

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2013
Vol.18 No.

1

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2013年1月 第18卷 第1期(总第201期)

目次

综述

- 正面人脸图像合成方法综述 赵林, 高新波, 田春娜(1)
SAR 图像道路网提取方法综述 程江华, 高贵, 库锡树, 孙即祥(11)

图像处理和编码

- 头部缺失的 JPEG 文件碎片恢复 徐明, 黄立, 张海平, 徐建, 郑宁(24)
安全高效的可撤销指纹模板构造 喻建平, 张鹏, 王瑶, 杨懿竣(36)
基于残差的图像超分辨率重建 陈华华, 姜宝林, 刘超, 陈伟强, 陆宇, 张嵩(42)
旋转的 Wang Tiles 纹理合成算法 王继东, 庞明勇, 赵瑞斌(49)
基于圆形约束快速水平集的原生质体细胞分割 王晓飞, 庞全(55)

图像分析和识别

- 采用压缩传感的鲁棒的视频指纹方案 孙锐, 李超, 蒋飞云(62)
基于图像显著性的路面裂缝检测 徐威, 唐振民, 吕建勇(69)
基于局部熵的主动轮廓模型 潘改, 高立群, 赵爽(78)
基于算子的图像分解 李峰, 曾晓辉, 陈盛霞, 沈玉娟(86)

图像理解和计算机视觉

- 中值流辅助在线多示例目标跟踪 王德建, 张荣, 尹东, 张智瑞(93)
四叉树直方图的特殊方向关系表达 张珂, 王小捷, 靳越(101)

计算机图形学

- 协同进化的近似规则纹理合成 王相海, 陶兢喆(107)
反走样直线的灰度循环生成算法 牛连强, 张胜男, 钟玲(115)

地理信息技术

- 微博客蕴含交通信息的提取 张恒才, 陆锋, 陈洁(123)

-
- “计算机视觉前沿论坛”专栏征文通知 (130)

Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 18 No. 1 January 2013

Contents

Review

- Review of frontal face image synthesis methods Zhao Lin, Gao Xinbo, Tian Chunna(1)
Review of road network extraction from SAR images Cheng Jianguhua, Gao Gui, Ku Xishu, Sun Jixiang(11)

Image Processing and Coding

- Rrecovery method for JPEG file fragments with missing headers
..... Xu Ming, Huang Li, Zhang Haiping, Xu Jiang, Zheng Ning(24)
Secure and efficient scheme to construct a cancelable fingerprint template
..... Yu Jianping, Zhang Peng, Wang Yao, Yang Yijun(36)
Image super-resolution reconstruction based on residual error
..... Chen Huahua, Jiang Baolin, Liu Chao, Chen Weiqiang, Lu Yu, Zhang Song(42)
Texture synthesis using rotational Wang Tiles Wang Jidong, Pang Mingyong, Zhao Ruibin(49)
Protoplasm somatic cells segmentation based on circle dependent fast level-set segmentation Wang Xiaofei, Pang Quan(55)

Image Analysis and Recognition

- Robust video fingerprinting via compressed sensing Sun rui, Li Chao, Jiang Feiyun(62)
Pavement crack detection based on image saliency Xu Wei, Tang Zhenmin, Lv Jianyong(69)
Active contour model driven by local entropy energy Pan Gai, Gao Liqun, Zhao Shuang(78)
Operator-based image decomposition Li Feng, Zeng Xiaohui, Chen Shengxia, Shen Yujuan(86)

Image Understanding and Computer Vision

- Median flow aided online multi-instance learning visual tracking Wang Dejian, Zhang Rong, Yin Dong, Zhang Zhirui(93)
Expression of special directional relation based on quadtree histogram Zhang Ke, Wang Xiaojie, Jin Yue(101)

Computer Graphics

- Fast near-regular texture synthesis based on the concept of co-evolution Wang Xianghai, Tao Jingzhe(107)
Integral algorithm for generating anti-aliased straight line controlled by gray iteration
..... Niu Lianqiang, Zhang Shengnan, Zhong Ling(115)

Geoinformatics

- Extracting traffic information from massive micro-blog messages Zhang Hengcai, Lu Feng, Chen Jie(123)

中图分类号: TP391 文献标识码: A 文章编号: 1006-8961(2013)01-0001-10

论文引用格式: 赵林, 高新波, 田春娜. 正面人脸图像合成方法综述[J]. 中国图象图形学报, 2013, 18(1): 1-10.

正面人脸图像合成方法综述

赵林, 高新波, 田春娜

西安电子科技大学工程学院, 西安 710071

摘要: 随着多视角人脸识别技术需求的不断增长, 正面人脸图像合成成了热门的研究课题。然而, 从侧面人脸图像准确地合成出正面人脸是一个典型的逆向问题, 具有较大的挑战性。对目前正脸图像合成方法进行了系统总结, 介绍了几种典型的合成策略, 并对这些方法按原理分成了基于图形学的方法和基于统计学习的方法两类分别进行研究。此外, 还从算法复杂性、鲁棒性, 以及图像合成效果等方面对现有正面人脸图像合成算法进行了对比研究, 给出了未来可能的研究方向。

关键词: 人脸合成; 多视角; 逆问题; 综述

Review of frontal face image synthesis methods

Zhao Lin, Gao Xinbo, Tian Chunna

School of Electronic Engineering, Xidian University, Xi'an 710071, China

Abstract: With the increasing demand for multi-view face recognition techniques, frontal face image synthesis has been one of the most interesting research topics nowadays. However, it is a classic inverse problem to synthesize frontal images from profiles accurately, and there are some challenges. In this paper, we present a systematical summary of the current frontal image synthesis methods. Furthermore, some classic synthesis strategies are introduced. Moreover, according to the theories used, the methods are classified into two categories: graphic based methods and statistical learning based methods. Additionally, these approaches are compared in three aspects: complicity, robustness, and performance. At last, some potential pointers towards future research topics are given.

