Markus H Goss, Friedrich R Carls, Daniel Fvon Buren, George Fankhauser, Yoav I H Parish. Simulating facial surgery using finite element methods [A]. Computer Graphics Annual Conference Series, SIGGRAPH '96 [C]. New Orleans, Louisiana: ACM SIGGRAPH, August, 1996. 421 - 428.
- [16] Jean Paul Gouret, Jamal Kham Lichi. A model for compression and classification of face data structures [J]. Computer & Graphics, 1996, 20(6): 863 - 879.
- [17] R M Koch, M H Goss, F R Carls, D F von B uren, G Fankhauser, Y Parish. Simulating facial surgery using finite element methods [A]. Computer Graphics Annual Conference Series, SIGGRAPH '96 [C]. New Orleans, Louisiana: ACM SIGGRAPH, 1996. 421 - 428.
- [18] Lee W, Kalra P, Magnenat Thalmann N. Model based face reconstruction for animation [A]. MMM '97 [C]. Singapore, 1997. 323 - 338.
- [19] 尹宝才, 高文. 利用 Bézier 曲面的面部表情和口形几何造型 [J]. 计算机学报, 1998, 21(Suppl.): 347 - 350.
- [20] F Pighin, J Hecker, D Lischinski, R Szeliski, DH Salesin. Synthesizing realistic facial expressions from photographs [A]. Computer Graphics Annual Conference Series, SIGGRAPH '98 [C]. Orlando, Florida: ACM SIGGRAPH, 1998. 75 - 84.
- [21] A Yilmaz, M Shah. Automatic feature detection and pose recovery for faces [A]. Asian Conf. on Computer Vision (ACCV) [C]. Melbourne, Australia, 2002. 284 - 289.
- [22] Manjula Patel, Willis P J. FACES: facial animation, construction and editing system [A]. Computer Graphics Annual Conference Series, SIGGRAPH '91 [C]. Las Vegas, Nevada: ACM SIGGRAPH, 1991. 33 - 45.
- [23] Lee Y C, Terzopoulos D, Waters K. Constructing physics based facial models of individuals [A]. Proceedings of Graphics Interface '93 [C]. Toronto, 1993. 1 - 8.
- [24] 付昀. 人脸建模技术 [R]. 西安: 西安交通大学人工智能与机器人研究所, 2003.
- [25] 邹北骥, 彭永进, 伍立华. 基于物理模型的特定人脸模型与表情动画模型 [J]. 系统仿真学报, 2001, 13(3): 385 - 387.
- [26] Y Wu, N M Thalmann, D Thalmann. A plastic-visco-elastic model for wrinkles in facial animation and skin aging [A]. Proc. 2nd Pacific Conference on Computer Graphics and Applications [C]. Beijing: Pacific Graphics, 1994. 201 - 213.
- [27] D DeCarlos, D Metaxas, M Stone. An anthropometric face model using variational techniques [A]. Computer Graphics Annual Conference Series, SIGGRAPH '98 [C]. Orlando, Florida: ACM SIGGRAPH, 1998. 67 - 74.
- [28] M P Young, S Yamane. Sparse population coding of faces in the inferotemporal cortex [J]. Science, 1992, 256: 1327 - 1331.
- [29] Takashi T, Takio K. A self-organization model of feature columns and face responsive neurons in the temporal cortex [A]. 2001 Int. Joint Conf. On Neural Networks (IJCNN2001) [C]. Washington, DC, 2001. 82 - 87.
- [30] S Ullman, R Basri. Recognition by linear combinations of models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13: 992 - 1006.
- [31] 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.
- [32] 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.
- [33] David Beymer. Vectorizing face images by interleaving shape and texture computations [R]. A I Memo 1537. Massachusetts: MIT, 1995.
- [34] Michael J Jones, Tomaso Poggio. Model-based matching by linear combinations of examples [R]. A I Memo 1583. Massachusetts: MIT, 1996.
- [35] 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.
- [36] 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.
- [37] 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.
- [38] Jeffrey Ng, Sing Kwong, Shaogang Gong. Learning support vector machines for a multi-view face model [A]. Proceedings of the Tenth British Machine Vision Conference, BMVC99 [C]. British, Nottingham, 1999. 503 - 512.
- [39] Vetter T. Synthesis of novel views from a single face image [J]. International Journal of Computer Vision, 1998, 28(2): 103 - 116.
- [40] S M Seize, C R Dyer. View morphing [A]. Computer Graphics Annual Conference Series, SIGGRAPH '96 [C]. New Orleans, Louisiana: ACM SIGGRAPH, August, 1996. 21 - 30.
- [41] 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.
