

一种多特征提取及融合的虹膜识别方法

张震, 刘博, 李龙

(郑州大学 电气工程学院, 河南 郑州 450001)

摘要:经典虹膜识别算法只提取单一的虹膜特征,提出一种采用多种特征提取及融合的虹膜识别算法.该算法首先对采集到的虹膜图像进行预处理,然后分别采用 Log-Gabor 滤波和 Haar 小波对预处理后图像进行特征提取,并用 Hamming 距离和加权欧式距离计算得到两个相似度,通过支持向量机(support vector machine, SVM)对相似度进行融合,得到匹配结果.实验证明,该算法有效,并能取得较高的识别率.

关键词:虹膜识别;多特征提取;Log-Gabor 滤波;Haar 小波;特征融合

中图分类号: TP391 **文献标志码:** A **doi:**10.13705/j.issn.1671-6833.2017.01.004

0 引言

随着信息技术和网络技术的发展,钥匙、密码等传统的身份认证方法易遗失、易伪造,已经不能满足现代社会生活的需要,而有效、可靠的身份认证方法是保证信息系统安全的必要前提.基于生物特征识别的身份认证中,虹膜识别作为重要的生物特征识别技术,具有普遍性、唯一性、稳定性、非侵犯性等优点,已成为近些年的热点研究课题.虹膜识别原理为对虹膜采集设备采集到的虹膜图像进行处理、特征编码和模式匹配来实现身份的识别认证,因此系统可以分为虹膜图像采集、虹膜图像预处理、特征提取及模式匹配 4 个过程.

目前常用的虹膜识别方法主要由以下几种经典的方法发展而来:Daugman^[1]首先采用边缘检测和 Hough 变换的方法定位分割虹膜图像,并提出了基于二维 Gabor 滤波虹膜特征提取方法,将滤波结果的相位信息进行编码,并通过 Hamming 距离计算分类.Wildes 等^[2]采用微积分算法确定虹膜内外边界,并采用拉普拉斯金字塔算法获得不同分辨率的虹膜纹理图像,然后计算相关系数并利用 Fisher 线性变换进行分类.Boles^[3]提出了基于小波变换过零点的算法,检测小波变换的过零点以及相邻过零点之间的小波变换积分后的平

均值作为虹膜特征,构造相似形函数进行匹配.此外还有 Lim^[4]提出了以 Haar 小波对虹膜进行滤波分解并二值化编码,以及基于多尺度 Gabor 滤波器的虹膜识别算法,基于灰度共生矩阵虹膜识别算法,基于尺度不变特征转换的虹膜识别算法等.

与只提取单一的虹膜特征相比,笔者提出了一种多特征提取及融合的虹膜识别方法,与文献[5]在提取阶段进行特征融合不同,笔者分别采用一维 Log-gabor 滤波器和二维 Haar 小波提取虹膜的两种特征,在匹配阶段进行特征融合,在通过使用 Hamming 距离和加权欧氏距离与虹膜特征库匹配计算相似度的基础上,构造支持向量 SVM,对计算得到的两个相似度进行融合,实现识别分类,得出识别结果.在匹配阶段采用相似度融合的方法,即在提取阶段兼顾了虹膜局部特征和全局特征,并且有效地解决了不同算法提取特征的异构问题,本文算法系统框图如图 1 所示.

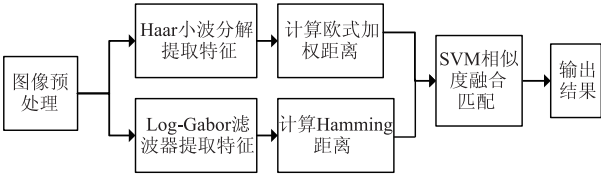


图 1 虹膜识别方法框图

Fig. 1 Block diagram of iris recognition method

收稿日期:2016-10-19;修订日期:2016-11-13

基金项目:河南省科技攻关计划项目(152102210032)

作者简介:张震(1966—),男,河南郑州人,郑州大学教授,博士,主要从事信号处理,图像处理与模式识别研究,
E-mail:zhangzhen66@126.com.

1 虹膜图像的采集

虹膜图像的采集环境复杂,高质量的虹膜图像使识别算法能够提取到足够多的虹膜特征,以便在模式匹配阶段有较高的识别率.但是由于虹膜直径约为 11 mm,面积较小,在采集过程中易受到各种因素的干扰,使采集到的虹膜图像有光斑、旋转、偏移、睫毛干扰、眼皮遮盖等现象,影响虹膜特征提取,降低了虹膜的正确识别率,所以虹膜图像采集技术是虹膜识别关键技术之一.由于虹膜采集设备的复杂性,笔者只对虹膜识别算法进行研究,采用中科院自动化研究所提供的 CASIA V1.0 虹膜图像数据库作为实验样本,进行虹膜识别算法的实验.