Key words: face synthesis; multi-view; inverse problem; review

0 引言

随着人脸识别技术不断走向成熟, 并逐渐走向商业化, 人们希望设计出能够适用于非受控环境下的人脸识别方法。所谓非受控环境指的是在人脸识别过程中系统不需要被识别人的合作, 即被识别人可以处于任何自由状态。而在实际应用中, 识别技术需要满足对环境光照的鲁棒性、对被识别人的视角和表情的鲁棒性^[1]。这就对人脸识别提出了很

高的要求, 成为限制人脸识别技术顺利走向商业化的瓶颈^[2]。在非受控环境的应用需求下, 人们希望通过图像变换技术来实现非受控环境与现有成熟识别技术的衔接。因而, 如何从一张任意光照、任意姿态、任意表情的人脸图像准确地合成出正常光照、正脸和正常表情的图像成为了一个热门而又富有挑战的研究课题。其中, 又以从任意侧脸到正脸的合成最为典型。本文将主要介绍正常光照、正常表情下正脸图像合成的研究进展。

一般来说, 人脸合成是利用一个或多个视角的

收稿日期: 2012-05-08; 修回日期: 2012-07-24

基金项目: 国家杰出青年科学基金项目(61125204); 国家自然科学基金项目(61172146); 陕西省自然科学基金(2009JM8004)

第一作者简介: 赵林(1988—), 男, 西安电子科技大学模式识别与智能系统专业博士研究生, 主要研究方向为模式识别、计算机视觉。

E-mail: sharling1988@gmail.com

人脸图像合成出所需视角的人脸图像。该问题可划分为两类:一是由一张正面的人脸图像合成其他侧面的人脸图像;二是由一个或几个侧面视角的人脸图像合成其正面的人脸图像。对于第 1 类问题,由于侧脸图像可以看成正脸图像经过形变压缩后的图像,因而是正向问题,通过线性或非线性的几何变换等即可实现合成。而对于第 2 类问题,由于从侧面图像合成正面图像是一个逆向问题,在此过程中,需要恢复经形变压缩后的人脸纹理以及由于人脸自遮挡而不可见的纹理,因此实现起来相对困难,且其困难程度也会随着侧脸视角的加大而增加。可以看出这两类问题在本质上是截然相反的,因此第 1 类问题的解决方案一般无法用来解决第 2 类合成问题,而第 2 类的合成方法往往可以用于第 1 类的合成问题。

正面人脸合成是一项极具挑战的任务。首先,合成的正脸图像中人脸的局部纹理不能有不自然的形变等;其次,合成的正脸图像要尽可能地与真实人脸相像,也就是说由于人脸的特殊性要求这个逆向问题得到准确求解。对现有的正面人脸合成方法进行了对比研究,简要介绍了各方法的优点以及局限性,进而对每一类方法的合成效果进行了评估,并指出了未来的研究可能值得进行探索的方向。

现有的正面人脸图像合成方法按其基本原理大概可以分为两类:

1) 图形学方法:通过图形学和几何变换的方法来还原由 3 维人脸到 2 维图像投影而产生的纹理形变。但是由于从侧脸到正脸的还原是逆问题,因而简单的 2 维几何线性或非线性仿射变换难以实现。而 3 维图形由于可以直接获取人脸的 3 维数据,因而能够顺利得到正面图像。但是 3 维人脸数据的合成一般需要很大的运算量,实现也相对困难。

2) 统计学习方法:这类方法不具体地描述由转动带来的人脸纹理形变,而希望用某种线性或非线性的模型来描述相应侧脸和正脸之间的映射关系。然后,通过样本学习来寻找最合适的模型。按学习策略的不同,又可分为横向学习和纵向学习方法。

以下将具体地介绍这两类方法。其中图形学的方法又将分为基于 3 维可变模型的方法和基于 2D+3D AAM (active appearance model) 的方法这两种来进行探讨,而统计学习的方法将分为未使用 AAM 的方法和使用 AAM 的方法这两种来进行探讨。具体地分析每一类方法的研究动机,讨论研究

进展及现状。

1 图形学方法

侧脸与正脸图像的不同,本质上是由不同 2 维投影造成的纹理的压缩形变和自遮挡。因而通过图形学方法来对这些形变建模,进而恢复这些形变,是最直接最有效的方法。这些形变模型,如线性仿射变换,在由正脸到侧脸图像的合成中也的确十分成功,如文献[3-4]。然而,由侧脸合成正脸图像是一个逆过程,其中需要恢复由形变和自遮挡导致的在侧脸图像中不存在的纹理,而 2 维图形学的方法难以完成这一任务。如果强行利用线性仿射变换由侧脸合成正脸,那么必然造成合成图像的非自然扭曲,如图 1 所示。因此,用 3 维图形学的方法来解决这一问题必然是选择。因为侧脸图像和正脸图像仅仅是同一个 3 维人脸的不同投影,如果能够由侧脸图像获取到人脸的 3 维数据,那么正脸图像自然就可以获取到。



(a) 侧面人脸图像 (b) 合成正面人脸图像 (c) 真实正面人脸图像

图 1 线性仿射变化合成正面人脸图像的结果

Fig. 1 Frontal face image synthesized by affine transformation

3 维人脸数据的获取一般有两种方式:一是利用已有的 3 维人脸模型,通过优化模型参数来匹配被合成人的侧脸图像,进而获取被合成的 3 维人脸数据;二是由多张侧脸图像和 3 维到 2 维的投影模型,通过求解投影矩阵,从而得到 3 维人脸数据。这两种方法获取的 3 维数据性质是不同的,第 1 种方式获取的 3 维数据包含人脸纹理每一个像素的 3 维坐标以及颜色值,是完备的 3 维数据;而第 2 种方式得到的是合成 3 维数据,它一般只包含形状(特征点)的 3 维坐标而没有纹理信息。下面将分别介绍这两种方式下的正脸图像合成方法。

