

基于特征曲线的自动人面识别研究*

章高清 王申康 叶懋东

(浙江大学计算机科学与工程系 杭州 310027)

E-mail: gqzhang@cenpok.net

摘要 提出了一种基于特征曲线的快速人面识别方法.该方法首先对不同姿态和大小的人面进行定位和归一化,然后在归一化后的人面图像上建立多种特征曲线.作为对人面特征的描述,最后利用傅里叶描绘子对特征曲线进行解析提取关键特征,得到人面的表征向量.使用此方法在1300幅人面图像上进行了实验,结果表明,此方法在速度和准确度方面都具有较好的性能,而且对有不同姿态和表情的人面的识别具有一定的鲁棒性.

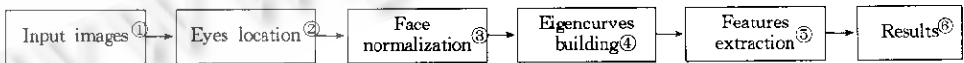
关键词 自动人面识别,人面归一化,积分投影,特征曲线,傅里叶描绘子.

中国法分类号 TP391

人面是人类独一无二的特征,即使一对双胞胎,其面部也一定存在着某方面的差异.虽然人类在表情、年龄或发型等发生了巨大变化的情况下都可以毫无困难地检测和识别出人面,但要建立一个能够完全自动进行识别的系统却是非常困难的,它们涉及到以下几个子问题:(1)把一种模式检测为人面,(2)面部表情分析,(3)基于特征的分类.由于能够实现上述功能的系统具有广泛的应用前景,如嫌疑犯照片识别、丢失儿童信息检索、建立安全系统、信用卡的确认等.在过去的20年中,人面自动识别领域吸引了众多的研究者,但目前能完全解决上述问题的有效系统尚未推出.

现有的自动人面识别方法基本上可分为两类^[1,2]:分离的、基于局部特征的方法和整体的、基于全局特征的方法.前者抽取离散的局部特征作为人面的索引,然后用它们作为度量,采用标准模式识别技术来识别.后者则将整个人面作为一个全局的描述,通常采用模板匹配策略,其特点主要是:互连方法、主元分析(PCA)^[3]、后向反馈的神经网络(BP)算法^[4]以及奇异值分解(SVD)^[5]等.但这些方法一般都需要较大的计算量和一组与特定环境相关的实验参数,并且对人面的姿态、大小或光照条件非常敏感.例如,典型的PCA就存在下面几个缺点:(a)算法利用KL变换,其速度在一般微机上难以忍受;(b)对没有训练过的人面,识别率较低;(c)对人面姿态变化较敏感.

针对上述问题,我们设计并实现了一种新的人面识别方法(eigencurve,简称ECV).ECV通过构造人面的特征曲线进行识别,其主要特点是:(a)与人面姿态、大小及表情无关,但光照条件变化非常大时会影响识别率;(b)速度快(检索每幅图像约需5秒);(c)无需大量的样本学习,减少了系统复杂度,同时增强了系统的鲁棒性.系统只假定:图像中只有一个人面,且背景较为简单.ECV的简单流程如图1所示.



①输入图像,②眼睛定位,③人面归一化,④特征曲线构造,⑤特征抽取,⑥识别输出.

Fig. 1 Face recognition process

图1 人面识别流程图

* 本文研究得到浙江省公安厅资金(No. 1997-7-15)资助.作者章高清,女,1972年生,博士生,主要研究领域为多媒体,数据库,数字图像处理,人工智能.王申康,1945年生,教授,博士生导师,主要研究领域为机器学习,知识获取和专家系统,多媒体/超媒体技术等.叶懋东,1945年生,副教授,主要研究领域为数字图像处理.

本文通讯联系人:章高清,杭州 310027,浙江大学计算机科学与工程系

本文 1998-10-22 收到原稿,1999-04-06 收到修改稿

1 归一化

对大小不同的人面或存在不同倾斜程度的人面图像进行归一化工作是非常必要的,而且归一化的结果直接影响到最后的识别效果.在本系统中,归一化的同时也完成了人面的自动定位,其基本思想是基于一个常常被忽略的事实:即对所有的人面图像,头发和眼睛都很容易与脸颊明显分开,因此,只要人面图像能被很好地分割,就可将眼睛准确地检测出来,并同时完成人面的定位,这样,准化问题便可较容易地得到解决.被归一化后的图像大小为 128×192 像素,人面中两眼在同一水平位置.