2 虹膜图像预处理

原始和定位后的虹膜图像如图 2 所示.采集得到的虹膜图像中不仅仅有虹膜,还包括了眼睑、睫毛、巩膜、瞳孔等部分,如图 2(a)所示,因此不能直接用于虹膜图像的特征提取和模式匹配,需要对采集到的虹膜图像进行预处理.虹膜图像预处理包括虹膜定位分割、图像归一化和图像增强 3 个步骤.

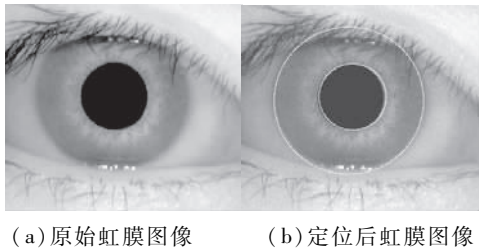


图 2 虹膜图像
Fig. 2 Iris image

2.1 虹膜图像的定位分割

虹膜图像定位分割即通过定位算法在采集到的虹膜图像上确定虹膜与瞳孔边界和虹膜与巩膜边界,即虹膜内、外边界的圆心及半径,根据得到的定位结果进行图像分割,如图 2(b)所示.笔者首先确定虹膜内边界,然后确定外边界,定位分割算法步骤如下.

1) 建立采集到的虹膜图像的灰度直方图,因为图像中瞳孔像素灰度小于虹膜灰度,虹膜灰度又大于巩膜灰度,所以,根据灰度直方图中灰度的峰谷值计算阈值,并根据确定的阈值对虹膜图像进行二值化操作,如图 3 所示.

2) 应用形态学中闭运算算法对二值化的虹膜

图像进行处理,如图 4 所示,去除睫毛等噪声^[6].

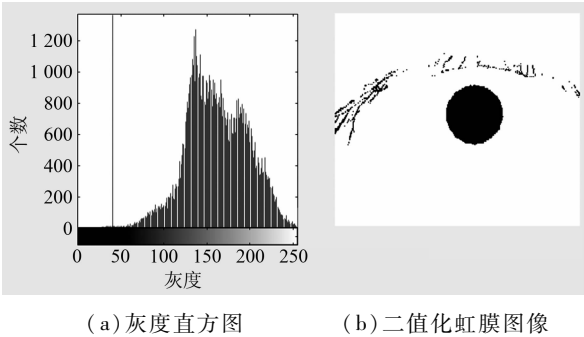


图 3 二值化的虹膜图像
Fig. 3 Binarized iris image

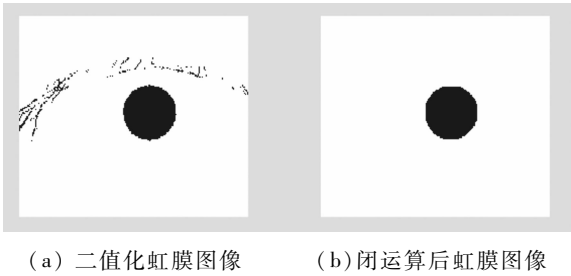


图 4 闭运算后的虹膜图像
Fig. 4 Iris image after closing operation

3) 将闭运算处理后的虹膜图像按行和列进行像素点灰度累加计算,峰值即瞳孔内一点,通过 Canny 边缘检测算法提取边缘,由已知的瞳孔内一点作水平和垂直的两条直线与虹膜内边缘相交,根据已确定的 4 个边缘点和圆形的几何特性^[7]求取虹膜内边界和圆心.

4) 由确定的内边界在虹膜图像的边缘二值化图像上截取感兴趣区域用于计算虹膜外边缘,如图 5(b)所示,以最大可能排除睫毛、眼睑等干扰,提高虹膜定位算法的效率,采用最小二乘拟合圆算法求取虹膜外边界及圆心.