1.1 基于 3 维可变模型的方法

3 维可变模型(3DMM)^[5-6]最早由 Blanz 等人提出,模型的构建需要使用大量 3 维人脸数据的样本。3DMM 中形状和纹理是分别进行建模的。用形状向量 $\mathbf{S} = (X_1, Y_1, Z_1, X_2, \dots, Y_n, Z_n) \in \mathbf{R}^{3n}$ 来描述人脸

的几何结构,其中 X, Y, Z 是顶点坐标。相应地,用向量 $T = (R_1, G_1, B_1, R_2, \dots, G_n, B_n) \in \mathbf{R}^{3n}$ 表示人脸的纹理,其中 R, G, B 是顶点的颜色值。对所有人脸的形状和纹理分别做主成分分析(PCA)^[7],就得到了3DMM的形状和纹理模型

$$\begin{aligned} S_{\text{model}} &= \bar{S} + \sum_{i=1}^{m-1} \alpha_i s_i \\ T_{\text{model}} &= \bar{T} + \sum_{i=1}^{m-1} \beta_i t_i \end{aligned} \quad (1)$$

式中, \bar{S} 是形状的均值, s_i 是正交的基向量,用于描述由训练样本可产生的形变, α_i 是形状子空间中形状的参数;同样的, \bar{T} 是纹理均值, t_i 是正交的基向量,用于描述由训练样本可产生的纹理的变化, β_i 是纹理子空间中纹理的参数。

构建完3DMM后,合成正面人脸图像就转换成了模型参数的优化问题^[8]。Blaiz等人提出了用3DMM匹配2维人脸图像的方法来获取人脸的3D形状和纹理参数。将3DMM匹配2维图像的过程可以归结为最小化方程

$$E_l = \sum_{x,y} \| I_{\text{input}}(x,y) - I_{\text{model}}(x,y) \|^2 \quad (2)$$

式中, E_l 表示欧几里德距离。从图2可看出,由3DMM得到的正面人脸图像,主观效果十分理想,整

体而言其合成结果非常接近真实情况,且在局部的细节特征上也非常的自然。在文献[6]中,Blanz等人提出了基于3DMM的视角鲁棒的人脸识别,进行了丰富的实验来验证此方法的识别性能,结果在CMU-PIE和FERET数据库上都得到超过95%的识别率。可以说3DMM很好地利用了3维数据天然的优势,实现了由侧面到正面人脸图像的合成。不过,3D数据虽然信息丰富完备,但是其数据量相当大,因而训练和优化过程十分耗时。另一方面,3D数据的获取需要3维扫描仪,相对2D图像而言要复杂得多。

针对上述问题,不少研究者提出了改进3DMM的办法。Romdhani等人^[9]扩展了AAM匹配的改进算法——逆组合的图像配准算法^[10],并将其用于3DMM对2D图像的配准中,从而实现了更优的效率和准确性。在3DMM的基础上,Haar等人^[11]提出了用3DMM匹配到的新实例来自动更新3DMM模型的方法,从而使得3DMM得到更强的描述能力。Zhang等人^[12]结合球谐函数基和3DMM提出了3D球谐函数可变模型,实现了对任何光照下人脸图像的匹配。而Chai等人^[13]简化了3DMM的模型参数,实现了快速匹配。

1.2 基于2D+3D AAM的方法

在描述2D+3D AAM^[14]之前,先回顾一下AAM^[15-16]。AAM的形状是由三角剖分的面膜来描述的,在数学模型中,用面膜的顶点坐标构成的向量来定义形状

$$s = (x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_n) \quad (3)$$

AAM的纹理 g 是用基准面膜内每一个像素点的颜色值来描述的,需要注意的是对于每一张人脸都需要将其面膜内的纹理变形到基准面膜内,来保证每一张人脸都有同样数目的像素点和相同物理意义的纹理。AAM的形状和纹理都允许线性变化

$$\begin{aligned} s &= s_0 + \sum_{i=1}^m p_i s_i \\ g &= g_0 + \sum_{i=1}^m \lambda_i g_i \end{aligned} \quad (4)$$

同3DMM一样,AAM形状和纹理的基也通过PCA而得到。式中, p_i 是形状参数, s_0 是基准形状或称为均值形状, s_i 是特征值最大的 m 个形状基向量; λ_i 是纹理参数, g_0 是纹理均值, g_i 是特征值最大的 m 个纹理基向量。

AAM与3DMM在很多方面都十分相似,都包含



图2 3DMM合成人脸过程及结果^[6]

Fig.2 The process of synthesizing frontal faces by 3DMM

线性的形状模型和线性的纹理模型。其中最主要的区别在于 AAM 的描述是 2D 的,而 3DMM 是 3D 的。由于 3DMM 天然的 3D 优势,3DMM 可以处理视角鲁棒、遮挡等 AAM 无法处理的问题。关于 AAM 的现状和研究进展可以参考 Gao 等人^[17]的综述。

考虑到 3D 模型在 AAM 匹配过程中的优势,受 3DMM 的启发, Xiao 等人^[14]提出了 2D + 3D AAM。他们证明了 AAM 的 2D 形状模型可以描述 3D 形状模型能够描述的任何形状,并且 2D 形状模型还可能会描述出一些不合理的非人脸形状。因此,在原 AAM 的匹配方程基础上加入了 3D 形状的约束项

$$K \left\| \mathbf{N}(s_0 + \sum_{i=1}^m p_i s_i; \mathbf{q}) - \mathbf{P}(\bar{s}_0 + \sum_{i=1}^m \bar{p}_i \bar{s}_i) \right\|^2 \quad (5)$$