1.1 阈值分割

由上所述,首先需对人面图像 $I(x, y)$ 进行阈值分割,使头发、眼睛与脸颊明显分开.由于系统假定图像中只存在一个人面,分割问题变得相对简单.图 2(b) 是人面图像的典型直方图,其中两个明显的峰值分别对应于人面的暗部(头发、眼睛、眉毛、嘴等)和亮部(额头、脸颊等).而且第 1 个峰值应对应于人面的暗部,因为一般背景总是会比人面的暗部要亮,这样,我们只需找到前两个峰值,在其间取一个灰度值作为阈值即可.由于存在多个可选阈值,可通过自动阈值调整,使图像的二值化过程非常可靠.

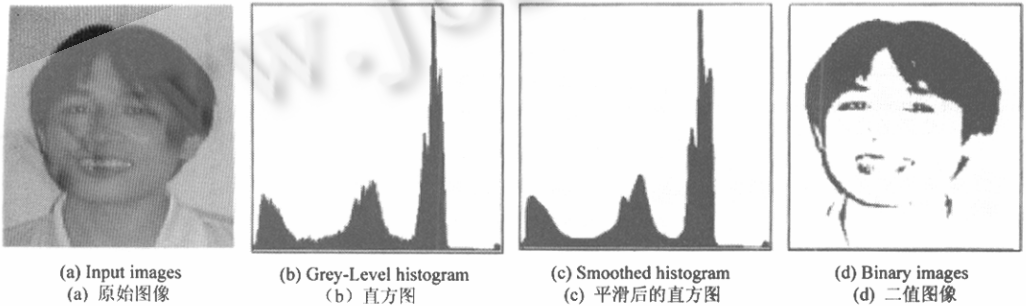


Fig. 2 The thresholding of input images

图 2 原始图像的阈值分割

直方图中常会因图像本身的灰度分布及噪声干扰等因素的影响而出现一些类似峰点的随机干扰点,它们的存在会导致对峰点的误判,从而影响阈值的选取.因此,在自动检测峰点之前,我们需对直方图进行移动平均平滑预处理:

$$G_i = \frac{1}{n}(g_i + g_{i-1} + g_{i-2} + K + g_{i-n+1}) = \frac{1}{n}(g_{i-1} + g_{i-2} + g_{i-3} + K + g_{i-n}) + \frac{1}{n}g_i - \frac{1}{n}g_{i-n}$$

$$= G_{i-1} + \frac{1}{n}(g_i - g_{i-n}),$$

其中 G_i 表示灰度值为 j 点的移动平均值, g_i 表示直方图中灰度值为 i 的像素出现的频数, n 为每次移动地求算术平均值采用的灰度级个数.移动平均的实质是对离散数列加以平滑处理来消除随机扰动点.若一次移动平均仍不能满足需要,可对 G_i (n 取相同值)数列作二次或三次移动平均,方法类似.一般来讲,进行二次移动平均即可达到满意效果.图 2(c) 是对图 2(b) 作二次移动平均后的结果,可见其整体上与原直方图无明显改变,但许多干扰点被消除了.

所以,通过对原始图像直方图作二次移动平均,可以准确地检测出前两个峰值,从而得到可靠的阈值进行分割,使头发、眼睛与脸颊明显分开.图 2(d) 是阈值分割后的二值图像 $B(x, y)$.

1.2 眼睛定位

眼睛定位一般采用模板匹配方法^[1,2].显然,单一模板匹配很难适应人面大小不同的情况,而由 Yille 等人^[6]提出、Xie 等人^[7]进行改进的基于弹性模板匹配的面部特征抽取方法虽然稳定、可靠,但仍存在着计算量大、能量函数很难适应一般情况的缺点.我们提出了一种称为面积法的眼晴定位方法,此方法基于阈值分割后的二值图

像 $B(x,y)$, 利用人面的一些简单先验知识, 可以快速而准确地定位眼睛. 而且, 眼睛定位的过程也是人面定位的过程(如图 3 所示). 以下是眼睛定位的主要步骤:

(1) 通过第 1 次扫描二值图像 $B(x,y)$ 定位头发.

