

人脸识别研究的新进展及趋势

林景亮¹, 陈岳林², 蔡晓东²

(1. 桂林电子科技大学 机电工程学院, 广西 桂林 541004;

2. 桂林电子科技大学 信息科技学院, 广西 桂林 541004)

摘要: 简要回顾了人脸识别技术的研究背景及发展历程, 总结了近三四年人脸识别方法的研究进展, 根据三维人脸重构方法的数据来源不同, 将其分为基于静态图像和视频序列的三维重构技术。重点对目前多特征和多模态识别技术进行了分类阐述, 分析了一些有代表性的算法并对其识别结果进行了比较。最后, 总结了人脸识别技术现存的研究难点, 并探讨了其未来的发展方向。

关键词: 人脸识别; 三维人脸重构; 多特征; 多模态

中图分类号: TP391. 41

文献标识码: A

New development and trends of face recognition

LIN Jing Liang¹, CHEN Yue Lin², CAI Xiao Dong²

(1. School of Mechanical & Electrical Engineering, Guilin University of Electronic Technology, Guilin 541004, China;

2. School of Information & Technology, Guilin University of Electronic Technology, Guilin 541004, China)

Abstract: In this paper, research background of face recognition and its development are briefly reviewed, and the latest progress of existing face recognition methods are also summarized. Firstly, the current 3D face reconstruction technology is roughly introduced and is divided into picture-based and video-based methods according to the resource. Then, face recognition using hybrid features and multimodal approach are mainly introduced and some typical algorithms are analyzed and listed. Finally, based on the frontal analysis, current challenges of the research are concluded and distilled as suggestion to future research.

Key words: face recognition; 3D face reconstruction; hybrid features; multimodal

在现有的众多生物特征(指纹、虹膜、视网膜、掌形等)识别技术中, 人脸识别技术具有传统的识别技术无法比拟的优点, 如直接、友好、对用户干扰少、更易于被接受等。人脸识别技术是一门融合了多学科(生物学、心理学、认知学等)、多技术(模式识别、图像处理、计算机视觉等)的新的生物识别技术, 它具有广泛的应用和巨大的市场前景, 可用于身份确认(verification or authentication 一对一比较)、身份鉴别(identification or recognition 一对多匹配)、访问控制(门禁系统)、安全监控(银行、海关监控)、人机交互(虚拟现实、游戏)等。在人脸识别研究的早期阶段, 识别技术主要依赖于人的先验知识, 也就是二维人脸识别方法。20世纪80年代, 二维图像处理技术日趋成熟, 在一定约束条件下人脸识别已经能取得较好的效果, 但同时也暴露了二维人脸识别技术对环境人脸姿态变化鲁棒性差的缺点。从20世纪90年代开始, 随着计算机性能的飞跃性发展和成像技术的进一步提高, 三维人脸识别技术成为目前最受关注的新

《微型机与应用》2010年第4期

方法^[1]。在国内, 比较有代表性的成果是北京奥运会使用的自动人脸识别系统。

1 三维人脸重构方法

三维人脸重构是指通过含有人脸的图像或视频中有限的人脸信息来建立人脸三维模型。根据人脸信息来源的不同, 本文将三维人脸重构方法分为从静态图像重构和从视频序列重构两大类。

1.1 从静态图像重构三维人脸

传统的三维重建技术大多是通过特征点的提取和匹配计算特征点的三维坐标来获取脸部的三维结构。特征点定位有手工标定和自动检测两种, 特征点数目比较大时, 手工标定很难严格定义特征点之间的对应关系。ZHANG C等人^[2]采用点对点集的距离来间接描述特征点之间的相似度, 但是, 这种局部相似度不能保证全局最优匹配, 且有可能造成畸变。

