

2000, 5A(11)

第5卷(A版) 第11期  
2000年11月中国图象图形学报  
Journal of Image and GraphicsVol. 5(A), No. 11  
Nov. 2000

## 人脸识别技术综述

张翠平 苏光大

(清华大学电子工程系“智能技术与系统”国家重点实验室图形图象分室, 北京 100084)

TP391.41

885-894

**摘要** 首先对计算机人脸自动识别技术的研究背景及发展历程做了简单回顾, 然后对人脸正面像的识别方法, 按照识别特征的不同进行了分类综述, 主要介绍了特征脸(Eigenface)方法、基于小波特征的弹性匹配(Elastic Matching)的方法、形状和灰度模型分离的可变形模型(Flexible Model)以及传统的部件建模等分析方法。通过对各种识别方法的分析与比较, 总结了影响人脸识别技术实用化的几个因素, 并提出了研究和开发成功的人脸识别技术所需要考虑的几个重要方面, 进而展望了人脸识别技术今后的发展方向。

**关键词** 人脸识别, 特征脸, 小波特征, 形状无关模型

**中图分类号**: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2000)11-0885-10

计算机自动识别

## Human Face Recognition: A Survey

ZHANG Cui-ping, SU Guang-da

(Electronic Engineering Department, The State Key Laboratory of Intelligent Technology and System,  
Tsinghua University, Beijing 100084)

**Abstract** In this paper, Research background of automatic face recognition and its relation to human vision system are briefly reviewed. Then current face recognition technologies are roughly introduced and classified according to different recognition features. Four main algorithms are analyzed and compared. The first is eigenface, which is extraction of global features using the PCA. In this approach, a set of faces is represented using a small number of global eigen vectors, which encode the major variations in the input set. The second is flexible model, which separate shape and gray parameter. The third is wavelet-based elastic graph matching, in which memorized faces are represented by regular graphs, whose vertices are labeled by a multi resolution description in terms of localized spatial frequencies. Spatial relationships within the object are labeled by geometrical distance vectors. The last method is traditional analytical techniques. Based on the analysis and comparison, key factors in face recognition technologies are concluded and distilled as suggestion to future research.

**Keywords** Face recognition, Eigenface, Wavelet-based feature, Shapeless model

## 0 引言

计算机人脸识别技术也就是利用计算机分析人脸图象, 进而从中提取出有效的识别信息, 用来“辨认”身份的一门技术。人脸识别技术应用背景广泛, 可用于公安系统的罪犯身份识别、驾驶执照及护照等与实际持证人的核对、银行及海关的监控系统及自动门卫系统等。虽然人类的人脸识别能力很强, 能够记住并辨别上千个不同人脸, 可是计算机则困难

多了。其表现在: 人脸表情丰富; 人脸随年龄增长而变化; 人脸所成图象受光照、成像角度及成像距离等影响; 而且从二维图象重建三维人脸是病态(ill-posed)过程, 目前尚没有很好的描述人脸的三维模型。另外, 人脸识别还涉及到图象处理、计算机视觉、模式识别以及神经网络等学科, 也和人脑的认识程度紧密相关, 这诸多因素使得人脸识别成为一项极富挑战性的课题。

计算机人脸识别技术是近 20a 才逐渐发展起来的, 90 年代更成为科研热点。仅 1990 年到 1998 年

基金项目: 清华大学科技发展基金, 公安部资助项目

收稿日期: 1999-11-12; 改回日期: 2000-04-18

之间, EI 可检索到的相关文献就多达数千篇. 由于人脸识别实验所采用的人脸库通常不大, 最常见的人脸库仅包括 100 幅左右的人脸图象, 如 MIT 库、Yale 库、CMU 库等人脸库均为小型库, 且由于不同人脸库之间的输入条件各异, 因此不同的识别程序之间很难进行比较. 为促进人脸识别算法的深入研究和实用化, 美国国防部发起了人脸识别技术 (Face Recognition Technology 简称 FERET) 工程<sup>[1]</sup>, 它包括一个通用人脸库和一套通用测试标准. 该 FERET 库可用于各种人脸识别算法的测试比较. 1997 年, FERET 人脸库存储了取自 1 199 个人的 14 126 幅图象, 其中同一人的图象差异, 包括不同表情、不同光照、不同头部姿态以及不同时期 (相隔 18 个月以上) 拍摄差异等. 如今 FERET 人脸库仍在扩充, 并定期对各种人脸识别程序进行性能测试, 其分析测试结果对未来的工作起到了一定的指导作用. 由于 FERET 库中包括军人的图片, 不能在美国以外获得, 因此其他国家的研究只能采用本地的人脸库, 如英国的 Manchester 人脸库<sup>[2]</sup>.

通常, 人类进行人脸识别依靠的感觉器官包括视觉、听觉、嗅觉、触觉等, 一般人脸的识别可以用单个感官完成, 也可以是多感官相配合来存储和检索人脸, 而计算机的人脸识别所利用的则主要是视觉数据. 另外, 计算机人脸识别的进展还受限于对人类本身识别系统的认识程度. 研究表明<sup>[3]</sup>, 人类视觉数据的处理是一个分等级的过程, 其中最底层的视觉过程 (视网膜功能) 起信息转储的作用, 即将人眼接收的大量图象数据变换为一个比较规则的紧凑表达形式. 生理学的研究表明, 人眼视网膜上存在着低层次和高层次的细胞. 其中, 低层次的细胞对空间的响应和小波变换的结果相似<sup>[4]</sup>; 而高层次的细胞则依据一群低层次细胞的响应, 而作出具体的线、面乃至物体模式的响应. 以此为依据, 在计算机人脸识别中, 可以将那些通过大量图象数据简单处理后获得的特征定义为低层次特征, 而将线、面、模式等描述特征定义为高层次特征. 由此, 图象 KL 变换后的系数特征、小波变换特征及一些统计特征均属低层次特征的范畴, 而人脸部件形状分析的结果则为高层次特征. 由于视觉数据经传输后的重建, 需依赖于人脑中早期形成的先验知识, 因此在人的识别系统中, 人脸的检测是一个整体识别和特征识别共同作用的结果<sup>[5]</sup>; 具体说来, 远处辨认人, 主要是整体识别, 而在近距离的人脸识别中, 特征部件的识别则更重要.