式中,前半部分是 AAM 匹配物体 2D 形状的目标方程;后半部分是 3D 形状模型对 2D 形状的约束, K 是约束的权值或者称为惩罚系数。2D + 3D 的 AAM 结合了 2D AAM 匹配更有效率的优点和 3D 模型匹配对视角鲁棒的优点。

由式(5)匹配人脸获取的 3D 人脸参数 \bar{p} , 结合不同视角的投影矩阵 \mathbf{P} , 就可以生成所需视角的 AAM 实例。生成新视角下的 AAM 实例需要分两步进行:1) 对被匹配人脸的 3D 形状模型用投影矩阵 \mathbf{P} 进行投影,得到新视角下人脸的 2D 形状;2) 对式(5)匹配到的人脸的纹理进行仿射变换,将纹理形变到新视角的 2D 形状下。由 2D + 3D AAM 生成的新视角下的人脸实例如图 3 所示。从图 3 中可以看出新视角下的人脸比较自然,纹理也很平滑,整体而言也十分像被合成人,因此合成效果也是比较理想的。

但是,2D + 3D AAM 只能合成 3D 的形状,即三角面片的顶点(特征点)的 3 维坐标。在由侧面合成正面图像的过程中,虽然由 3D 形状可以很好地生成正脸的 2D 形状,但是获取正脸形状的每个三角面片内的纹理依然是逆过程。因此,2D + 3D 的 AAM 模型无法像 3DMM 一样理想地实现正面人脸图像的合成。

在 Xiao 等人^[14]的工作基础上,学者又提出了许多方法来改进 3D 模型对 2D 形状的约束,从而得到对视角更好的匹配鲁棒性。其中最主要的方法是

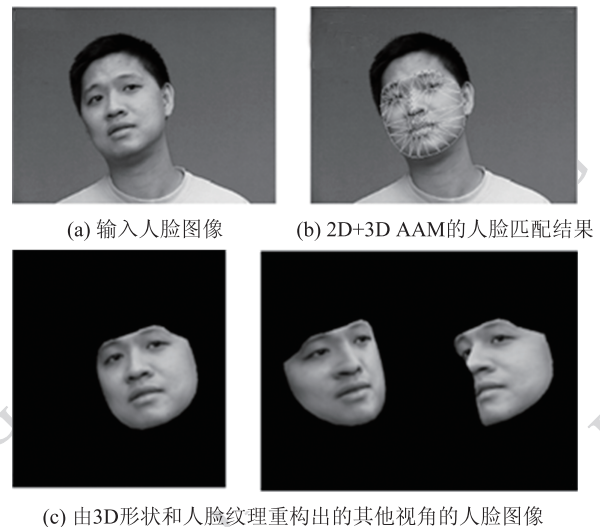


图 3 2D + 3D AAM 合成人脸示例^[14]

Fig. 3 An example of 2D + 3D AAM based face synthesis

通过多个相机同时拍摄的图像来更准确的生成 3D 形状。Ramnath 等人^[18]提出了使用 3 幅同时拍摄的图像来对相机进行标定,从而获取更准确的 3 维形状模型,实现更好的视角匹配鲁棒性。类似地, Sung 等人^[19]使用两个相机的几何关系来进一步限制 2D 形状模型,从而减少模型的参数,优化匹配的效率和。而 Liebelt 等人^[20]提出了另一种限制 2D 形状的方式,他们使用由深度相机获取的深度信息来表示 3D 形状,进而约束 2D 形状模型,同样取得了很好的匹配性能。但是这些方法,同样只能由侧面合成正面图像的 2D 形状,而无法自然地实现正脸纹理的合成。

1.3 讨论

由以上分析可以看出,由于 3D 数据对视角变化天然的鲁棒性,3D 图形学的方法可以十分理想地解决正脸合成问题。基于 3DMM 的方法由于使用的是致密的 3D 数据,因而可以实现正脸形状和纹理的理想合成;而基于 2D + 3D AAM 的方法仅能合成 3D 的形状,因而只能实现正脸形状的理想合成,而无法理想地合成正脸的纹理。因此在目前情况下,基于 3DMM 的方法更适合于进行正面人脸图像的合成。而对于基于 2D + 3D AAM 的方法,需要进一步地研究是否可以通过多幅 2D 图像致密的合成出 3 维的纹理数据,或者是否可以通过其他方法由侧面图像得到正面图像的纹理。关于基于 3DMM 和 2D + 3D AAM 的合成方法的比较见表 1。

表 1 3DMM 与 2D +3D AAM 合成方法的比较
Table 1 Comparison of frontal face synthesis methods based on 3DMM and 2D +3D AAM

模型	训练数据	合成人脸形状	合成人脸纹理	匹配效率	正脸合成效果
3DMM	3D	3D	3D	低	好
2D +3D AAM	2D	3D	2D	高	一般

2 基于统计学习的方法

与图形学方法不同,统计学习方法希望通过对大量样本的学习,来寻找到人脸图像由侧脸到正脸的变化规律。进而在输入某幅未知人脸的侧脸图像时,运用学习到的规律合成出其正脸图像。一般而言,基于统计学习的正脸合成有两种学习策略:一种学习策略希望直接寻出人脸图像从某一侧面视角转到正面的变化规律;另一种学习策略则回避了规律的直接学习,而假设人脸纹理的相似性在不同视角间是类似的,因而其学习目的是寻找相像的人脸并求解相像的程度。将这两种学习策略分别称为纵向学习和横向学习。图 4 给出了基于这两种学习策略的正脸合成示意图。