为了减少点对应性的困难, 参考文献[3]中提出了以形状匹配为相似性度量的通用头部形变模型, 在不需要

欢迎网上投稿 www.pcachina.com 1

综述与评论

Review and Comment

严格的特征点对应甚至某些特征点缺失的情况下,完成姿态估计和三维重建。为了解决搜索点之间的最佳相似性,文中还提出了一种多级搜索的方法,大大减少了搜索时间,但这只是基于多幅图像的情况。虽然多幅图像可以消除人脸特征部件检测的不确定性,但特定脸的多幅图像一般难以获得,所以当前的很多研究都是基于单一图像的。为了能获得人脸的完整的脸部特征信息,单一图像一般要求是正面人脸图像且是中性表情。

BREUER P^[4]和胡元奎^[5]等人都是基于单一图像进行建模。参考文献[4]中提出了一种融合支持向量机(Support Vector Machine)和三维形变模型(3D Morphable Model)的方法,分别用不同方法检测人脸和人脸的局部特征(鼻尖、眼角、嘴角),然后确定人脸特征可能的位置并评估特征点的轮廓,通过迭代处理提高算法对头部方向的鲁棒性,最后初始化形变模型的模型试配流程来产生高分辨率的三维人脸模型。但是,形变模型算法需要花费很长的时间对大量的三维人脸数据进行训练。参考文献[5]基于通用三维人脸模型的三维人脸合成方法能很好地减少算法的复杂性和训练时间。他们利用了基于知识的特征点定位算法和ASM(Active Shape Model)方法进行人脸特征点的定位,用SFS(Shape From Shading)算法恢复人脸表面深度,并利用内插算法对通用三维人脸模型进行变形处理以生成适用于特定人脸的三维模型。此算法的优点是只需要一个通用的三维人脸模型即可,不需要进行额外的训练,而且对训练数据以及存储空间的实际需求很容易满足,具有明显的优势。

无论是单一图像还是多幅图像,静态图像提供的信息都是相对有限的,例如无法提供连续多帧图像和时间相干性。于是,在研究从静态图像重构人脸模型的同时,少数研究尝试了从视频图像序列重构三维人脸模型的方法。

1.2 从视频序列重构三维人脸

从视频重构人脸的过程和从单一图像重构人脸的过程基本上一样(如图1所示),只是源图像不同。视频序列虽然也可以使用适合于单一图像的方法,如参考文献[4]从图像序列中选出最合适的一帧图像(例如正面图像)来重构三维人脸模型,但这显然不是应用视频序列图像的目的。



图1 三维人脸重构算法框图

PARK U等^[6]用通用三维人脸模型和两个视频帧来重构特定用户三维人脸模型。他们从立体视频中重建脸部标记点的稀疏集合,将其用于薄板样条TPS(Thin Plate Spline)的试配过程,在TPS试配的基础上对一般人脸模型进行非线性变换,得到合适的三维人脸模型,将视频中人脸纹理信息对三维人脸模型进行映射,从而获得真实的三维人脸模型。该方法应用比较广泛,但在

重建的初始化阶段,初始值与通用模型非常相似,导致重建的模型与视频中的人脸相比更近似于通用人脸模型。此缺点可以利用SFM(Structure From Motion)算法^[7]解决。首先,SFM算法能够保留通用人脸模型的特定特征;其次,通过与通用人脸模型比较,两脸间的误差在能量函数最小化过程中都被修正。但是,不使用通用模型,单纯用SFM算法对视频图像进行三维估计会使深度估计变得困难,可能会带来其他信息不足或估计误差等问题。文中的算法流程如图2所示^[7]。



图2 三维建模算法流程图

2 多特征融合人脸识别方法

重构三维人脸模型只是人脸识别的重要手段之一,但其算法相对复杂。目前,使用多方法(二维)融合来提高识别性能还是人脸识别领域研究的热点^[8-15]。多方法融合主要分为融合多种脸部特征(肤色、轮廓和纹理等信息)和融合多种模态(二维和三维信息)两种。由于图像与形状信息相对独立,多特征融合的人脸识别方法曾经很少使用。