另外, 人脸的各部件对识别的贡献也不相同, 如眼睛和嘴巴的重要程度大于人的鼻子, 人脸上半部分重要性大于人脸下半部分, 其中特别的人脸更容易被识别记住<sup>[6]</sup>, 比如说歪嘴, 或是独眼龙等人脸就更容易为人记起, 没有个性的人脸相对就需要更长的时间来辨认. 根据对人脑的研究表明<sup>[7]</sup>, 人脸的表情识别和人脸识别虽然存在联系, 但总体说是分开的、并行的处理过程. 这些结论对于设计有效的识别方法起到了一定启发作用. 在现有的识别算法中, 特征脸方法<sup>[4]</sup>和神经网络方法<sup>[5]</sup>是基于整体人脸的识别, 而基于提取眼睛等部件特征而形成特征向量<sup>[6]</sup>的方法就是基于人脸特征的识别.

人脸识别的研究始于 60 年代末, 最早的研究见于文献<sup>[7]</sup>, Bledsoe 以人脸特征点的间距、比率等参数为特征, 建成了一个半自动的人脸识别系统. 而且早期人脸识别研究主要有两大方向: 一是提取人脸几何特征的方法<sup>[7]</sup>, 包括人脸部件规一化的点间距离和比率以及人脸的一些特征点, 如眼角、嘴角、鼻尖等部位所构成的二维拓扑结构; 二是模板匹配的方法, 主要是利用计算模板和图象灰度的自相关性来实现识别功能. Berto 在 1993 年对这两类方法作了较全面的介绍和比较后认为, 模板匹配的方法优于几何特征的方法<sup>[8]</sup>. 目前的研究也主要有两个方向: 其一是基于整体的研究方法, 它考虑了模式的整体属性, 包括特征脸 (Eigenface) 方法、SVD 分解的方法<sup>[9]</sup>、人脸等密度线分析匹配方法<sup>[10]</sup>、弹性图匹配 (elastic graph matching) 方法<sup>[11]</sup>、隐马尔可夫模型 (Hidden Markov Model) 方法<sup>[12]</sup>以及神经网络的方法等; 其二是基于特征分析的方法, 也就是将人脸基准点的相对比率和其它描述人脸脸部特征的形状参数或类别参数等一起构成识别特征向量. 这种基于整体脸的识别不仅保留了人脸部件之间的拓扑关系, 而且也保留了各部件本身的信息, 而基于部件的识别则是通过提取出局部轮廓信息及灰度信息来设计具体识别算法. 文献<sup>[8]</sup>认为基于整个人脸的分析要优于基于部件的分析, 理由是前者保留了更多信息, 但是这种说法值得商榷, 因为基于人脸部件的识别要比基于整体的方法来得直观, 它提取并利用了最有用的特征, 如关键点的位置以及部件的形状分析等, 而对基于整个人脸的识别而言, 由于把整个人脸图象作为模式, 那么光照、视角以及人脸尺寸会对人脸识别有很大的影响, 因此如何能够有效地去掉这些干扰很关键. 虽然如此, 但对基于部件分析的人

脸识别方法而言也有困难,其难点在于如何建立好的模型来表达识别部件。近年来的一个趋势是将人脸的整体识别和特征分析的方法结合起来,如 Kin-Man Lam 提出的基于分析和整体的方法<sup>[13]</sup>, Andreas Lanitis 提出的利用可变形模型(Flexible Models)来对人脸进行解释和编码的方法<sup>[14]</sup>。

在介绍重要的人脸识别方法之前,先扼要说明一下应用于人脸识别的其它方法。其中 SVD 方法和特征脸识别方法同属统计分析的范畴,都是将表达人脸的大量图象数据降维后进行模式分类,其区别仅是变换基的给出不同;而等密度线的分析方法则试图通过从二维的人脸图上抽取等密度线(即等灰度线)来反映人脸的三维信息,其根据是地图上的等高线能反映地形特征,那么通过不同人脸的等密度线也可比较人脸的相似度;HMM 是语音处理中成功的一种统计方法;而神经网络方法通常需要将人脸作为一个一维向量输入,因此输入节点庞大,其识别重要的一个目标就是降维处理。根据文献[15]对于自组织神经网络方法的分析,该文认为可采用自组织神经网络的  $P$  个节点来表达原始的  $N$  个输入( $P < N$ ),但由于将  $P$  个输出进行分类,其识别的效果仅相当于提取人脸空间特征向量后进行的识别分类,因此采用此类神经网络进行识别的效果只能是特征脸的水平,所以本文将不对神经网络作专门介绍。需要说明的是,由于人脸处于高维空间,如  $100 \times 100$  的图象为 10 000 维,这样神经网络的输入节点将很庞大,因此实际训练网络的时候参数繁多,实现起来很困难,但神经网络方法的优点是可以针对特定的问题进行子空间设计,如神经网络的方法可以用作性别识别等问题<sup>[15]</sup>。

## 1 常用的人脸识别方法简介

### 1.1 基于 KL 变换的特征脸识别方法

#### 1.1.1 基本原理

KL 变换是图象压缩中的一种最优正交变换,人们将它用于统计特征提取,从而形成了子空间法模式识别的基础。若将 KL 变换用于人脸识别,则需假设人脸处于低维线性空间,且不同人脸具有可分性。由于高维图象空间 KL 变换后可得到一组新的正交基,因此可通过保留部分正交基,以生成低维人脸空间。而低维空间的基则是通过分析人脸训练样本集的统计特性来获得。KL 变换的生成矩阵可以

是训练样本集的总体散布矩阵,也可以是训练样本集类间散布矩阵,即可采用同一人的数张图象的平均来进行训练,这样可在一定程度上消除光线等的干扰,且计算量也得到减少,而识别率不会下降。