在扫描 $B(x,y)$ 的过程中, 用相邻点标记算法标记出各个 1 值块. 一般来说, 在背景较为简单的条件下, 我们认为, 面积最大的一块即为头发, 这个假设通常都是成立的. 即使对光头或白发人来说, 其头部灰度也是比较小的, 因此用自动阈值调整法也能将头部找出来.

(2) 通过对头发块的处理, 定位头发内边界, 生成人面内轮廓图 $I'(x,y)$.

在定位头发块后, 再作一次从上到下的局部扫描, 可以得到头发块的下边界曲线. 对于头发块中可能出现 0 值斑点的情况, 我们用平均值方法进行修正. 观察这条曲线, 我们发现, 在左右两边各有一个极小值(即在鬓角的部位人们的头发延伸到下方), 找出这两个点. 现在可以再生成一个人面内轮廓图 $I'(x,y)$, 该图顶端为头发下边界曲线的最高点, 左右两端为找到的两个极小值点位置, 底端为顶端加上左右端间距的 1.5 倍与整幅图像上三分之二分界的较小值.

(3) 对人面内轮廓图 $I'(x,y)$ 应用拉普拉斯算子, 作第 2 次扫描, 最后利用人面的一些先验知识定位双眼.

先使 $I'(x,y)$ 中的每点灰度乘以如图 3 所示的矩阵:

| | | |
|---|----|---|
| 1 | 1 | 1 |
| 1 | -8 | 1 |
| 1 | 1 | 1 |

经过上述处理的图像是边界敏感的, 同样对它作阈值分割, 然后进行全图扫描, 方法与定位头发块相同, 这次找出面积最大的 10 个块, 作分对处理, 标记出可能为双眼的成对块, 经过下面的规则判别出最有可能成为眼睛的一对, 然后取出这一对块中最黑的点作为眼睛的中心点.

判别眼睛对所用的主要规则是:

(a) 眼睛块宽度(width)、高度(height)及面积(area)应满足:

$$width > height \text{ and } area > 0.5;$$

(b) 眼睛块到 $I'(x,y)$ 上边界的距离 $d1$ 与其到 $I'(x,y)$ 下边界距离 $d2$ 应满足:

$$d1 < d2 \text{ (眼睛应在人面的上半部)};$$

(c) 眼睛块对应位于图像 $I'(x,y)$ 中心线的两侧.

眼睛定位的结果如图 4 所示.

Fig. 3 The distorted differentia kernel
图3 变形的二阶微分模板

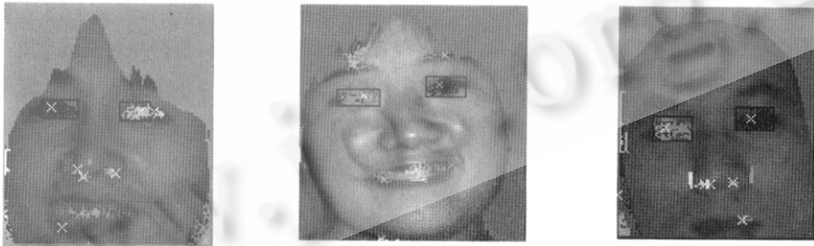


Fig. 4 The examples of eyes location
图 4 眼睛定位示例

1.3 归一化

首先利用眼睛的中心点对人面姿态进行归一化: 将图像 $I'(x,y)$ 随着两眼中心点连线 l 旋转至 l 达到水平位置, 生成图像 $I''(x,y)$; 然后对 $I''(x,y)$ 进行缩放变换, 使图像 $I^{norm}(x,y)$ 最后的尺寸为 128×192 像素, 完成人面大小的归一化. 一般说来, 图像 $I^{norm}(x,y)$ 的顶端在眉毛的上边界, 底端在下唇至下颏间(因人而异), 左右两端均在颧骨附近(如图 5 所示), 这样就使脸部的关键区域凸现出来, 大大缩小了图像识别的范围.

