

基于信息融合的人脸人耳多模态识别技术研究

于改红

(北京航空航天大学 计算机学院, 北京市, 100191)

摘要:单一模式生物特征识别系统存在一些固有的局限性,有时难以满足实际应用的需求,因此本文研究了基于人脸和人耳信息融合的多模态生物特征识别方法。人脸人耳融合识别,融合策略包括数据层融合、特征层融合和决策层融合三种。研究内容的三个关键部分为图像库预处理、特征提取和不同层次的信息融合识别。图像预处理阶段需要经过图像灰度化、图像与修复、直方图均衡化、尺寸归一化等五个步骤。采用标准的LBP描述符提取特征,特征融合采取简单串联算法形成新的特征。分类决策时采用了三种模板匹配算法,分别是LBP模板匹配,PCA模板匹配,及欧氏距离模板匹配。利用Matlab开发工具,设计实现了基于信息融合的人耳人脸多模态识别系统原型。识别结果表明基于人脸人耳信息融合的多模态识别的识别率优于单体的人耳或人脸识别,并扩大了识别范围。这说明融合多种生物特征的多模态识别可以提高身份认证的准确率。

关键词: 人耳识别 人脸识别 模板匹配 信息融合 多模态识别

Multimodal Recognition Technique of Face and Ear Based on Information Fusion

Abstract: Unimodal biometric systems have to contend with a variety of problems and sometimes cannot satisfy application requirements. In this paper, a novel method of multimodal recognition using face and ear based on information fusion was proposed. The multimodal recognition was studied on the data level fusion, feature level fusion and decision level fusion. The contents of this thesis refer to three main stages of multimodal recognition system using face and ear based on information fusion. These are image preprocessing, feature extraction and classifier designing based on information fusion of different levels. The preprocessing procedure includes five steps mainly, converting an RGB ear image to gray one, denoising and restoration, histogram equalization and size normalization. We attempt to adopt the normal LBP descriptor to extract image feature, and use simple cascade algorithm to make the new feature of feature level. We adopt three template matching methods contains LBP, PCA and Euclidean templet. We designed and performed the recognition system in the environment of Matlab 7.0. The experimental results based on the given ear and face database of different angles and single ear and face database cut manually show that the multimodal recognition outperforms the unimodal biometric recognition and the scope for identifying is expanded. This work shows that multi-biometric system can increase the accuracy of overall system recognition, and provides an effective approach of non-intrusive biometric recognition.

Key words: Ear recognition Face recognition Template matching Information fusion Multimodal recognition

引言

随着国民经济和社会信息化的飞速发展,金融机构、政府机关、企业及其个人之间通过互联网日益紧密地联系在一起,一方面为信息的共享提供了条件,另一方面也为心怀叵测的人试图非法获取他人信息提供了机会。因此,如何自动、高效和准确地鉴别人的身份是信息安全领域的重要问题。为克服传统身份鉴别容易遗忘、容易丢失、容易被破译和仿制等缺点,

人们提出了生物特征识别方式。所谓生物特征识别技术是指通过计算机利用人体所固有的生理特征或行为特征来进行个人身份鉴别。与传统的个别鉴别技术相比，生物特征识别技术是一种更安全、更方便的身份鉴别技术。由于没有任何一个单一的生物特征识别系统足够精确和可靠，因此多特征识别系统的出现是一个可选的策略^[1]。

人脸识别具有友好性、非打扰性的优点，一直是研究和应用热点。人耳不受表情、化妆的影响，也不易损伤，具有一定的研究价值和应用前景。又由于人耳和人脸这两种特征体在生理位置有特殊关系，所以就提出了将两者进行融合识别，是对生物特征识别领域的一种拓展。基于多生物特征融合来进行身份鉴别有好多优点，多个生物特征的运用可以提高整个身份鉴别的准确性，伪造多个生物特征显然比伪造单个生物特征更困难使之可靠性强，融合多种生物特征可以扩大单个生物特征的适用范围使之适用性强。而多生物特征是靠信息融合来实现的。信息融合可解决来自多元数据与信息的相关、相关和组合问题，以实现了对多研究实体的精确定位和特性估计^[2]。