也就是说,根据总体散布矩阵或类间散布矩阵可求出一组正交的特征向量  $u_1, u_2, \dots, u_n$ , 其对应的全部特征值分别为  $\lambda_1, \lambda_2, \dots, \lambda_n$ , 这样,在新的正交空间中,人脸样本  $X$  就可以表示为

$$X = \sum_{i=1}^n \hat{X}_i u_i \quad (1)$$

若通过选用  $m$  ( $m < n$ ) 个特征向量作为正交基,则在该正交空间的子空间中,就可得到以下近似表达式

$$X = \sum_{i=1}^m \hat{X}_i u_i \quad (2)$$

如将子空间的正交基按照图象阵列排列,则可以看出这些正交基呈现人脸的形状,因此这些正交基也被称作特征脸,这种人脸识别方法也叫特征脸方法。关于正交基的选择有不同的考虑,即与较大特征值对应的正交基(也称主分量)可用来表达人脸的大体形状,而具体细节还需要用与小特征值对应的特征向量(也称次分量)来加以描述,因此也可理解为低频成分用主分量表示,而高频成分用次分量表示。其中,采用主分量作正交基的方法称为主分量方法(PCA)。同时,也有人采用  $m$  个次分量作为正交基,原因是所有人脸的大体形状和结构相似,真正用来区别不同人脸的信息是那些用次分量表达的高频成分。由训练得到特征脸后,将待识别人脸投影到新的  $m$  维人脸空间,即用一系列特征脸的线性加权和来表示它,这样即得到一投影系数向量来代表待识别人脸,这时候,人脸识别问题已转化为  $m$  低维空间的坐标系数矢量分类问题,而分类最简单的做法是最小距离分类。

KL 变换在 90 年代初受到了很大的重视,实际用于人脸识别也取得了很好的效果,其识别率从 70~100% 不等,这取决于人脸库图象的质量。从压缩能量的角度来看,KL 变换是最优的,它不仅使得从  $n$  维空间降到  $m$  维空间前后的均方误差最小,而且变换后的低维空间有很好的脸表达能力,然而这不是说已经具有很好的人脸辨别能力。选择训练样本的散布矩阵作为 KL 变换的生成矩阵,是由于其最大特征向量抓住了该样本集合的主要分布,但这是图象统计,而不是人脸统计方法。它虽然考虑了图象之间所有的差异,但由于它不管这样的差异是由照明、发型

变更或背景导致,还是属于人脸的内在差异,因此特征脸识别的方法用于人脸识别存在理论的缺陷.研究表明,特征脸的方法随着光线、角度及人脸的尺寸等因素的引入,识别率急剧下降.虽然可通过采用同一人的训练样本的平均来计算类间散布矩阵,但也只能在一定程度上纠正这个缺点.研究结果表明,主分量的方法使得变换后表达能力最佳,次分量的方法则考虑了高频的人脸区分能力.由于对KL变换而言,外在因素带来的图象差异和人脸本身带来的差异是不加任何区分的,因此,不管如何选择正交基,也不能根本解决问题.其改善的一个思路是针对干扰所在,对输入图象作规范化处理,其中包括将输入图的均值方差归一化、人脸尺寸归一化等;另一种改进是考虑到局部人脸图象受外在干扰相对较小,在进行人脸识别时,除计算特征脸之外,还可利用KL变换计算出特征眼睛、特征嘴巴等,然后将局部特征向量加权进行匹配,就能够得到一些好的效果.

#### 1.1.2 对特征脸方法的改进

一种较好的特征脸改进方法是fisher脸方法(fisherface)<sup>[17]</sup>,众所周知,fisher线性判别准则是模式识别里的经典方法,一般应用fisher准则是假设不同类别在模式空间是线性可分的,而引起它们可分的主要原因是不同人脸之间的差异,fisher的判别准则是:不同类样本尽可能远,同类样本尽可能近.文献[17]对用KL变换和fisher准则分别求出来的一些特征脸进行比较后得出如下结论,即认为特征脸很大程度上反映了光照等的差异,而fisher脸则能压制图象之间的与识别信息无关的差异,Belhumeur的试验<sup>[17]</sup>,是通过对160幅人脸图象(一共16个人,每个人10幅不同条件下的图象)进行识别,若采用KL变换进行识别,其识别率为81%;若采用fisher方法则识别率为99.4%,显然fisher方法有了很大的改进,Chengjun Liu在KL变换基础上提出了PRM(Probabilistic Reasoning Models)模型<sup>[18]</sup>,并在PRM中采用了贝叶斯分类器,它是利用最大后验概率进行分类,其类条件概率密度的方差参数用类内散布矩阵来估计,而且,PRM是采用马氏距离,而不是采用最小欧氏距离的判别准则,并且特征脸和fisher脸均可以看成是PRM的特殊情况.

文献[19]的改进方法是将人脸图象进行差异分类,即分为脸间差异和脸内差异,其中脸内差异属于同一个人脸的各种可能变形,而脸间差异则表示不同人的本质差异,而实际人脸图的差异为两者之和.

通过分析人脸差异图,如果脸内差异比脸间差异大,则认为两人脸属于同一人的可能性大,反之属不同人的可能性大.假设该两类差异都是高斯分布,则先估计出所需的条件概率密度<sup>[19]</sup>,最后也归为求差图在脸内差异特征空间和脸间差异特征空间的投影问题.如果说fisher脸的方法是试图减少光照等的外在干扰,那么文献[19]则是解决表情干扰的一点有效尝试,虽然这样的尝试还很初步.文献[19]中提到,ARPA在1996年进行的FERET人脸识别测试中,该算法取得了最好的识别效果,其综合识别能力优于其它任何参加测试的算法.