- [42] Boissieux L, Kiss G, Thalmann N M, Kalra P. Simulation of skin aging and wrinkles with cosmetics insight [A]. Proceedings of Eurographics Workshop on Animation and Simulation (EGCAS 2000) [C]. Interlaken: EUROGRAPHICS, 2000. 21 - 22.
- [43] O'Toole A J, Vetter T, Volz H, Salter. As we get older, do we get more distinct[R]? MPI, Germany, Memo TR49, 1997.
- [44] O'Toole A J, Vetter T, Volz H, Salter E M. Three-dimensional caricatures of human heads: Distinctiveness and the perception of facial age [J]. Perception, 1997, 26: 719 - 732.
- [45] Platt S M. A structural model of the human face [D]. University of Pennsylvania, Dept. of CIS, 1985.
- [46] Waters K. A muscle model for animating three-dimensional facial expressions [A]. Computer Graphics Annual Conference Series, SIGGRAPH '87 [C]. Anaheim, California: ACM SIGGRAPH, 1987. 21 (4): 17 - 24.
- [47] Terzopoulos D, Waters K. Physically-based facial modeling, analysis, and animation [J]. Journal of Visualization and Computer Animation, 1990, 1(4): 73 - 80.
- [48] Nahas M, Hutric H, Rioux M, Domey J. Facial image synthesis using skin texture recording [J]. The Visual Computer, 1990, 6(6): 337 - 343.
- [49] Nahas M, Hutric H, Saintourens M. Animation of a B-spline figure [J]. The Visual Computer, 1988, 3(5): 272 - 276.
- [50] Singh K, Fiume E Wires. A geometric deformation technique [A]. Computer Graphics Annual Conference Series, SIGGRAPH '98 [C]. Orlando, Florida: ACM SIGGRAPH, 1998. 405 - 414.
- [51] Cohen M, Massaro D. Synthesis of visible speech [J]. Behavior Research Methods: Instruments & Computers, 1990, 22(2): 260 - 263.
- [52] Kalra P, Mangili A, Thalmann N M, Thalmann D. Simulation of facial muscle actions based on rational free form deformations [J]. Computer Graphics, 1992, 11(3): 59 - 69.
- [53] Beier T, Neely S. Feature-based image metamorphosis [J]. Computer Graphics, 1992, 26(2): 35 - 42.
- [54] Braun C, Gruendl M, Marberger C, Scherber C. Beautycheck-causes and consequences of human facial attractiveness (Summary) [R]. Germany: The German Students Award, Regensburg and Rostock, 2001.
- [55] Light L L, Stuart F K, Hollander S. Recognition memory for typical and unusual faces [J]. Journal of Experimental Psychology: Human Learning and Memory, 1979, 5: 212 - 228.
- [56] O'Toole A J, Deffenbacher K, Valentin D, Abdi H. Structural aspects of face recognition and other race effect [J]. Memory and Cognition, 1994, 22: 208 - 224.
- [57] Weiss Y, Edelman S. Representation of similarity as a goal of early visual processing [J]. Network, 1995, 6: 19 - 41.
- [58] A Shashua, T Riklin-Raviv. The quotient image: class-based rendering and recognition with varying illuminations [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(2): 129 - 139.
- [59] 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.
- [60] 张强, 郑南宁, 张婷, 付昀, 卓峰. 基于稠密特征对应的真实感人脸图像表情合成 [R]. 西安: 西安交通大学人工智能与机器人研究所, 2003.
- [61] 张强, 郑南宁, 卓峰, 张婷, 付昀. 基于平均脸的人脸图像年轻变换研究 [R]. 西安: 西安交通大学人工智能与机器人研究所, 2003.
- [62] 张强, 郑南宁, 付昀, 张婷, 卓峰. 基于形状和纹理矢量的人脸图像衰老变换 [R]. 西安: 西安交通大学人工智能与机器人研究所, 2003.
- [63] 付昀. 人脸图像渐变算法研究 [J]. 中国图像图形学报 (A), 2003, (Special.) 8: 174 - 182.
- [64] 付昀. 基于原型样本线性组合的人脸图像 CD-SGD 鲁棒对准算 [R]. 西安: 西安交通大学人工智能与机器人研究所, 2002.
- [65] Fu Yun, Zheng Nanning. Sequential view synthesis from a single view [R]. AIAR. Xi'an, 2003.
- [66] Fu Yun, Zheng Nanning, Zhang Qiang. Non-uniform sampling based dynamic matching for the linear model of human faces [R]. AIAR. Xi'an, 2003.
- [67] Fu Yun, Zheng Nanning. Advanced framework for the error-resilient parameter analysis-synthesis system of facial animation [A]. 2003 IEEE International Conference on Systems, Man & Cybernetics (IEEE SMC '03) [C]. Washington, D. C. USA, 2003. 5. 4528 - 4534.
- [68] 付昀, 郑南宁, 张婷. 人脸的表情与年龄变换和非完整信息的重构技术 (下) [J]. 电子学报, 2003, 31(12A): 1963 - 1970.
- [69] 荣建群. "口述记忆组合画像" 年龄和表情变化的修整 [R]. 郑州: 河南省公安厅刑事科学技术研究所, 2002.

作者简介:



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。



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。