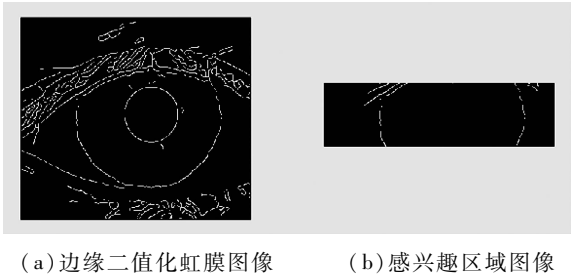


图 5 虹膜边缘检测图像
Fig. 5 Iris edge detection image

2.2 图像归一化

采用 Rubber Sheet 模型将分割得到的近似环形的虹膜图像由直角坐标转换为极坐标.

$$I(x(r, \theta), y(r, \theta)) \rightarrow I(r, \theta). \quad (1)$$

如图 6 所示,用 Rubber sheet 模型消除采集虹膜图像时的平移、缩放和旋转等干扰,为虹膜图像的特征提取做准备,提高编码效率。

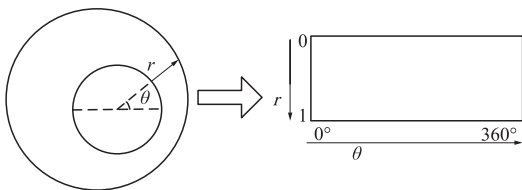


图 6 Rubber Sheet 模型

Fig. 6 The Rubber Sheet model

2.3 图像增强

由于归一化后的虹膜图像对比度较低,并不利于直接提取特征纹理,因此需要对归一化的图像进行处理,增强图像对比度,笔者采用直方图均衡化^[8]的方法,如图 7 所示,增强虹膜纹理的对比度。

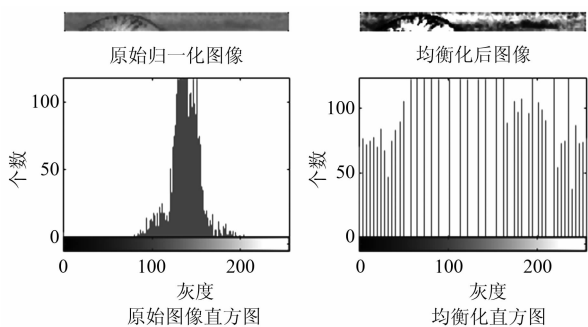


图 7 直方图均衡化

Fig. 7 The histogram equalization schematic

3 虹膜特征提取和匹配

3.1 Log-Gabor 滤波和 Hamming 距离

Field 提出的 Log-Gabor 滤波器^[6]:

$$G(f) = \exp\left(\frac{-(\log(f/f_0))^2}{2(\log(\beta/f_0))^2}\right), \quad (2)$$

式中: f_0 代表滤波器的中心频率; β/f_0 为常数,保证了滤波器的形状不改变。Log-Gabor 滤波器除了有良好的多通道和多分辨率特性,并且没有直流分量,是严格的带通滤波器,在高频处弥补了 Gabor 滤波器对高频分量表达不足的缺点。

笔者采用一维 Log-Gabor 滤波器对归一化的虹膜图像进行滤波,将归一化图像中的每一行看作一维的信号,分别于一维 Log-Gabor 滤波器进行卷积,即先将归一化图像进行傅里叶变换,与一维 Log-Gabor 滤波器的频谱表达式在频域相乘,然后再将相乘结果进行傅里叶反变换^[7]:

$$h_{[Re, Im]} = \text{sgn}_{[Re, Im]} \text{ifft}(G(f) \cdot \text{fft}(I(n))). \quad (3)$$

根据 Daugman 等人提出的虹膜识别算法^[1],对滤波结果进行相位编码,得到虹膜纹理二值化编码,编码规则为:

$$h_{Re} = \begin{cases} 1, & \text{Reifft}(G(f) \cdot \text{fft}(I(n))) \geq 0 \\ 0, & \text{Reifft}(G(f) \cdot \text{fft}(I(n))) < 0 \end{cases}; \quad (4)$$

$$h_{Im} = \begin{cases} 1, & \text{Imifft}(G(f) \cdot \text{fft}(I(n))) \geq 0 \\ 0, & \text{Imifft}(G(f) \cdot \text{fft}(I(n))) < 0 \end{cases}.$$

笔者计算待识别虹膜纹理二值化编码与虹膜特征数据库中二值化编码的 Hamming 距离,即统计两个二值化编码对应位不同的个数占总编码位数的比例,以此作为两个二值化编码的距离,即两个虹膜图像的相似度度量。设 A 、 B 是两个二值化特征模板,模式距离为:

$$H = \frac{1}{N} \sum_{i=1}^n A_i \oplus B_i, \quad (5)$$

式中: \oplus 为异或操作符; N 表示特征维数; H 表示模式距离,两个二值化编码对应位编码不同所占比例越大, H 值越大,相似度越小, H 值越小,则相似度越大,为消除虹膜采集的旋转、偏移,在计算时进行左右各 3 位的循环移位,取最小 H 值作为待识别虹膜与特征库虹膜的相似度。