#### 1.1.3 特征脸方法小结

如今特征脸方法用于人脸识别仍存在如下一些弊病:首先,由于作为一种图象的统计方法,图象中的所有像素被赋予了同等的地位,可是角度、光照、尺寸及表情等干扰会导致识别率急剧下降,因此较好的识别算法<sup>[19]</sup>都对人脸进行了矫正处理,且只考虑裸脸;其次,根据文献[2],人脸在人脸空间的分布近似高斯分布,且普通人脸位于均值附近,而特殊人脸则位于分布边缘.由此可见,越普通的人脸越难识别,虽然特征脸的方法本质上是抓住了人群的统计特性,但好的表达能力不等于好的区分能力;特征脸虽反映了特定库的统计特性,但不具有普遍代表性,而广泛的应用,则需要训练出的特征脸具有普遍意义;采用此方法的重要假设是人脸处于低维线性空间,即人脸相加和相减后还是人脸<sup>[2]</sup>,显然这是不可能的,因为即使在定位和尺寸相同的情况下,由于部件的相对位置不同,相加、相减后的人脸也一样存在模糊,因此文献[14]提出形状无关人脸(shapeless face)的概念,即依据脸部基准点将人脸变形到标准脸,再进行特征脸处理.总之,有效的特征脸识别方法需要做大量预处理,以减少干扰.而如何表达,并去除表情因素则是识别的另一关键.

#### 1.2 形状和灰度分离的可变形模型

文献[14]提出了一个形状和灰度分离的模型,即从形状、总体灰度、局部灰度分布3个方面来描述一个人脸(如图1、图2、图3所示).其中,点分布模型(图1)用来描述人脸的形状特征,该点分布模型中是用每点的局部灰度信息(图3是采用耳朵上一点附近的方向投影)来描述人脸的局部灰度特征;然后用点分布模型将图象进行变形,以生成形状无关人脸(图2),再做特征脸分析,从而得到人脸的总体灰度模式分布特征.这种三者相结合的识别方法,识



图 1 提取点分布模型

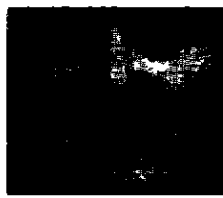


图 2 用于特征脸分析的形状无关脸

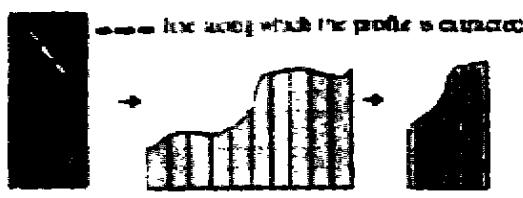


图 3 局部灰度投影提取

别率为 92% (300 个人脸), 虽然该方法作了一些改进, 但构成该方法的基础仍是 KL 变换。一般在特征脸的方法中, 是由行或列扫描后的人脸图象数据来生成特征脸子空间, 这里则对应于 3 种由不同类型参数生成的 3 种特征子空间。该方法首先是循序取每点坐标位置信息, 并将其排列成待训练数据以生成形状特征子空间; 然后对点分布模型的每一点 (如图 3 中耳朵附近一点) 取局部投影信息来代表该点附近的局部灰度特征, 再通过训练后生成与该点相对应的局部灰度分布特征子空间。若将所有人脸的关键点都变形到规定位置, 则生成形状无关人脸, 然后对所有的形状无关人脸进行特征脸分析, 以生成特征脸子空间。虽然每一个特征子空间都可以单独用来识别人脸, 但若完整描述一个人脸, 则需要 3 个特征子空间的人脸参数。文献[14]还试图通过形状特征子空间来分离和表情相关的参数, 而设计形状和灰度分离的模型是希望能够有一个好的人脸模型。试验中, 将这样的模型用于三维姿态复原、身份识别、性别识别、表情识别以及人脸的重建, 均取得了一定的效果。

### 1.3 基于小波特征的弹性匹配方法

#### 1.3.1 基本原理

在 KL 变换中, 待识别人脸  $X$  和库中人脸  $C$  之间采用了通常的欧氏距离来进行匹配。虽然欧氏距离计算简单, 但是当  $X$  和  $C$  只有位移、膨胀 (如 affine 变换) 或是表情不同时, 则欧氏距离不会等于零, 甚至很大, 此外, 若  $C$  作为人脸库中的已知人脸模板, 应该是描述人脸的关键特征, 它的维数并不需要和待识别人脸一样, 因而此时欧氏距离就不合适; 而弹性图匹配法是在二维的空间中定义了这样一个距离, 它对通常的人脸变形具有一定的不变性, 也不要求  $C$ 、 $X$  维数一定相同。可采用属性拓扑图来表达人脸 (图 4 采用的是规则的二维网格图), 其拓扑图的任一顶点均包含一特征矢量, 它记录了人脸在该顶点位置的分布信息 (如图 5), 如文献[11]中介绍的二维拓扑图的顶点矢量就是人脸经小波变换后的

特征矢量。在图象的敏感位置 (如轮廓线、突出点等), 小波变换后生成的特征矢量的模较大。用拓扑图分别代表已知和待识别人脸, 还可根据匹配拓扑图算出它们的“距离”, 作为人脸的相似度准则。由于篇幅所限, 详细的拓扑图生成过程文献[11]、[15]。



图 4 定义在人脸上的二维拓扑图

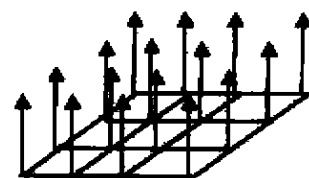
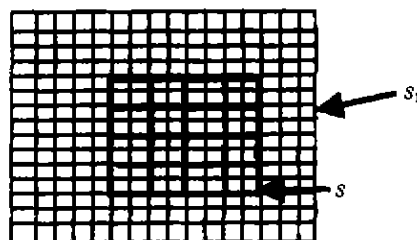


图 5 表达人脸特征的二维矢量

人脸的相似度可用拓扑图的“距离”来表示, 而最佳的匹配应同时考虑顶点特征矢量的匹配和相对几何位置的匹配。由图 6 (和图 5 一样, 它们的每一顶点均为一特征矢量) 可见, 特征匹配即:  $S_1$  上的顶点  $i$ , 与  $S$  中相对应的顶点  $j$  ( $j=M(i)$ ,  $M$  为匹配函数), 其特征的匹配度则表示  $i$  和  $j$  顶点的特征矢量相似度, 而几何位置的匹配则为  $S$  中相近的两顶点, 匹配后,  $S_1$  中对应的两顶点也应该相近, 因此文献[11]用了以下能量函数  $E(M)$  来评价待识别人脸图象矢量场和库中已知人脸的矢量场之间的匹配程度