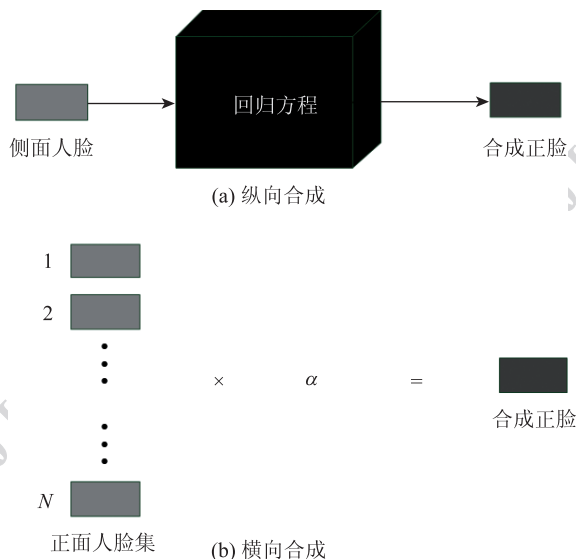


图 4 基于统计学习的合成示意图

Fig. 4 Illustration of statistical learning based frontal face synthesis

基于统计学习的方法还存在预处理方式的不同,简单的可以分为使用 AAM 的和未使用 AAM 的。二者的本质区别在于是否对不同视角间的图

像进行严格对齐。另外,还存在全局合成与局部合成的不同,全局的合成立足于整个正面人脸作为整体一次性的合成,而局部的合成着重于人脸各局部的分别合成,再拼接成完整的正面人脸图像。

下面按预处理方式的不同来分别介绍不同的合成方法,说明学习策略和合成策略,并讨论其优缺点。

2.1 未使用 AAM 的方法

这类方法一般对不同视角的人脸图像间的配准要求不高,只需要人脸大小基本一致和人眼位置大概在同一水平线上。因而,一般通过手动的方式进行粗糙对齐。

Chai 等人^[21]使用了线性回归模型来学习不同视角的人脸图像间存在的规律,根据光照模型和圆柱模型推导出不同视角的人脸图像间的关系可以用线性模型来逼近

$$\min_{A_p} \sum_{i=1}^N \| I_0^i - A_p I_p^i \| \quad (6)$$

式中, I_0^i , ($i = 1, \dots, N$) 为正面人脸图像, I_p^i , ($i = 1, \dots, N$) 为侧面人脸图像, A_p 是要求解的线性变换矩阵。式(6)由最小二乘回归即可求解,得到 A_p 后,输入对应视角的测试人脸 I_p , 则合成的正脸图像为 $I_0 = A_p I_p$ 。

然而这种全局的线性回归方法要求人脸严格对齐,而这对于手动的粗糙对齐是不可能满足的。因此,他们提出了局部线性回归的方法(LLR),将图像划分为许多有重叠的大小相等的矩形块,希望用局部的线性来逼近全局的非线性。由 LLR 合成的正脸图像效果如图 5 所示。图中输入侧脸的视角分别为正脸向左各偏转 22.5°、45°和 20°,所有图像都经过了手动的裁剪,将人脸区域限定在眼睛和嘴巴之间。

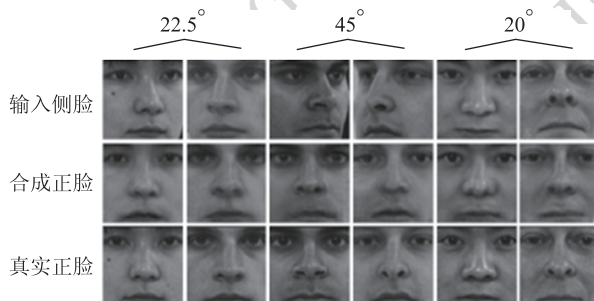


图 5 LLR 合成正脸结果^[21]

Fig. 5 The synthesized frontal faces by LLR

从图中可以看出,这种方法合成的效果还是十分不错的,图像自然、平滑,而且较接近真实值。在合成正脸图像的基础上,Chai 等人^[21]做了基于 PCA + LDA 的多视角人脸识别实验,结果也比较理想。但是,可以看到这种矩形分块的方法需要去除脸颊的纹理来尽量保持视角间对应纹理块的物理意义一致性。因而,其合成的正脸图像是不完整的,而只包含人脸正面中心区域。

类似地, Kim 等人^[22]提出了子空间回归方法 (SR)。他们同样认为不同视角的人脸图像间存在线性关系,与 LLR 不同的是, SR 在图像经 PCA 降维后的数据上做线性的回归。这样提高了回归的效率,但是由于这是一种全局的回归方法,其合成图像的效果不如 LLR 的好。

此外, Prince 等人^[23]提出了一种基于概率框架的学习方法。他们认为同一个人的不同视角的图像间存在共有的隐含的信息,即人的身份。使用 EM 算法^[24]来学习隐含的身份信息,而后由身份信息结合视角信息,从而合成出对应视角的人脸图像。但是这里求解的身份变量只是一种大概的估计,因而其合成的图像效果不是很理想。然而在基于后验概率估计的视角鲁棒的人脸识别实验中,该方法取得了成功。

以上总结了在统计学习框架下未使用 AAM 进行预处理,而实现正面人脸图像合成的方法。总的来说,这些方法合成正面人脸图像的效果一般,但是它们都较好地实现了视角鲁棒的人脸识别。这是因为识别不完全依赖于正面人脸图像合成的结果,而只需所用的识别信息对于不同的人具有足够的区分性。

2.2 使用 AAM 的方法

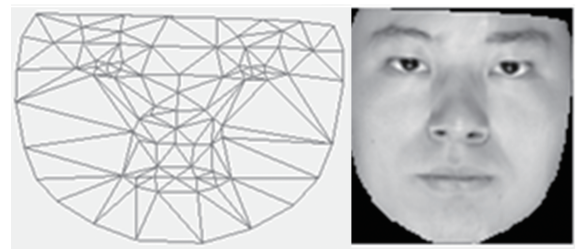
使用 AAM 对人脸图像进行预处理,主要是因为经 AAM 抠取的人脸图像,可以保证严格的配准和对齐。这一特性是保证人脸图像可以准确合成的非常重要的前提。首先描述一下 AAM 是如何实现同一视角内不同人脸图像的严格对齐的。