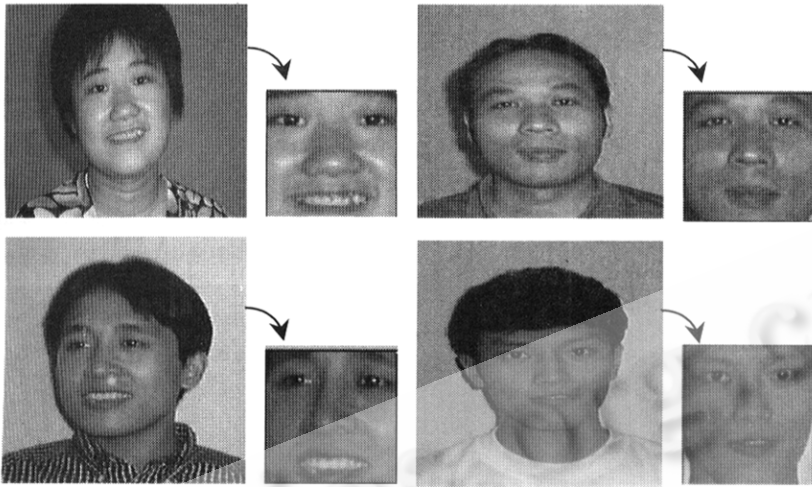


Fig. 5 The examples of face normalization
图 5 人面图像归一化示例

2 人面识别

ECV 人面识别算法是一种基于整体图像的全局特征匹配方法。它在归一化图像上建立若干特征曲线,然后对它们进行傅里叶分析,得到关键的全局特征值,构成人面的表征向量,最后利用这些人面表征向量进行匹配识别。

2.1 特征曲线

特征曲线的思想是基于积分投影方法。在图像的某个区域内进行积分投影的方法在人面识别的早期就已被提出,用于抽取人面局部特征。Takeo Kanade 等人^[8]用此方法抽取人面局部特征(嘴、鼻等),准确率达到 75%。后来, Brunelli 等人^[9]对其作了改进,准确率提高到 80%左右。显然,这种准确率尚未达到令人满意的程度,而我们在采纳了积分投影思想的基础上对其进行了改进,构造出若干种全局积分投影函数,即特征曲线,用来描述人面的特征。

特征曲线可分为 3 类,其构造如下:

(1) 对图像 $I^{\text{norm}}(x, y)$ 分别作行、列积分投影(如图 6(a)所示),得到特征曲线 c_1 和 c_2 。

$$C_1: H(x) = \sum_{y=0}^N I^{\text{norm}}(x, y), \quad (1)$$

$$C_2: V(y) = \sum_{x=0}^M I^{\text{norm}}(x, y). \quad (2)$$

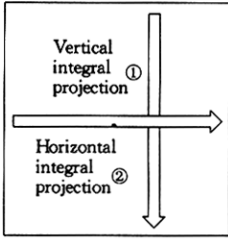
(2) 以图像中心为圆心, q 为半径作一个圆,将该圆半径分为 p 个分点,在每个分点上仍以图像中心为圆心,圆心到该分点的距离为半径作同心圆,将圆每隔 1 度取一个点,共 360 个点的灰度相加,得到曲线 c_3 。显然,该曲线在 X 轴上的总点数为 p (如图 6(b)所示)。设 $g_i(\alpha)$ 表示第 i 个圆上 α 角度处的灰度值,则 c_3 如式(3)所示。

$$C_3: c(i) = \sum_{\alpha=1}^{360} g_i(\alpha), \quad 0 < i \leq p. \quad (3)$$

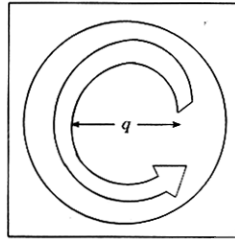
(3) 以图像中心为圆心, q 为半径作一个圆,将该圆半径分为 p 个分点,在每个分点上仍以图像中心为圆心,圆心到该分点的距离为半径作同心圆,将圆每隔 1 度取一个点,分别作上 180 度共 180 个点和下 180 度共 180 个点,将其灰度相加,得到两条曲线 c_4 和 c_5 ,显然,这些曲线在 X 轴上的总点数为 p (如图 6(c)所示), $(0 < i \leq p, g_i(\alpha)$ 同式(3))。

$$C_4: c(i) = \sum_{\alpha=1}^{180} g_i(\alpha), \quad (4)$$

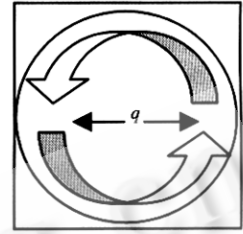
$$C_s:c(i) = \sum_{\alpha=181}^{360} g_i(\alpha). \tag{5}$$