图 6 定义在已知脸和待识别脸上的二维网格  $S$  和  $S_1$

$$E(M) = \sum_i \left[ -\frac{\langle C_i, X_i \rangle}{\|C_i\| \|X_i\|} \right] + \lambda \sum_{i_1, i_2} [(i_1 - i_2) - (j_1 - j_2)]^2 \quad (5)$$

式中的第一项是计算两个矢量场中对应的局部特征  $X_i$  和  $C_i$  的相似程度,第二项则是计算局部位置关系和匹配次序。由此可见,最佳匹配也就是最小能量函数时的匹配。

在求能量函数实现匹配的时候,可以有如下两种匹配的方法:其中一种是严格的匹配方法;另一种匹配即所谓弹性图匹配方法(见图7)。由图7可见,网格  $S$  经过了变形,即由原来网格  $S$  中的一点对  $S_1$  中一点的严格匹配,变成了  $S$  中一点和  $S_1$  中一点领域范围内的匹配,其目的是为了进一步减小能量函数,通过最终收敛到一个最小值,来实现弹性匹配。正是这样的匹配容忍了表情的细微变化。



图7 弹性匹配

根据 Jun Zhang<sup>[15]</sup> 对综合 MIT、Olivetti、Wwizmann、和 Bem 等人脸库所形成的包括 272 幅照片的综合人脸库,分别用 KL 方法和弹性匹配方法进行识别试验比较<sup>[15]</sup>,所得的识别率分别为 66% 和 93%。其中 KL 变换的识别率很低,其原因主要是由于综合库里来自 4 个人脸库的人脸图象在光照上有很大的差异所造成的,文献[15]之所以作出了弹性图形匹配优于 KL 变换的结论,其原因之一是由于拓扑图的顶点采用了小波变换特征,因为它对于光线、变换、尺寸和角度具有一定的不变性。大家知道,小波特征分析是一种时频分析,即空间-频率分析,若空间一点周围区域的不同的频率响应构成该点的特征串,则其高频部分就对应了小范围内的细节,而低频部分则对应了该点周围较大范围内的概貌。根据该原理,文献[20]提出了用数学形态学上的腐蚀扩张方法形成的多尺度(多分辨率)特征矢量来取代小波特征,并证明了它具有和小波特征相似的效果,它能够反映空间一点周围的高低频信息。现已证明,弹性图形匹配能保留二维图象的空间相关性信息,而特征脸方法在将图象排成一维向量后,则丢失了很多空间相关性信息。这些都是弹性匹配方

法优于特征脸方法的原因,如向人脸库中加入新的人脸时,由于不能保证已有特征脸的通用性,因而有可能需要重新计算特征脸;而对于弹性匹配的方法,则不需要改变已有的数据,通过直接加入新的模板数据即可,但计算较复杂是弹性匹配的一大缺点。根据引言中提出的低层次特征和高层次特征的定义,这里的小波特征类似于外界景物在人眼视网膜上的响应,属低层次特征,没有线、面、模式的概念。由于低层次特征中信息的冗余不仅使得计算复杂,而且由于大量与识别无关的信息没有过滤掉,因而识别率会大打折扣,另外特征脸也存在这样的问题,其中典型的无用信息就是头发。

针对弹性匹配方法的缺陷,可从以下两方面进行改进:一是降低计算复杂度,即对表达人脸的二维矢量场进行特征压缩和提取;二是减少冗余信息,即将所提取出来的低层次特征和高层次特征(如眼角、鼻端的位置等)结合起来,以突出关键点的识别地位。

### 1.3.2 对弹性匹配方法的改进及分析

文献[20]提出了一种弹性匹配的改进方法,即将 KL 变换应用于小波变换,来生成二维网格中顶点的矢量串,以减少其维数,从而大大减少了表达一幅人脸所需要的特征数量,而识别率不会明显下降。

文献[21]是采用人脸基准点,而不是采用二维网格作为拓扑图的节点,同时节点特征也是小波变换特征,即它忽略了除重要人脸部件以外的特征数据,把研究的重点直接定位到感兴趣的区域(参照图8)。

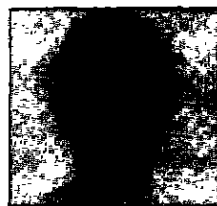


图8 基于部件的拓扑图

文献[21]还采用了和文献[11]不同的结构来存储人脸特征(如图9所示)。



图9 人脸特征库存储结构

由于文献[11]中特征库的存储是面向人脸的,即对每一张人脸都需要存储描述该人脸的整个拓扑图,因而导致了人脸的特征库很庞大,文献[21]中特征库的存储是面向人脸基准点的(如图9),且对应每个基准点有一串的特征矢量,当由某一人脸的对应基准点提取出来的矢量不同于库中已有的任意矢量时,就添入到该结构中存储起来,并编号.这样识别每个人脸只需知道人脸对应基准点在该存储结构中的特征矢量序号即可.该存储结构一个主要优点是,由于不同人脸在同一个基准点所对应的特征矢量可能相同,因此和面向人脸的存储形式相比,数据量会大大减少;另一优点是该存储结构有很强的表达潜力,设有10个基准点,如库中每一基准点都存储了50个特征矢量,那么该存储结构能表达 $50^{10}$ 个不同的人脸.由此可见,文献[21]对文献[11]方法的一大改进是结合了人脸的高层次特征.

另外,弹性匹配方法在实现时,需要考虑具体的参数选择,如二维网格的大小、小波变换参数的选择等,这些参数都会影响识别的效果.毫无疑问,有效的识别效果依赖于关键识别信息的提取,如采取多大的人脸分辨率?能否对提取出来的特征(具体的或抽象的)进行筛选?经验知识使我们关注人脸部件及其附近的特征,而能否再次对这些特征进行筛选?并有何依据?文献[2]正是希望能够回答这些问题.