在 AAM 模型训练过程中,每一张人脸都先要进行特征点的标定,目前已经有方法可以较好地实现特征点的自动标定^[25]和人脸的自动对齐^[26-27],并且每一张人脸的特征点个数必须是完全相等的,每一个特征点的位置也需要基本对应,训练人脸的标点如图 6(a) 所示。这些特征点就构成了每一张人脸的形状,然而 AAM 并不是按每张人脸自有的形

状进行抠图的。AAM 使用普鲁克分析 (Procrustes analysis^[28]) 对所有人脸形状进行标准化,消除所有形状的大小比例、角度和偏移量的不同,进而计算出均值形状,也就是参考形状,最后对参考形状进行德洛内三角剖分 (delaunay triangulation), 并生成参考形状的抠图表 (warptable), 参考形状的三角剖分示例如图 6(b) 所示。这样,所有人脸纹理的抠取就全都按照参考形状的 Warptable 进行,而每一个特征点会在人脸纹理获取时对应到其在图像中的实际位置中去。因而,每一张人脸的像素点个数是完全相同的,每一个像素点的物理意义也是完全一致的。所以,经 AAM 抠取的人脸图像是完全对齐的。当然,这仅限于同一视角的人脸图像,因为只有同一视角的人脸图像才能使用同样的参考形状。



(a) 输入人脸图像以及其人脸形状的标点



(b) 参考形状的三角剖分图 (c) 去除背景后的人脸图像

图 6 AAM 抠取人脸图像过程

Fig. 6 The process of extracting face textures by AAM

在使用 AAM 对训练集中所有人脸图像进行预处理的基础上, Huang 等人^[29]用实验表明,同一人的不同视角的人脸在一个统一的特征空间中可以构成一个平滑的流形,并且不同的流形由差异是相互分开的。进而,他们认为训练集外的人脸的流形可以由训练集内人脸的流形来合成。最后,结合特征空间就可以重构出输入人的其他视角的人脸图像。使用线性模型来求解合成的系数

$$\begin{aligned} \arg \min_{\pi_i} & \left\| \mathbf{u}^p - \sum_{i=1}^k \pi_i \mathbf{x}_i^p \right\|_2 \\ \text{s. t.} & \sum_{i=1}^k \pi_i = 1 \end{aligned} \quad (7)$$

式中, u^p 是输入的第 p 个视角的人脸图像的 AAM 纹理模型参数, $x^p = (x_1^p, \dots, x_k^p)$ 是训练集中第 p 个视角的人脸的 AAM 纹理模型参数, $\pi = (\pi_1, \dots, \pi_k)$ 是要求解的合成系数。在求得第 p 个视角下的合成系数 π 后, Huang 等人认为重构系数 π 在视角间是保持一致的。

由上述流形预测的方法 (ME) 合成的人脸图像如图 7 所示, 白色框内为输入的侧脸图像。可以看出, 合成人脸从图像质量上来说还是不错的, 图像比较平滑、自然。这与使用 AAM 对图像进行

预处理是密不可分的。但是, 可以看到合成人脸无论是在哪个视角都并不是很像被合成的人。其主要原因在于全局合成策略天然的缺陷。基于横向学习的合成在使用 AAM 对人脸图像进行配准后, 其视角间人脸相似一致性的假设, 在一般情况下是合理可行的。但是, 人脸的相似性主要体现于某些局部的相像, 而即使十分像的两张人脸, 也存在不相像的局部。那么, 全局的合成策略在合成人脸时, 必然会带入不相像的局部, 而影响总体合成的效果。



图 7 ME 合成人脸结果^[29]

Fig. 7 The synthesized faces by ME

Lee 等人^[30] 基于张量子空间提出了另一种合成正面人脸图像的策略。他们构建了张量主动表观模型 (TAAM), 实现了对人脸视角、表情、光照鲁棒的匹配。在正确匹配的基础上, Lee 等人认为, 所有人在对应不同视角下的形状差异是类似的, 而另一方面纹理的比率是类似的。因此, 合成正面人脸图像只需修正对应侧面人脸的形状和纹理即可。这种修正方法不需要大量样本来求解回归模型, 实现相对简单。但是, 该方法只适用于较小的视角、表情和光照的变化情况, 因此其应用的局限性较大。

还有其他一些方法也使用了 AAM 来进行人脸图像的合成, 如 Cootes 等人^[31] 使用了多个 AAM 模型来匹配多视角的人脸, 每个 AAM 模型适用于对一定视角范围内的人脸图像进行匹配。他们认为, 在不同视角的 AAM 模型中同一个人的形状和纹理参数是相同的。因此使用某一侧面的 AAM 模型对侧面人脸图像进行匹配, 得到 AAM 参数后, 直接可以由正面 AAM 模型重构出对应的正面人脸图像。实验结果表明了这种合成方法一定的可行性。这种方法本质上与流形预测的方法^[29] 是类似的, 在流形预测的方法中, 不同视角间保持一致的是输入人脸的 AAM 模型参数的重构系数, 而这里直接假设了

模型参数的一致性。这种假设虽然有一定的合理性, 但是并不能证明为普遍的规律, 因而也存在一定的局限性。

2.3 讨论

按预处理方式的不同, 分别介绍了未使用 AAM 的方法和使用 AAM 的方法。比较这两类方法可以发现, 使用 AAM 合成的人脸图像在图像质量上要优于未使用 AAM 的方法。这与经 AAM 预处理的人脸图像可保证严格的对齐有很大的关系。但是使用 AAM 来进行预处理, 需要 AAM 可以进行视角和环境鲁棒的人脸匹配, 而传统 AAM 很难做到这一点。目前已经有一些方法^[14, 18, 31-33] 很好地提高了传统 AAM 匹配鲁棒性的不足, 但是完全视角和环境鲁棒的 AAM 匹配同多视角人脸识别一样, 仍是有待解决的研究课题。