综上所述，基于信息融合多模态有很多优势，并且研究人脸人耳融合识别技术有很好的前景，能为信息安全的进步做贡献。

1 多模态识别研究现状

1.1 选题的背景

国际安全形势的日趋复杂化对身份验证技术的需求与日俱增，基于信息融合的多模态的生物特征识别技术也越来越重要。人耳识别技术是20世纪90年代末开始兴起的一种生物特征识别技术。人耳具有独特的生理特征和观测角度的优势，使人耳识别技术具有相当的理论研究价值和实际应用前景。但人耳识别很局限，很容易受到外界因素如头发遮挡，戴耳环的影响。人脸识别虽然有悠久的历史，一直以来人们就将对人脸的辨识作为判断人的身份的一种方法，进行了许多探索，但人脸识别容易受表情变化、年龄增长、遮挡物及光照、成像角度、成像距离的影响。目前的人脸识别系统，在待识别图像和训练图像姿态相近的情况下识别率很高，但是当姿态变化较大时，识别率急剧下降，因此，姿态的影响成为了人脸识别真正走向实用的瓶颈。

鉴于上述人脸人耳在识别中的问题，通过改善识别过程中的某一环节来提高系统的性能是非常困难的。由于人耳、人脸特征都是以图像信息给出的，因此样本采集和特征提取都存在相似性。另外，人耳、人脸位置具有关联性，识别时，只要用摄像头从不同角度对准人的面部即可，操作容易，甚至还可能在公共场所实现远距离的非打扰识别，因此基于信息融合的人脸人耳生物特征识别就越来也成为研究的主题。与传统身份验证方法相比，生物特征识别技术具有稳定、便捷、不易伪造等优点，得到了广泛的研究和应用。

1.2 人脸人耳多模态识别的研究现状

2005年燕山大学信息科学与工程学院的陈宝峰、罗白玲、王元卓、张大鹏在《计算机科学》中发表了题目为“基于人脸和人耳的组合生物特征识别方法”论文，文章提出了人脸人耳等单一生物特征识别的局限性和人脸人耳特征组合的方法及其可行性，并根据人头部的器官位置关系，对特征组合算法进行了描述，采用人脸识别的自动校准算法进行主元的提取，用神经网络方法完成特征降维处理，用子网组合式神经网络识别方法处理新样本加入时的问

题。这在一定程度上较好解决了对人脸图像进行特征提取时对位置要求较为严格的问题，同时也在一定程度上解决了对单一人耳识别时的定位问题，使特征的获取和识别具有较好的自适应性和较强的鲁棒性[5]。

2007年北京科技大学的徐晓娜、穆志纯在《计算机应用研究》上发表题为“基于CAA的人耳和侧面人脸特征融合的身份识别”论文，文章也是根据人脸人耳的特殊生理位置关系，从非打扰识别角度出发，提出仅采集侧面人脸图像，利用典型相关分析(CAA)的思想提取人耳和侧面人脸的关联特征，进行人耳和侧面人脸在特征层的融合。实验结果表明，此方法与单一的人耳或侧面人脸特征识别比较，识别率得到提高[6]。

2007年袁立，穆志纯、曾慧在《北京科技大学学报》上发表了题为“基于人脸和人耳的多模态生物特征识别”论文，文章为克服单一模式生物特征识别系统中存在的一些固有局限性提出了基于正面人脸和人耳信息融合的多模态生物特征识别方法。针对USTB人耳图像库和ORL人脸图像库，利用核Fisher鉴别分析方法分别进行了人耳识别、人脸识别和人脸人耳融合识别，融合策略包括图像层融合和特征层融合两种。识别结果表明基于人脸人耳信息融合的多模态识别的识别率优于单体的人耳或人脸识别。这说明融合多种生物特征的多模态识别可以提高身份认证的准确率，也为实现非打扰式识别提供了一种新的途径，满足了实际应用的需求[7]。

2009年北京科技大学信息工程学院的郭文杰、穆志纯在《天津大学学报》发表了题目为“基于特征融合的人脸人耳多生物身份鉴别”论文，文章针对人头旋转时单一人脸识别率较差的问题，基于人脸与人耳位置上的关联性，提出人脸人耳多生物特征级融合的身份鉴别方法以克服姿态带来的影响。首先采用传统独立成分分析(ICA)方法及其变形分别提取出图像的局部和全局特征，然后将这两种互补的特征进行多模态加权串联融合，并采用基于非线性核函数的主元分析法(KPCA)降维。在USTB图像库上的实验表明，两种独立成分特征具有很好的互补性，多生物识别大大优于单一生物识别，且提出的核非线性降维方法进一步改善了识别性能[8]。