文献[2]的方法称为紧凑多层图形方法,它是采用三维的拓扑图来表达人脸(如图10).

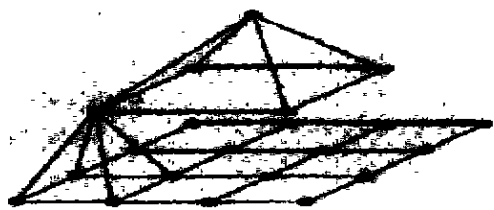


图 10 表达人脸的三维拓扑图

该图构成了一个金字塔的人脸模型,而且每一层中节点的特征矢量也是小波变换的结果.通过这样的金字塔模型就实现了同一个人脸的多分辨率表达.另外,文献[2]有如下两点创新:(1)将高低层特征联系起来,并通过手工选择一些关键点(如眼角、嘴角等)来定位三维拓扑图,同时去除了背景、头发等所在节点;(2)对三维拓扑图的特征进行了特征选择,选出了活跃的特征(包括节点内的特征分量和不同节点之间两种活跃性能比较),还去除了相当多的贡献不大的特征,从而形成了人脸的稀疏表达.由于

特征选择后,不同人脸的拓扑图保留的节点不完全一样,因此用于比较的两个人脸的三维拓扑图在数值上和结构上都不相同,为此,文献[2]定义了一种距离来计算它们的相似度.为提取活跃特征,我们曾尝试利用那些手工提取的关键点,来生成训练库的形状无关模型(不是形状无关人脸),即通过插值小波变换后生成的二维拓扑图来形成人脸的连续表达模型,并假设所有人的脸内差异相同(即表情等),然后根据训练库的统计形状无关模型,在一人一张照片的情况下,估计出个人表达模型中的活跃特征.打个比方,人的眼睛都是相似的,假设眼睛的分布为高斯分布,那么一个眼睛离平均眼睛越远,这个眼睛的特征就越显著,即,若有一定的与众不同性,就可以认为是该人的活跃特征,详细内容参考文献[2],该文有很多创新,它是以人脑对人脸的识别为依据,因此有很好的参考价值.

通过上述的介绍分析,可看出弹性匹配方法比特征脸识别方法前进了一大步.它是采用小波变换特征来描述人脸的局部信息,并和人眼视网膜对图象的响应相似<sup>[2]</sup>,而且一定程度上容忍光线等干扰,对细微表情也不敏感.而且弹性匹配中的人脸模型还考虑了局部人脸细节,并保留了人脸的空间分布信息,且它的可变形匹配方式一定程度上能够容忍人脸从三维到二维投影引起的变形.目前还没有见到国内有利用弹性匹配进行识别的相关报道,但是从国外众多的关于弹性匹配的研究结果来看,它在人脸识别众方法中具有重要地位.

#### 1.4 传统的部件建模的方法

文献[8]认为在人脸识别中,模型匹配方法要优于基于相对距离的特征分析方法.尽管如此,传统的部件分析方法还是被一些研究室用于人脸识别,究其原因,一方面是由于其它方法还处于摸索阶段,另一方面是利用曲线去拟合部件、分析部件的形状比较直观,也容易取得一定的成果<sup>[6]</sup>.

在各种人脸识别方法中,定位眼睛往往是人脸识别的第一步,由于两眼睛的对称性以及眼珠呈现为低灰度值的圆形,因此在人脸图象清晰端正的时候,眼睛的提取是比较容易的,如从400幅人像库上可取得96%的眼睛定位率<sup>[6]</sup>,但是如果人脸图象模糊,或者噪声很多,则往往需要利用更多的信息(如眼睛和眉毛、鼻子的相对位置等),而且这将使得眼睛的定位变得很复杂.由于通常眼睛的形状模型为椭圆<sup>[22]</sup>,嘴巴的形状模型为抛物线<sup>[22]</sup>,因此椭圆和

抛物线的参数和位置能够用作表达当前人脸的特征,文献[6]考虑到眼睛用椭圆表达过于简单,故又采用了二值化,并通过跟踪以得到眼睛形状的方法,由于眉毛和脸形的形状具有任意性,因此在一些研究中曾采用 snake 动态能量曲线来逼近形状<sup>[5,12]</sup>,如脸颊的形状采用了折线,下巴采用抛物线的模型.这些都是传统的提取和分析形状的方法.虽然人脸是刚体,可实际图象中,部件未必轮廓分明,有时人用眼看也只是个大概,计算机提取就更成问题.另外,由于用抛物线、椭圆或者直线作为模型也不能很好的表达反映变化多端的人脸部件,且由于人脸识别还受到表情的影响,且不能在模型中表达表情,因而导致描述同一个人的不同人脸时,其模型参数可能相差很大,而失去识别意义,这也是部件建模识别方法近年受冷落的原因.尽管如此,在正确提取部件以及表情变化微小的前提下,该方法依然奏效,因此在许多方面仍可应用,如对标准身份证照的应用.

## 2 人脸识别方法的分析和总结

### 2.1 特征来源以及特征的后处理

众所周知,人脸的结构大体相同,所不同的是一些细节上的差异,原始的人脸图象不仅数据庞大,而且还会随着拍摄条件及表情神态变化而变化,这就使得人脸的识别成为模式分析中的一个难题.一般从人脸图象上进行有效的识别需要提取稳定的人脸特征,目前所利用的特征可以概括为形状、灰度分布、频域特征3种.其中,形状特征包括人脸各部件的形状以及人脸各部件之间的相对位置,这是最初研究所采用的特征;灰度分布特征,即将人脸图象看成一维或二维灰度模式,所计算出的不同灰度模式之间的距离就是整体的灰度分布特征,例如特征脸的方法,此外还有描述局部关键点领域的灰度分布特征的分析方法;频域特征,即将人脸图象变换到频域后所做的特征脸分析方法就是频域特征脸方法,此时的特征即为频域特征,如小波特征就是频域特征.虽然形状特征是3个特征中最具体形象的特征,但是它也和灰度特征一样受到光照、角度和表情的影响,而频域特征虽然相对较稳定,但作为低层次特征,不易直接用于匹配和识别,因此对它进行进一步的解释是目前需要解决的问题.