纵向和横向学习在以上方法中均有应用, 纵向方法如文献[21-22], 横向方法如文献[29]。但是无论纵向还是横向学习, 都仅使用了线性学习模型。而线性模型对于纵向和横向是等价的^[21], 都可以看成求解不同训练样本的重构权值。那么是否有合适的非线性模型可以更好地实现合成, 仍有待进一步的研究。

在全局合成^[22]和局部合成^[21]的比较中,由于全局合成天然的劣势,局部合成的人脸图像在纹理的细节上明显与被合成人更加相像。但是,如文献[21]一样的矩形划分策略,难以保证视角间相应的矩形块所对应纹理的物理意义一致性,因此也存在一定的局限性。综合上述合成策略,可将基于统计学习的合成方法比较结果如表 2 所示。

表 2 基于统计学习的合成方法的比较
Table 2 A comparison of different synthesis methods based on statistical learning

方法	使用 AAM	学习策略	合成策略	合成效果
LLR	否	纵向、线性	局部	较好
SR	否	纵向、线性	全局	一般
ME	是	横向、线性	全局	一般
TAAM	是	纵向	全局	一般

3 性能评估

全面而合理的评估正面人脸图像合成方法的性能,需要同时考虑合成的方法和合成图像的效果。人脸识别率在一定程度上反映了合成图像的质量,但是识别率毕竟还受识别方法、数据库等因素影响^[1],因

此仅用识别率来评价合成效果是不完备的。

对于合成方法,首先要评价的是对输入侧脸张数和视角范围的要求,越少的张数和越大的视角范围,则说明合成方法有更好的鲁棒性。另外,合成一张正脸图像所需的时间越少越好。对于合成图像的效果,首先要评价的是合成的人脸是否自然,其次要评价的是合成的人脸与真实人脸的相像程度。自然而相像的合成人脸图像是所希望得到的结果。

综合以上参数,表 3 从需要输入人脸的张数、输入侧脸视角范围、合成速度、合成图像质量、与真实人脸相像程度以及平均识别率这几个方面,对上述方法进行了总结和比较。其中输入人脸张数、视角范围、合成速度以及平均识别率这几项的数据来源于原文中的实验结果和描述。而合成图像质量以及与真实人脸的相像程度这两项,根据原文的实验结果中给出的正脸合成图像,采用主观实验对表中所列的方法进行了比较,并规定从好到差分别为 2, 1, 0, -1, -2 这 5 个分数。将所有方法给出的合成正面人脸图像混合到了一起,并让之前没有见过这些实验结果的 20 个人分别独立地对这些合成正脸图像进行了打分。最后,将这 20 个人的打分进行平均,并将数值归一化到 0 与 1 之间,给出了表中这两项的比较。

表 3 正面人脸合成方法性能评估

Table 3 Performance assessment of frontal face synthesis methods

合成方法	输入人脸张数	视角范围/(°)	合成速度	合成图像质量	相像程度	平均识别率/%
3DMM ^[6]	1	-90 ~ 90	慢	0.78	0.76	95
LLR ^[21]	1	-45 ~ 45	快	0.46	0.61	82.4
ME ^[29]	1	-90 ~ 90	快	0.75	0.46	—
TAAM ^[30]	1	-22.5 ~ 22.5	快	0.50	0.55	81.67

从表 3 中可以看出,基于 3DMM 的正面人脸合成方法除了运算速度外,其他都要优于基于统计学习的方法。而在基于统计学习的方法中,使用 AAM 的方法合成图像的质量要优于没有使用 AAM 的方法,局部合成的方法在相像程度上要好于全局合成的方法。

4 结 语

正面人脸图像的合成,无论从理论上还是从实际的应用角度来看,都是十分具有挑战性的研究课

题。尝试对这个问题目前的研究进展给出一个较完整的综述,介绍了目前解决这个问题的两类方法,图形学的方法和基于统计学习的方法。

从以上介绍的方法中可以看出,正面人脸图像的合成问题,已经得到了一定的研究并取得了较好的研究成果。当然也还存在一些可能的研究方向值得尝试。图形学的方法,相对而言,发展得已经比较成熟,然而仍可以进一步期待图形学的发展,从而可以加以利用来改善人脸合成方法的性能。而对于基于统计学习的方法,仍有一些未被尝试的方向值得加以研究:

1) 基于 AAM 三角剖分的纹理局部合成,可以保证对应局部纹理物理意义的一致性,值得加以研究来进行正脸图像的合成;

2) 非线性的学习模型相对于线性的模型是否可以更好的学习到视角间纹理变化的规律,从而得到更好的合成结果,值得进一步的探索;

3) 基于稀疏表示^[34]的合成方法可以充分利用人脸相似的稀疏性,因此值得进行研究。

还有一些区别于图形学方法和统计学习方法的新思路或许值得进行探索,为正面人脸图像的合成注入新的活力。比如将正面人脸图像合成问题看成相对于侧脸图像的纹理信息填充和补全问题,进而利用图像超分辨或者矩阵填充的思想来解决。相信随着正面人脸图像合成技术的不断发展,它必将服务于未来的生活。

参考文献 (References)