综合上述研究现状，根据生物特征识别系统的结构，多模态生物特征识别技术中的信息融合可分为四层，即数据层融合、特征层融合、匹配层融合和决策层融合。从数据层到决策层，实现的难度依次递减，但它们所利用的信息量的大小也依次递减。

1.3 本文主要研究内容

本文主要研究了基于信息融合的人脸人耳识别方法。信息融合包含数据层、特征层和决策层。利用基于LBP纹理特征提取方法分别提取人脸人耳整体特征，人脸人耳分开的特征，利用特征串联再降维的方法形成特征层的新特征。对于识别方法，文中研究了模板匹配识别法，重点研究了距离模式相似度测度下的模板匹配，采用了欧式距离模板匹配，基于PCA的模板匹配及LBP模板匹配。在特征提取方面，主要介绍了基于局部二值模式(LBP)提取特征，

以及加权串联然后降维提取融合特征的方法。

2 对多模态生物特征识别的认识

2.1 单生物特征识别

较早的生物特征识别都只是针对一个单独的特征。研究表明，能够用来进行身份确认的生物特征应该满足以下几点要求：

- (1) 普遍性：每个人都应该拥有该特征。
- (2) 唯一性：两个人之间不存在相同的该特征。
- (3) 不变性：该特征不随时间变化。
- (4) 可采集性：该特征可以定量采集。

在实际应用中，往往很难找到能够同时满足以上所有条件的生物特征。并且生物特征会受到各种因素的影响和制约。例如，有 5%左右的人群的指纹很难提取出有效特征，人脸会因为化妆、年龄增大等有所改变使识别率降低，虹膜会因为白内障等疾病有所改变，声音会因健康条件而改变[3]。

2.2 多生物特征识别

由于没有任何一个单一的生物特征识别系统足够精确和可靠，因此多特征识别系统的出现是一个科学的策略。例如声音和人脸可以结合在一起组成一个多特征识别系统，本文将人耳和人脸结合起来也可以。这些概念也可以扩展到那些对同一特征应用不同处理方法的系统。随着需求的增加，多模态生物特征识别（Multi-model Biometrics）的研究和应用逐渐兴起和深入[3]。基于多生物特征融合来进行身份鉴别的优点主要有以下几个方面：

- (1) 准确性：多个生物特征的运用可以提高整个身份鉴别的准确性。
- (2) 可靠性：伪造多个生物特征显然比伪造单个生物特征更困难。
- (3) 适用性：每种生物特征都存在应用的局限性。融合多种生物特征可以扩大单个生物特征的适用范围。

多生物特征融合与信息融合密切相关，本文就是基于信息融合的人脸人耳组合特征识别。下面介绍有关信息融合的知识。

2.3 信息融合的基础概念

2.3.1 信息融合定义

- (1) 以处理过程来定义，信息融合是将多源信息或多个传感器获取的信息进行有目的

的组合，将多种不同的数据组合成一个一致的，精确的和智能整体的处理过程。

(2) 以技术或理论方法和工具来定义，信息融合是协同利用多源信息（传感器、数据库、人为获取的信息）进行决策和行动的理论和工具，旨在比仅利用单信息源或非协同利用部分多源信息获得更精确和更稳定的性能[9]。

2.3.2 信息融合步骤

融合处理涉及的对象可能是：多源信息和多传感器直接提供的数据，需要决策的静态或相对静态的对象，本文的系统要融合的对象是人脸人耳的静态特征，静态图像数据。针对一个确定的具体决策任务，设计其信息融合处理一般流程如图 2-1 所示。

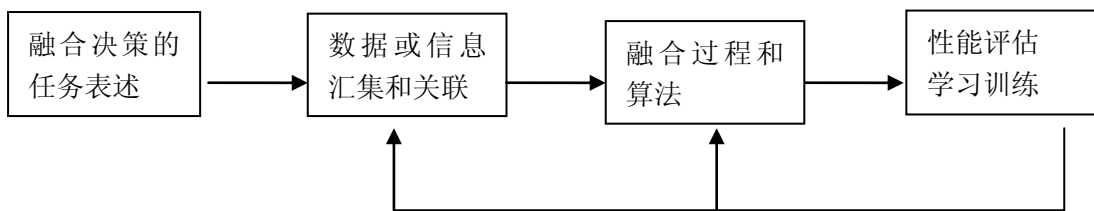


图 2-1 信息融合的设计流程图

2.3.3 融合层次分类

信息融合文献中经常提到信息融合层次的概念。因为融合单元的输出输入，都可能采取数据、特征、决策等不同的形式，这样就形成了多种不同形式的输入-输出成对关系。最好理解的三对：数据-数据对、特征-特征对和决策-决策对，于是就形成了数据层融合、特征层融合和决策层融合等划分。