在弹性匹配中,若对每个节点运用KL变换,则能够减少特征数,而不降低识别率.其特征后处理的

一个重要方面是特征的选择,也就是需选出最活跃的特征,去除对识别不重要的信息.在人脸识别的特征选择中,生物心理学家首先研究了人脸各部件对识别的重要性,接着文献[2]从模式识别的角度出发,结合人脸各部件信息,并运用最大后验概率,对表达人脸的低层次特征进行了筛选,从而减少了人脸信息的存储量,并改善了识别的效果.

### 2.2 人脸的定位问题

虽然人脸定位问题是人脸识别的第一步,但在前面介绍各种人脸识别方法的时候,并没有介绍具体的定位问题.事实上,对大多数方法而言,人脸的定位过程也就是人脸识别特征的生成过程,而且定位算法也是和识别算法密切相关的.为了说明这一点,下面给出一些人脸识别所采用的定位方法:

方法1 特征脸的方法也可用于定位人脸,这是因为人脸模式在特征脸子空间的投影系数基本相似,若先将子图在特征脸空间投影后重建,然后比较原图和重建图,就能够说明原图是否是人脸,这是因为特征脸空间能反映人脸的分布,而对于非人脸则没有很好的表达力,因此重建图和原图的差异会较大.

方法2 最初的模板匹配方法是直接计算人脸图象和模板人脸图象之间的相似度,匹配最好时,即是人脸在原图中的位置,如弹性图形匹配中,采用的也是一种模板匹配,但是其参与匹配的是用小波特征表达的二维(或三维)拓扑图,若将模板拓扑图在全图生成的拓扑图上移动匹配(严格的或弹性的),则其最佳匹配就给出了人脸的位置,如文献[2]就采用了多分辨率的三维模型,其定位的时候是从最低分辨率开始定位,然后依次增加分辨率,直到位置不变为止,这是由于文献[2]考虑的是定位的分辨率可以远小于识别所需要的分辨率.

方法3 在灰度和形状分离的可变形模型中,很关键的一步是将形状模型自动定位到未知图中,因此匹配采用了ASM(active shape model)的方法<sup>[14]</sup>,其在训练阶段,所有的形状模型都是手工定位的(见图1),由此生成形状特征子空间,且任何形状都可表达为 $X_i = \bar{X} + Pb$ ,其中, $\bar{X}$ 是平均形状; $P$ 是由训练生成的形状特征子空间的基所排列成的矩阵;也可将 $(X_i - \bar{X})$ 用形状特征子空间表达; $b$ 为坐标系数矢量,一般通过对训练库进行统计分析,就能够给出 $b$ 的统计特性,如均值、方差等,当 $b$ 在合理的范围内变化的时候,可以认为其能够表达所有人



的形状特征,具体的定位是采用形状模型和模型上点的局部灰度分布相结合的方法来实现的,而匹配度的计算则是通过用模型上点的当前局部灰度分布和经验灰度分布匹配的加权和来表示,在达到最佳匹配时,待识别人脸的形状模型参数和模型上每点的局部灰度分布参数也同时确定了,所以说,定位的过程也就是识别的过程。

方法 3 在定位人脸的同时也就定位出了具体的部件位置,虽然方法 1 和方法 3 的基本原理不需要定位人脸的部件,而依赖于部件分析来进行人脸识别的方法<sup>[6]</sup>通常是应用一些先验知识(如眼睛的投影直方图形状、人脸的部件分布比例等)来初步给出人脸的大致位置,然后再精确定位人脸的各个部件,这里部件的定位通常使用投影方法、hough 变换的方法以及构造模型能量函数的匹配方法。

### 2.3 识别效果的比较

由于采用的人脸库不同,因此不同识别算法之间的优劣没有可比性,前面的论述也是尽量从理论上进行比较,根据 Moghaddam 等在 1996 年进行的 FFEIT 人脸库测试<sup>[19]</sup>,结果说明区别脸内差异和脸间差异的 Bayesian 特征脸方法的表现最佳,即从 5 000 幅待识别人像中,第一候选的识别率为 89.5%,而灰度和形状分离的可变形模型在 300 幅人像中的识别率达到 92%,另根据文献[15]的测试,在 2 000 幅人脸图象的综合库中,利用小波特征弹性图形匹配的方法获得了 93% 的识别率,而 PCA 识别率只达 66%。

## 3 结 论

人脸识别是一个跨学科富挑战性的前沿课题,但目前人脸识别还只是研究课题,尚不是实用化领域的活跃课题,人脸识别难度较大,主要难在人脸都是有各种变化的相似刚体,由于人脸部件不仅存在各种变形,而且和皮肤之间是平缓过渡,因此人脸是不能用经典的几何模型来进行识别分类的典型例子,如今人脸识别研究人员已经慢慢地将研究重点从传统的点和曲线的分析方法,过渡到用新的人脸模型来表达和识别人脸,其中弹性图匹配就是较成功的尝试,虽然人脸识别算法的开发需要工程人员的努力,但也和解剖学、生理学等的研究密切相关,从目前的研究成果来看,就二维图象而言,成功的人脸识别至少需要考虑以下几个方面:(1)由于外部干