2005年前后,SU Hong Tao^[8]和周晓彦等人^[9]分别提出了融合主分量分析PCA(Principal Component Analysis)与线性判别式分析LDA(Linear Discriminant Analysis)混合特征和融合核主元分析KPCA(Kernel Principal Component Analysis)与核判别式分析KDA(Kernel Discriminant Analysis)的人脸识别算法。在参考文献[8]中,利用库中图像和被检测图像的交互信息进行粗分类,在图形数据的傅里叶频率区域进行PCA和LDA特征的抽取。由于PCA和LDA能分别反映图像的不同特性,故融合两子特征将取得比单一特征更好的分类性能。参考文献[9]首先求解KDA的最佳判别矢量,然后基于KPCA准则函数求得另一组投影矢量,最后将两组投影矢量融合成一组新的特征矢量用于特征的提取。

与参考文献[8]和参考文献[9]不同,LIU Zhi Ming等^[10]应用了颜色和频率特征。离散傅里叶变换将肤色RIQ空间转换到频域并分别求出各个颜色分量的掩面(mask),用增强Fisher模型EFM(Enhanced Fisher Model)抽取互补频率特征(包括检测脸、库中人脸和R分量大小),在特征水平上用级联的方法将其融合在一起,将得到的相似性结果用于分类。分别对各个分量进行互补频率特征抽取和分类,最后将分类结果通过加权因子再次融合在一起,用于人脸的识别。该方法比单颜色分量时的识别率有很大的提高,这也说明,单一特征所包含的信息都是有限的。充分利用人脸的肤色和纹理结构等各种特征,将会取得较好的识别效果。

《微型机与应用》2010年第4期

综述与评论

Review and Comment

3 多模态融合人脸识别方法

多模态融合的人脸识别方法与多特征融合方法一样,目的是融合二维和三维甚至四维的脸部信息,以提高识别的精度和算法对环境的鲁棒性^[11,14-15]。

在多模态融合的研究上,比较成功的是 MIAN A 等人^[12]提出的全自动三维人脸识别算法。该算法能全自动检测人脸鼻子区域,自动修正三维人脸姿态和进行标准化处理,可以在规模比较大的人脸识别中通过粗匹配快速拒绝大多数不适合的人脸并能自动分割易受表情影响和不易受表情影响区域。但是,算法在最后识别阶段容易受头发影响,且无法自动检测侧面图像。为了解决此问题,他们在前期研究的基础上提出了一种有效的多模态(二维/三维)融合和混合(局部/整体特征)匹配的方法^[13]。该方法在参考文献[12]的基础上,用三维球面人脸描述 SFR(Spherical Face Representation)和可变尺度特征变换 SIFT (Scale-Invariant Feature Transform)描述子来构建拒绝分类器,通过粗(整体)匹配快速拒绝大部分候选人脸并对剩下的人脸进行区域分割,得到对表情不敏感的局部特征区域(眼睛-前额/鼻子),再用修正的迭代最近点算法(Modified ICP)对这些局部特征进行单独匹配。

相对三维融合二维信息的方法,三维与四维信息的融合技术应用比较少。参考文献[14]中 PAPTAEODOROR T 等人利用人脸纹理结构和表面信息注册,提出一种自动的四维人脸识别方法。他们利用立体摄像系统,结合面部外观的二维纹理映射描述符和三维面部几何的致密三维网格顶点描述符,重构四维人脸数据。在识别阶段先进行三维或者四维刚性注册,然后通过 ICP 算法和欧氏距离计算两脸部图形对应点的距离,根据相似性判断来进行识别。表 1 所示是多特征融合与多模态融合的一些代表性算法在各自实验中的结果。

由于现在人脸数据库并不统一,不同文献的实验数据库一般不一样,且所使用的图像类型(颜色图像和灰

度图像)与大小不同,因此很难根据它的识别率来判断某个算法的优劣性。此外,即使使用同一个数据库,数据库中也往往含有各种变化剧烈程度不一致的表情、姿态和光照情况,在不同的环境和姿态表情下,算法的结果会有很大的差别。