①积分投影,②行积分投影.
(a) The first kind of eigencurves
(a) 第1类特征曲线



(b) The second kind of eigencurves
(b) 第2类特征曲线

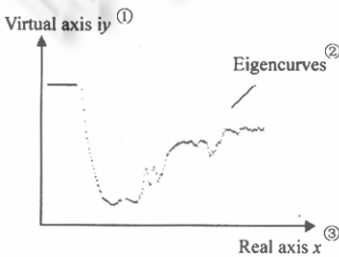


(c) The third kind of eigencurves
(c) 第3类特征曲线

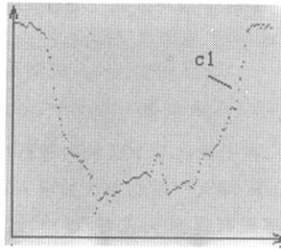
Fig. 6
图6

2.2 特征抽取

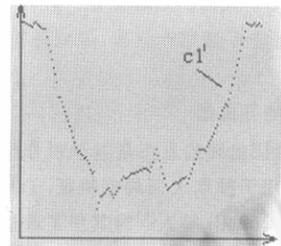
全局特征值的抽取是通过对上述 5 条特征曲线进行傅里叶分析而获得的. 傅里叶描绘子是区域边界的一种重要表示方法,在此我们用它来表示特征曲线.假如某特征曲线上有 M 个点可利用,可以把特征曲线看成是在复平面内(如图 7(a)所示),从特征曲线端点开始,沿曲线跟踪一遍就可得到一个复数序列,这个复数序列的傅里叶变换称为特征曲线傅里叶描绘子.



①虚轴iy,②特征曲线,③实轴x
(a) Eigencurves presentation in plus plane
(a) 复平面上特征表示



(b) Eigencurves
(b) 原特征曲线



(c) The restoration of eigencurves by first ten Fourier descriptors
(c) 前10对傅里叶描绘子得到的特征曲线

Fig. 7 Eigencurves
图7 特征曲线

设 $f(x)$ 表示特征曲线, N 为该特征曲线上的特征点个数,则其傅里叶描绘子可表示为

$$F(n) = \sum_{k=0}^{N-1} f(x) e^{-2\pi i k n / N}.$$

实验表明,前 10 对傅里叶描绘子即能很好地表示特征曲线.如图 7(b),(c)所示, c_1 是原特征曲线, c'_1 为由前 10 对描绘子经傅里叶反变换生成的特征曲线.这样,每条特征曲线只用 10 对傅里叶描绘子来表示即可.将 5 条特征曲线的傅里叶描绘子组合起来,最后得到一个维数为 50 的人面全局特征的表征向量.

2.3 识别

全局特征的抽取使每个人面与一个 50 维的表征向量对应起来,最后通过最近邻分类器进行识别.

两个人面表征向量 V_1, V_2 之间的相似性距离测度定义如下:

$$D(V_1, V_2) = |V_1 - V_2| = \sqrt{(V_1 - V_2)(V_1 - V_2)^T}.$$

采用这一测度可以测量两个人面的相似性.如果已知人面对应的表征向量用 X_1, X_2, \dots, X_L 表示,未知人面的表征向量用 X 表示,那么可计算 X 与各个已知人面的距离 $D(X, X_j)$,如果

$$D(X, X_i) < D(X, X_j) \quad (j=1, 2, 3, \dots, L, \text{且 } j \neq i),$$

则可判定第 i 幅人面即为首选匹配人面.