扰不可避免,预处理的效果将会影响到识别结果,好的人脸模型应能够在识别的同时,抑制分离外在干扰的影响;(2)细节是区分不同人脸的关键,因此很多识别方法都十分注重细节,如弹性图匹配中的局部细节,就是通过节点的小波变换特征加以表达,而在灰度形状分离的可变形模型中,局部灰度投影分布也描述了人脸细节,另外,传统的点和曲线的方法更是直接从局部细节入手,可是特征脸方法则缺少对细节的考虑,故需和别的方法相结合,才能取得好的识别效果;(3)在匹配的时候,不仅要考虑各种因素所导致的人脸微小变形,而且在容忍变形的同时,还不能损害到人脸识别的有效性,如弹性图匹配的方法不论从特征的选择上,还是从匹配的方法上都力图遵循这一原则,由此可见,人脸变形在人脸识别中具有重要意义,因为人脸丰富的变形就是导致传统的点线分析方法失败的原因;(4)对于表达人脸的各种特征需要进行比较和选择,以找出人脸最活跃的特征,这可以通过如下两种途径:一是比较同一个人的多张图片,以得到稳定的特征;另一种方法就是比较不同人的图片,以得出该人最“与众不同”之处<sup>[2]</sup>。

此外,实用的识别系统还必须考虑计算复杂度,现有的识别方法中,通过从人脸图中提取出特征串,来对数据库进行检索的方法速度快,而利用拓扑属性图匹配来确定匹配度的方法则相对慢,而且随数据库增加,前者的识别率要比后者下降得快,因此改进的思路是将两者相结合,首先用快速的特征串匹配,来缩小检索范围,再进行拓扑图慢匹配,此外,用减小拓扑图存储量的方法也能够加快匹配速度,但这需要提取有效特征和去掉冗余信息。

本文介绍和分析的各种人脸识别方法同样可用于摄像机输入人脸的识别,而对于摄像机图象而言,人脸的定位和表情的分析还可以利用序列图象之间的相关性信息,如从摄像机输入动态图可以进行二维及三维的运动估计,从而建立三维的人脸模型,由于从摄像机动态输入图中得到的信息很多,故还有可能进行有效的表情分析,以作为身份辨认的辅助手段,本文只是对目前应用于人脸识别的技术作了选择性的介绍,也是对文献[3]、[15]的一点补充,由于人脸识别的理论还不完善,具体算法的实现也有很多的因素待研究,因此计算机人脸识别的实用化还需要众多研究人员的不懈努力。

## 参考文献

- 1 Phillips P J, Moon H, Rauss P *et al*. The FERET evaluation methodology for face-recognition algorithms. In: IEEE Computer Society Conference on CVPR, June 17-19 1997; 137~143.
- 2 Grudin M A. Compact multi-level representation of human faces for recognition. Ph. D dissertation, Liverpool John Moores University, Nov. 1997.
- 3 Chellappa R, Wilson C L, Sirohey S. Human and machine recognition of faces; A Survey. Proceedings of the IEEE, 1995, 83(5); 705~740.
- 4 Turk M, Pentland A. Face recognition using eigenfaces. In: Proc. of IEEE Conf. on CVPR, 1991; 586~591.
- 5 Valentin D, Abdi H, O'Toole A J *et al*. Connectionist models of face processing; A Survey. Pattern Recognition, 1994, 27(9); 1209~1230.
- 6 严超, 苏光大. 人脸特征的定位与提取. 中国图象图形学报, 1998, 3(5); 375~379.
- 7 W Bledsoe. Man-machine facial recognition. Panoramic Research Inc, Palo Alto, CA, 1956. Rep PR1; 22.
- 8 Berto R, Poggio T. Face recognition: Feature versus templates. IEEE Trans. on PAMI, 1993, 15(10); 1042~1052.
- 9 Ziquan Hong. Algebraic feature extraction of image for recognition. Pattern Recognition, 1991, 24(3); 211~219.
- 10 Nakamura O, Mathur S, Minami T. Identification of human faces based on isodensity maps. Pattern Recognition, 1991, 24(3); 263~272.
- 11 Lades M, Vorbuggen J, Buhmann J *et al*. Distortion invariant object recognition in the dynamic link architecture. IEEE Trans. on Computers, 1991, 42(3); 300~311.
- 12 Samaria F, Young S. HMM-based architecture for face identification. Image and Vision Computing, 1994, 12(8).
- 13 Lam K M, Yan H. An analytic-to-holistic approach for face recognition based on a single frontal view. IEEE Trans. on PAMI, 1998, 20(7).
- 14 Lanitis A, Taylor C J, Cootes T F. Automatic interpretation and coding of face images using flexible models. IEEE trans. on PAMI, 1997, 19(7).
- 15 Jun Zhang, Yong Yan, Martin Lades. Face recognition: Eigenface, Elastic Matching, and Neural Nets. Proceedings of the IEEE, 1998, 85(9); 1422~1435.
- 16 Luebbers P G, Uwechue O A, Pandya A S. A neural network based acial recognition system. Proc. SPIE, 1994, 2243; 595~606.
- 17 Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs fisherfaces; Recognition using calss specific linear projection. In: IEEE Tran. on PAMI, 1997, 19(7).
- 18 Chenjun Liu, Wechsler H. Probabilistic reasoning models for face recognition. Proceedings of the 1998, IEEE Computer Society Conference on CVPR, Jun 23-25, 1998; 827~832.
- 19 Moghaddam B, Pentland A. Beyond linear eigenspaces; Bayesian matching for face recognition. In: Face Recognition- from Theory to Applications, NATO ASI series: Series F, 1997; 230~243.
- 20 Kotropoulos C, Pitas I. Face authentication based on morphological grid matching. IEEE Inter. Conf. on Image Processing, 1997, 1; 105~108.
- 21 Wiskott L, Fellous J M, Kruger N. Face recognition by elastic bunch graph matching. IEEE Trans. on PAMI, 1997, 19(7).
- 22 Chunglin Huang, Chingwen Chen. Human facial feature extraction for face interpretation and recognition. Pattern Recognition, 1992, 25(12); 1435~1444.



张翠平 1974年生, 1997年获得清华大学电子工程系学士学位。现为清华大学电子工程系硕士生。主要研究方向为图象处理。目前正在从事人脸识别的研究。



苏光大 清华大学电子系副教授。长期从事图象处理系统的科学研究, 目前主要研究方向为人像识别和高速图象处理。先后5次获部委级科技成果奖, 发表论文30余篇。