目前,三维人脸识别方法已经代替二维人脸识别方法成为研究热点。克服表情姿态和环境影响是目前三维识别研究的主要目的,其手段就是从图像中重构人脸三维模型。图像采集设备的差异和成像原理的不同,造成了采集数据的差异,如何更好地获取更多的有效信息与对数据的正规化一样成为难题。同时,特征点定位和人脸特征的提取对三维人脸重构非常重要,有效地监测定位和特征提取算法有待进一步地完善。

虽然三维数据获取技术有了飞速发展,但远没有达到像获取二维图像那么方便和普及,且基于三维信息的识别技术同样会受到人脸姿态和环境的影响。此外,由于三维识别算法在某些特定环境下(视频监控、受限制区域)无法像二维识别技术一样取得令人满意的结果,它需要二维方法甚至更高维方法的辅助来提高它的识别精度和鲁棒性,故融合多特征的识别技术和融合多模态的人脸识别技术在很长的一段时间内将是最有效的人脸识别方法之一。

参考文献

- [1] 段锦,周春光,刘小华. 三维人脸识别研究进展 [J]. 小型微型计算机系统, 2004, 25 (5): 886-890.
- [2] ZHANG C, COHEN S. 3-D face structure extraction and recognition from images using 3-D morphing and distance mapping [J]. IEEE Trans on PAMI, 2002, 11 (11): 1249-1258.
- [3] XUE Feng, DING Xiao Qing. 3D reconstruction of human face based on shape match morphing model [J]. ACTA ELECTRONICA SINICA, 2006, 34(10): 1896-1899.

表 1 多特征与多模态识别算法一览表

类别	关键技术	数据库	人数(图像)	识别率	表情	分辨率	文献
多特征	PCA+LDA+RBFN	ORL	40(400)	90%以上	有	92×112	[8]
	DCT+EFM	FRGC v2.0	12 776	80.3%	有	64×64	[10]
	KPCA+KDA+GNFL	ORL	40(400)	96.7%	有	112×92	[9]
		Yale	11(165)	87.33%	有	243×32	
	Gabor+PCA+HTPSO	自建库	10(200)	96%以上	有	100×100	[11]
多模态	P ² CA(2D+3D)	UPC	23(621)	91.91%	有	122×100	[15]
	SFR+SIFT+MICP (2D+3D)	FRGC v2.0	466(4 007)	99.74%	无	480×640	[13]
				98.31%	有		
ICP+ED (3D+4D)	自建库	62	100%	无	未给出	[14]	