- [1] Zhao W, Chellappa R, Rosenfeld A, et al. Face recognition: a literature survey [J]. *ACM Computing Surveys*, 2003, 35(4): 399-458.
- [2] Phillips P, Grother P, Micheals R, et al. FRVT evaluation report [EB/OL]. [2012-05-02]. <http://www.frv.org/FRVT2002/documents.htm>.
- [3] Li Y, Su G. Face pose estimation and synthesis by 2D morphable model [C]//Proceedings of the International conference on Computational Intelligence and Security. Heidelberg: Springer-Verlag Berlin, 2007: 1001-1008.
- [4] Akshay A, Tom G, Roland G, et al. Learning-based face synthesis for pose-robust recognition from single image [C]//Proceedings of the British Machine Vision Conference. London: British Machine Vision Association and Society for Pattern Recognition, 2009: 1-10.
- [5] Blanz V, Vetter T. A morphable model for the synthesis of 3D faces [C]//Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM Press, 1999: 187-194.
- [6] Blanz V, Vetter T. Face recognition based on fitting a 3D morphable model [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(9): 1063-1074.
- [7] Duda R O, Hart P E, Stork D G. *Pattern Classification* [M]. 2nd ed. New York: John Wiley & Sons, 2001.
- [8] Jones M, Poggio T. Multidimensional morphable models: a framework for representing and matching object classes [J]. *International Journal of Computer vision*, 1998, 29(2):107-131.
- [9] Romdhani S, Vetter T. Efficient robust and accurate fitting of a 3D morphable model [C]//Proceedings of the 9th IEEE International Conference on Computer Vision. New York: IEEE Press, 2003: 59-66.
- [10] Matthews I, Baker S. Active appearance models revisited [J]. *International Journal of Computer Vision*, 2004, 60(2): 135-164.
- [11] Haar F B, Veltkamp R C. Automatic bootstrapping of a morphable face model using multiple components [C]//Proceedings of the 12th IEEE International Conference on Computer Vision Workshops. New York: IEEE Press, 2009: 1497-1504.
- [12] Zhang L, Wang S, Samaras D. Face synthesis and recognition from a single image under arbitrary unknown lighting using a spherical harmonic basis morphable model [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2005: 209-216.
- [13] Chai X J, Shan S G, Qing L Y, et al. Pose and illumination invariant face recognition based on 3D face reconstruction [J]. *Journal of Software*, 2006, 17(3): 525-534. [柴秀娟, 山世光, 卿来云, 等. 基于 3D 人脸重建的光照、姿态不变人脸识别 [J]. *软件学报*, 2006, 17(3): 525-534.]
- [14] Xiao J, Baker S, Matthews I, et al. Real-time combined 2D + 3D active appearance model [C]//Proceedings of 2004 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2004: 535-542.
- [15] Cootes T F, Edwards G J, Taylor C J. Active appearance models [C]//Proceedings of the 5th European Conference on Computer Vision. Heidelberg: Springer-Verlag Berlin, 1998: 484-498.
- [16] Cootes T F, Edwards G J, Taylor C J. Active appearance models [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(6): 681-685.
- [17] Gao X, Su Y, Li X, et al. A review of active appearance models [J]. *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, 2010, 40(2): 145-158.
- [18] Ramnath K, Koterba S, Xiao J, et al. Multi-view AAM fitting and construction [J]. *International Journal of Computer Vision*, 2008, 76(2):183-204.
- [19] Sung J, Kim D. Staam: Fitting a 2D + 3D AAM to stereo images [C]//Proceedings of 2006 IEEE International Conference on Image Processing. New York: IEEE Press, 2006: 2781-2784.
- [20] Liebelt J, Xiao J, Yang J. Robust AAM fitting by fusion of images and disparity data [C]//Proceedings 2006 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2006: 2483-2490.
- [21] Chai X, Shan S, Chen X, et al. Locally linear regression for pose-invariant face recognition [J]. *IEEE Transactions on Image Processing*, 2007, 16(7): 1716-1725.
- [22] Kim M, Zhang Z, Torre F, et al. Subspace regression: Predicting a subspace from one sample [EB/OL]. [2012-05-02]. <http://www.humansensing.cs.cmu.edu/projects/mikim.html>.
- [23] Prince S J D, Elder J H, Warrell J, et al. Tied factor analysis for face recognition across large pose differences [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(6): 970-984.

- [24] Dempster A, Laird N, Rubin D. Maximum likelihood from incomplete data via the EM algorithm [J]. *Journal of the Royal Statistical Society Series B*, 1977, 39(1): 1-38.
- [25] Asthana A, Lucey S, Goecke R. Regression based automatic face annotation for deformable model building [J]. *Pattern Recognition*, 2011, 44(10): 2598-2613.
- [26] Zhu X X, Ramanan D. Face detection, pose estimation, and landmark localization in the wild [C]// *Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE Press, 2012: 2879-2886.
- [27] Cao X D, Wei Y C, Wen F, et al. Face alignment by explicit shape regression [C]// *Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE Press, 2012: 2887-2894.
- [28] Cootes T F, Taylor C J, Cooper D, et al. Active shape models—their training and applications [J]. *Computer Vision and Image Understanding*, 1995, 61(1): 38-59.
- [29] Huang X, Gao J, Cheung S S, et al. Manifold estimation in view-based feature space for face synthesis across poses [C]// *Proceedings of the 9th Asian Conference on Computer Vision*. Heidelberg: Springer-Verlag Berlin, 2010: 37-47.
- [30] Lee H, Kim D. Tensor-based AAM with continuous variation estimation: Application to variation-robust face recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(6): 1102-1116.
- [31] Cootes T F, Walker K, Taylor C J. View-based active appearance models [C]// *Proceedings of the 4th IEEE International Conference on Auto Face and Gesture Recognition*. New York: IEEE Press, 2000: 227-232.
- [32] Zhao C, Cham W K, Wang X G. Joint face alignment with a generic deformable face model [C]// *Proceedings of 2011 IEEE Conference on Computer Vision and Pattern Recognition*. New York: IEEE Press, 2011: 561-568.
- [33] Navarathna R, Sridharan S, Lucey S. Fourier active appearance models [C]// *Proceedings of 2011 IEEE International Conference on Computer Vision*. New York: IEEE Press, 2011: 1919-1926.
- [34] Yang A Y, Ganesh A, Zhou Z H, et al. Fast l_1 -minimization algorithms and application in robust face recognition [EB/OL]. [2012-05-02]. <http://www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-13.pdf>.