3.2 Haar 小波和加权欧式距离

通过小波变换提取图像纹理特征,是图像识别中常用的分析方法。Haar 小波是一种最基本的正交小波、具有紧凑性、构造简单、计算方便等特点,笔者采用 Haar 小波对归一化的虹膜进行 3 层分解,从高尺度向低尺度进行,得到如图 8 所示的小波分解结果,分解出来的子图像称为小波分解通道:LL、HL、LH、HH。每个通道子图像是原始图像在不同空间频率和方向选择下的信息。根据原始图像的纹理在不同频率和方向下具有不同的能量特性,采用以下公式计算一系列小波通道的能量和方差作为虹膜图像的特征值。

$$f = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N x(i, j), \quad (6)$$

$$\delta = \sqrt{\frac{\sum_{i=1}^M \sum_{j=1}^N [x(i, j) - EC_n]^2}{MN - 1}}, \quad (7)$$

式中: MN 为小波通道的尺寸; i, j 分别代表了通道中行和列; $x(i, j)$ 是该通道小波分解系数^[5]。

笔者舍弃包含图像大部分噪声的 HH 通道,对如图 8 所示 LL3、HL3、LH3、HL2、LH2、HL1 和 LH1 通道进行特征值计算,从 7 个小波通道中提取到 14 个特征值作为虹膜特征向量。

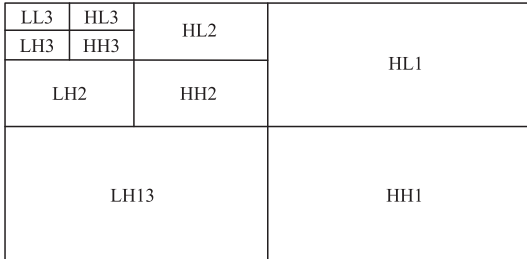


图 8 小波分解示意图

Fig.8 The sketch map of wavelet transform

欧几里得度量又称作欧式距离,是适用于多维向量的距离定义,公式为:

$$W(k) = \sum_{i=1}^N \frac{(f_i - f_i^{(k)})^2}{(\delta_i^{(k)})^2}, \tag{8}$$

式中: f_i 为未知样本的第 i 个特征; $f_i^{(k)}$ 、 $\delta_i^{(k)}$ 表示第 k 类虹膜的第 i 个特征的均值和方差; N 表示特征总数.把方差的倒数作为权值,使用加权欧氏距离来计算虹膜相似度^[9].

3.3 SVM 特征相似度融合及分类

支持向量机是 Cortes 和 Vapnik 于 20 世纪初首先提出的,它是一种二类分类模型,学习策略为特征空间上间隔最大化,最终可以转化为一个凸二次规划问题,在解决小样本非线性及高维模式识别中表现出许多特有的优势,其分类的查全率和查准率几乎超过了现有的所有方法,具有很好的泛化能力及其他机器学习方法不可比拟的优势^[10].虹膜识别系统的特征库作为小样本系统,十分适合采用 SVM 进行模式匹配.

笔者使用多项式核函数实现认证模式下虹膜的两种特征相似度融合识别,将采用 Log-Gabor 滤波器处理和 Hamming 距离计算得到的相似度 H_i ,采用 Haar 小波处理和欧式加权距离计算得到的相似度 W_i ,作为 SVM 的二维输入样本,输入 SVM 分类器进行匹配分类^[10].

4 实验结果及分析

笔者采用中科院自动化研究所提供的 CASIA V1.0 虹膜数据库图像(108 类虹膜,每类虹膜有 7 幅,共 756 幅虹膜图像)作为本算法实现的实验样本,通过 MATLAB7.0 编程实现文中虹膜识别算法.以正确识别率(correct recognition rate, CRR)、识别时间作为衡量算法性能的指标.笔者将数据库中每个类的虹膜图像训练样本数由 1 个开始,增加到 6 个,剩下一个虹膜图像作为待识别实验样本,识别结果如图 9 所示,用于训练的样本数量越多,正确识别率越高,当训练样本数为 6 个

时,已经取得较好的识别效果.