- [4] BREUR P, KIM Kwang-In, KIENZLE W, et al. Automatic 3D face reconstruction from single images or video[C]. Automatic Face & Gesture Recognition, 2008. FG '08. 8th IEEE International Conference on 17-19 Sept. 2008; 1-8.
- [5] 胡元奎. 可变光照和可变姿态下的人脸图像识别研究[D]. 合肥: 中国科学技术大学, 2006.
- [6] PARK U, JAIN A K. 3D face reconstruction from stereo video [C]. Proceedings of the 3rd Canadian Conference on Computer and Robot Vision (CRV '06), 2006; 41 - 41.
- [7] CHOWDHURY A R, CHELLAPPA R, VO T, et al. 3D face reconstruction from video using a generic model[C]. Multimedia and Expo, 2002. ICME '02. Proceedings. 2002 IEEE International Conference on Volume 1, 26-29 Aug, 2002, 1; 449 - 452.
- [8] SU Hong Tao, FENG D D, WANG Xiu Ying, et al. Face recognition using hybrid feature [C]. Machine Learning and Cybernetics, 2003 International Conference, 2003 (5); 3045 - 3049.
- [9] 周晓彦, 郑文明. 一种融合 KPCA 和 KDA 的人脸识别新方法[J]. 计算机应用, 2008, 28(5): 1263-1266.
- [10] LIU Zhi Ming, LIU Cheng Jun. A hybrid color and frequency features method for face recognition [J]. Image Processing, IEEE Transactions, 2008, 17(10): 1975 - 1980.
- [11] LIN C J, CHU C H, LEE C Y, et al. 2D/3D Face Recognition Using Neural Networks Based on Hybrid Taguchi-Particle Swarm Optimization [C]. Intelligent Systems Design and Applications, 2008. ISDA '08. Eighth International Conference on Volume 2, 26-28 Nov, 2008; 307 - 312.
- [12] MIAN A, BENNAMOUN M, OWENS R. Automatic 3D face detection, normalization and recognition [J]. 3D Data Processing, Visualization, and Transmission, Third International Symposium on 14-16 June, 2006; 735 - 742.
- [13] MIAN A S, BENNAMOUN M, OWENS R. An efficient multimodal 2D-3D hybrid approach to automatic face recognition [J]. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 2007, 29(11): 1927-1943.
- [14] T PAPTHERODOROU, RUECKERT D. Evaluation of automatic 4D face recognition using surface and texture registration [C]. Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on 17-19 May, 2004; 321 - 326.
- [15] RAMA A, TARRES F, ONOFRIO D, et al. Mixed 2D-3D information for pose estimation and face recognition[C]. Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on Volume 2, 2006; 14-19.

(收稿日期: 2009-11-09)

作者简介

林景亮, 男, 1985年生, 在读硕士, 主要研究方向: 图像处理与模式识别;

陈岳林, 男, 1963年生, 教授, 硕士生导师, 主要研究方向: 图像处理与模式识别;

蔡晓东, 男, 1972年生, 博士, 主要研究方向: 多媒体信号处理、视频监控等。

当你脑海中闪过某个画面的时候, 就是画面开始发声的时候。
画面的声音真实吗? 你的耳朵感受到了什么? 你有没有听到, 那分明是一种精致和格调!
耳边的精致生活, 是要重现声音自然的本质态度。说到底, 给我震撼真实的声音, 其余免谈。

家庭影院总动员: 从耳边开始精致你的生活

由《家庭影院技术》杂志、影音中国(www.hdavchina.com)网站、杜比实验室联合多家合作机构和媒体共同发起的高清家庭影院大赛于2010年1月1日正式拉开帷幕。本次大赛以“我爱家庭影院”为主题, 全面征集家庭影院构建中的各种想法、理念、心得和解决方案, 为广大喜爱家庭影院的朋友提供一个展现自我创意、提升专业视野的自由舞台, 释放精彩的影像与真实的声音, 摩擦出最炫最酷的真实感; 分享你和家庭影院的故事, 沉淀出精致格调的影音娱乐生活!

本次大赛将从2010年1月1日持续到2010年4月18日, 在全国范围内同步展开。杜比实验室将邀请相关领域的资深人士和著名的影像、声音艺术家, 讲述如何打造最酷的家庭影院, 享受最炫最真实的环境声。此次大赛力求搭建酷爱娱乐、懂得享受好声音的爱好者之间、爱好者与专家之间相互交流的平台, 进而为广大爱好者提供更好的享受和体验。

比赛期间, 《家庭影院技术》杂志将对赛事进行全程报道, www.hdavchina.com 影音中国网站也将同步传递最新赛事消息。家庭影院征总动员现已全面启动, 赶快来展示你参与构建的真实的家庭影院和大家分享你的影音生活, 你的作品可以是独立家庭影院视听室、客厅影院, 也可以是书房、卧室影院或其他类影音系统。即刻登录 www.hdavchina.com, 上传作品照片和说明或者邮寄材料至《家庭影院技术》杂志社, 即可加入。

我们拭目以待, 期待共同分享优质影音体验, 从耳边开始, 精致生活!

(杜比实验室供稿)