2.2.3.1 数据层融合

数据层融合一般是指基于相同维数多源数据的融合。数据层融合分类的计算对象是由多源数据融合形成的一个新数据。可以采用全数据范围对应时空位置的直接融合，也可以采用互补式插入的组合融合。本文采用前者。数据层融合就是利用人脸人耳的位置将人脸人耳作为时空整体来进行特征提取和识别。

2.2.3.2 特征层融合

特征层融合的直接概念是将由多源数据中分别提取的特征汇总在一起形成新的特征集，以后根据新的特征在进行决策。关键是特征融合算法的选取，例如最简单多模态加权串联融合、典型相关分析（CAA）的思想提取人耳和侧面人脸的关联特征以及 ICA 实现特征融合等，通过多源数据分析与处理产生新的更具有表征能力的超级特征，提高了识别率更具有说服

力。本文中的特征层融合是将人脸的特征和人耳的特征用加权串联的算法进行融合得到新的特征，然后再用 PCA 降维进行识别。

2.2.3.3 决策层融合

决策层融合分类在很大程度上是指多类分类器融合分类，每个分类器的输出都具有一定层次的局部决策。利用多分类器进行融合分类，是根据两个应用情况提出的。一是实际应用中常常使用不同类型的传感器，直接在各类传感器数据分类基础上进一步融合分类十分自然的，这样可以克服单传感器数据性能限制的弱点。而是从实际计算处理的需求考虑，利用单一分类器处理所有信息源提供信息的分类，不但给计算、通信等带来负担，而且需要对任意初始条件或参数做适合的承诺，采用多分类器组合可以减轻这方面的压力，提高融合分类效率。本文采用了最简单的分类决策方法，就是分别对人耳和人脸进行识别，也可用不同的分类方法对他们分类，对分类结果进行总决策。

3 各融合层次的特征提取原理算法

本文中提取特征环节都是采用 LBP 算法。形成了数据层特征人脸人耳库(data_whole)、人耳库(data_ear)、人脸库(data_face)三个库。特征层融合特征是将人脸人耳库作为多源信息来源进行融合，形成融合特征库(data_link)。下面分别介绍 LBP 提取特征和加权串联然后 PCA 降维的融合特征算法。

3.1 数据层融合特征提取

本文数据层属于空间图像融合，直接使用 LBP 提取人耳和人脸合在一起的特征。局部二值模式(LBP)作为图像局部对比度的互补算法。LBP 最初定义于像素的 8 邻域中，以中心像素的灰度值为阈值，将周围 8 个像素的值与其比较，如果周围的像素值小于中心像素的灰度值，该像素位置就被标记为 0，否则标记为 1；将阈值比较后的值(即 0 或者 1)分别与对应位置像素的权值相乘，8 个乘积的和即为该邻域的 LBP 值(如图 3-6)。由定义可以看出，灰度值的单调变化不会引起 LBP 值的变化，因此，它与局部对比度互相独立并互相补充，图中显示了局部对比度(Contrast, C)的计算方法，它实际上是两组数据平均值的差，即大于等于中心像素灰度值的所有灰度取平均值，减去小于中心像素灰度值的所有像素值的平均值[13]。