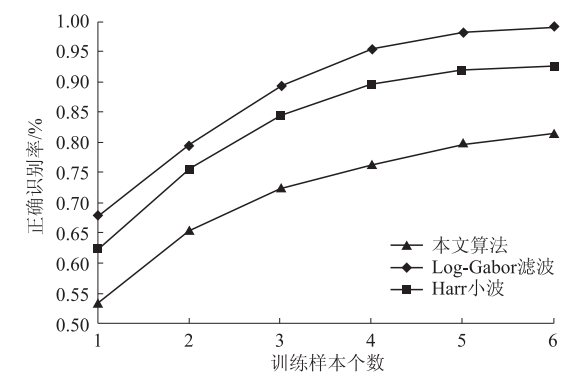


图 9 算法比较折线图

Fig.9 The broken line graph of algorithm comparison

实验结果如表 1 所示,可以看出笔者提出的多特征提取融合匹配的方法虽然由于算法复杂使识别时间有所增加,但相较于单独特征提取匹配的方法,正确识别率有了大幅度的提高.

表 1 不同识别方法时的识别率和识别时间

Tab.1 The recognition rate and time of different methods

算法	正确识别率/%	识别时间/ms
Log-Gabor 滤波	92.59	69.7
Haar 小波	81.48	48.3
本文算法	99.17	82.4

与经典算法的对比结果如表 2 所示,可以看出笔者提出的多特征提取融合匹配的算法虽然在正确识别率上与 Daugman 算法相比有所不足,但识别时间大大缩短,系统识别速度更快,与 Boles 算法相比,识别效果更好.

表 2 与经典算法的对比

Tab.2 The comparison of classical methods

算法	识别率/%	识别时间/ms
Daugman 算法	99.31	365.3
Boles 算法	93.56	150.7
本文算法	99.17	82.4

5 结论

笔者分别采用 Log-Gabor 滤波器和 Haar 小波特征提取归一化虹膜图像的纹理特征,使用 Hamming 距离和欧式加权距离计算相似度,并运用生物识别技术中多特征融合的思想,通过 SVM 对相似度进行融合匹配分类.实验结果证明了该算法的有效性,并能取得较高的识别率.

参考文献:

[1] DAUGMAN J. How iris recognition works[J]. IEEE transactions on circuits and systems for video technology, 2004, 14(1):21-30.

[2] WILDES R P. Iris recognition: An emerging biometric technology [J]. Proceedings of the IEEE, 1999, 85(9):1348-1363.

[3] BOLES W, BOASHASH B. A human identification technique using images of the iris and wavelet transform[J]. IEEE transactions on signal processing, 1998, 46(4):1185-1188.

[4] LIM S, LEE K, BYEON O, et al. Efficient iris recognition through improvement of feature vector and classifier[J]. Journal of electronics and telecommunication research institute, 2001, 23(2):61-70.

[5] 周治平,吴会军. 结合小波变换和 Log-Gabor 滤波的虹膜识别方法[J]. 计算机应用, 2009, 29(8):2095-2097.

[6] 许威. 虹膜识别算法的改进与实现[D]. 合肥:安徽理工大学电气与信息工程学院, 2013.

[7] 陈文静, 金华. 基于人脸与虹膜的多生物特征融合识别算法研究[J]. 中国人民公安大学学报(自然科学版), 2015, 83(1):77-79.

[8] 张震,张英杰. 基于支持向量机与 Hamming 距离的虹膜识别方法[J]. 郑州大学学报(工学版), 2015, 36(3):25-29.

[9] 王蕴红,朱勇,谭铁牛. 基于虹膜识别的身份鉴别[J]. 自动化学报, 2002, 28(1):1-10.

[10] 唐银凤,黄志明,黄荣娟,等. 基于多特征提取和 SVM 分类器的纹理图像分类[J]. 计算机应用与软件, 2011, 28(6):22-25+46.

An Iris Recognition Algorithm of Multiple Features Extraction and Fusion

ZHANG Zhen, LIU Bo, LI Long

(School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: In this paper, an iris recognition method of multiple features extraction and fusion was proposed, which was different from the classical iris recognition methods that focused on a single texture feature. Firstly, the collected iris images were preprocessed. Secondly, features of the iris were extracted by using Log-Gabor filter and Haar wavelet respectively, and two similarity values were calculated by adopting Hamming Distance and Weighted Euclidean Distance. Finally, the Support Vector Machine was used to fuse the values of similarity and classify. The experimental results verified that the proposed method is effective and a higher recognition rate was achieved.

Key words: iris recognition; multiple feature extraction; Log-Gabor filter; Haar wavelet; feature fusion