3 实验结果

测试实验建立在一个人面数据库 FACEDB 之上。FACEDB 包括 260 个人, 每个人有 5 张不同的人面图像, 其中 3 张是不同表情的正面图像, 另两张则是不同姿态(分别向左、向右倾斜)的图像, 如图 8 所示。所以, 在 FACEDB 中共有 1 300 张人面图像, 这些人面图像分成两个互不相交的集合: 样本集和测试集。样本集由 250 个人的两张正面图像(一般选取表情较平和、光线较均匀的人面图像)组成, 作为已知人面图像数据库; 测试集则包括其他所有人面图像, 作为未知待测图像库。



Fig. 8 The face images in database

图 8 数据库中人面示例

测试实验分两部分进行: (1) 测试系统在无“拒绝”条件下的识别性能; (2) 测试系统在引入“拒绝”可能条件下的识别性能。测试性能结果见表 1, 系统的平均识别率达到 95% 左右, 每幅人面的识别速度在 3 秒以内。同时, 实验结果还表明, 系统对人面表情的变化、姿态的变化具有较好的鲁棒性, 但对光照条件的鲁棒性则较差。

Table 1 Test results

表 1 实验结果

| Test ^① | Correct ^② (%) | Incorrect ^③ (%) | Reject ^④ | Threshold ^⑤ |
|--|--------------------------|----------------------------|---------------------|------------------------|
| Test 1 (no rejection) ^⑥ | 94.53 | 5.47 | 0 | — |
| Test 2 (permit rejection) ^⑦ | 96.38 | 2.71 | 1.20 | 0.84 |

①实验, ②识别率, ③错误率, ④拒绝率, ⑤阈值, ⑥测试 1(无拒绝), ⑦测试 2(允许拒绝)。

4 总结

基于特征曲线的 ECV 人面识别方法原理简单, 有效、实用, 尤其无需大量的样本学习, 具有速度快、准确率高的特点。实验结果表明, 此方法对于不同表情、姿态的人面的识别具有很好的鲁棒性。其缺点是, 差异较大的光照条件会影响识别率。

今后, 算法将在以下两方面进行扩展: ① 优化或构造性能更好的特征曲线, 以克服光照条件带来的影响; ② 取消对背景的限制, 以适应更复杂的场景。

参考文献

- Chellappa R, Wilson C L, Sirohey S. Human and machine recognition of faces: a survey. *Proceedings of the IEEE*, 1995, 83(5):705~740
- Samal A, Iyengar P A. Automatic recognition and analysis of human faces and facial expressions; a survey. *Pattern Recognition*, 1992, 25(1):65~77
- Turk M, Pentland A. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 1991, 3(1):71~86
- Hong Z. Algebraic feature extraction of image for recognition. *Pattern Recognition*, 1991, 24(3):211~108
- Narendra K S, Parthasarathy K. Identification and control of dynamic system using neural networks. *IEEE Transactions on Neural Network*, 1990, 1(1):4~27
- Yuille A L, Conen D S, Halinan P W. Feature extraction from faces using deformable templates. In: Harvard Robotics Lab. ed. *Proceedings of the CVPR'89 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, D C: IEEE Computer Document Press, 1989. 104~109

- 7 Xie X, Sudhakar R, Zhuang H. On improving eye feature extraction using deformable templates. Pattern Recognition, 1994,27(6):791~799
- 8 Kanade T. Picture processing by computer complex and recognition of human faces. Technical Report, Department of Information Science, Kyoto University, 1973
- 9 Brunelli R, Poggio T. Face recognition: features versus templates. IEEE Transactions on Pattern Analysis Machine Intelligence, 1993,15(4):1042~1052

Automatic Human Face Recognition Based on Eigencurves

ZHANG Gao-qing WANG Shen-kang YE Mao-dong

(Department of Computer Science and Engineering Zhejiang University Hangzhou 310027)

Abstract A fast method for face recognition based on eigencurves is presented in this paper. First all faces with different poses and sizes are located and normalized. Then several kinds of eigencurves are built to describe characteristics of faces. And finally, the dominant characteristics are extracted which are composed of representative vectors of faces through analyzing the eigencurves using Fourier describer. This method has been tested on 1 300 facial images and shown a good performance of speed and accuracy, and the result also shows it is robust for various head poses and expressions.

Key words Automatic face recognition, face normalization, integral projection, eigencurve, Fourier describer.