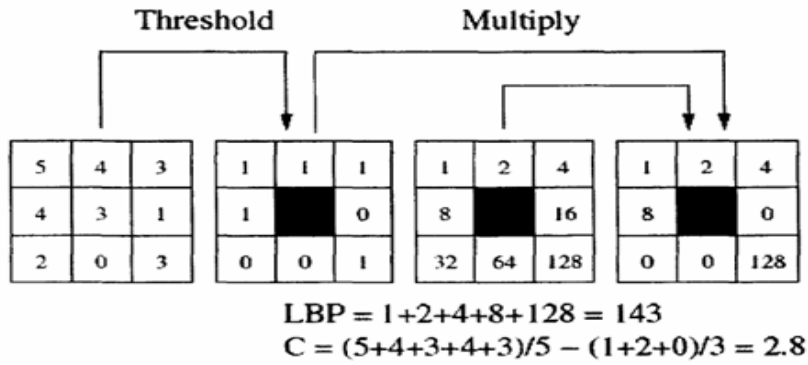


图 3-6 原始 LBP 值与对比度的定义

以下我们从对纹理的常规定义中引出 LBP 纹理分析方法。

在某一灰度图像中，定义一个半径为 $R(R>0)$ 的圆形邻域， $P(P>0)$ 个邻域像素均匀分布在圆周上(图 3-7)。设该邻域中心像元的纹理为 T ，则 T 可以用该邻域中 $P+1$ 个像素的函数来定义，即：

$$T = t(g_c, g_0, \dots, g_{p-1})$$

(3-1)

其中， g_c 为该邻域中心像素的灰度值， $g_i (i=0, \dots, P-1)$ 是 P 个邻域像素的灰度值。

在数字图像中，设邻域中心像素的坐标为 (x_c, y_c) ，则其邻域像素的坐标 (x_i, y_i) 可以表示为：

$$\begin{cases} x_i = x_c + R \cos(2\pi p/p) \\ y_i = y_c + R \sin(2\pi p/p) \end{cases} \quad (3-2)$$

图 3-5 显示几种不同 P, R 值对应的圆环形邻域，图中没有落在像素中心的邻域的灰度值可通过双线性内插得出。

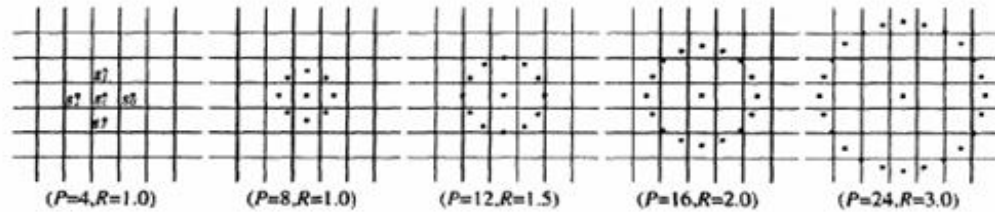


图 3-7 几种不同 P, R 值对应的圆形邻域

由于随着距离的增大，像素之间的相关性逐渐减小，因此，在较小的邻域中即可获得绝大部分纹理信息。在不丢失信息的前提下，如果将邻域像素的灰度值分别减去邻域中心的灰度值，局部纹理特征则可以表示为：

$$T = t(g_c, g_0 - g_c, \dots, g_{p-1} - g_c)$$

(3-3)

假设各个差值与 g_c 相互独立，则式(3-3)可以分解为：

$$T \approx t(g_c)t(g_o - g_c, \dots, g_{p-1} - g_c)$$

(3-4)

鉴于 $t(g_c)$ 代表图像的亮度值，且与图像局部纹理特征无关，式(3-4)描述的纹理特征可以直接表示为差值的函数：

$$T \approx t(g_o - g_c, \dots, g_{p-1} - g_c)$$

(3-5)

这一 P 维差值函数记录了邻域中每个像素的纹理模式(Pattern)。例如，对于相对平滑，变化较小的区域，这些差值接近 0；在某点处，所有的差值相对较大；在边缘处，某方向的差值会比其他方向大。

虽然公式(3-5)中定义的纹理不受灰度值变化的影响，即邻域中所有 $P+1$ 像素同时加上或者减去某个值，其表征的纹理不变，但是当所有像素的值同时放大或者缩小相同倍数后，其纹理特征发生变化(be affected by scaling)。为了使定义的纹理不受灰度值单调变化的影响，只考虑差值的符号：

$$T \approx t(s(g_o - g_c), \dots, s(g_{p-1} - g_c))$$

(3-6)

其中

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

函数 $s(g_i - g_c)$ 乘以因子 2^i ，则得到唯一表征局部纹理特征的 LBP 值：

$$LBP_{P,R} = \sum_{i=0}^{P-1} s(g_i - g_c)2^i$$

(3-7)

可以看出， $LBP_{P,R}$ 算子与其最初的定义原理上是相似的：描述局部纹理模式的二值编码都是通过邻域像素与中心像素的灰度值比较得到。两者的主要区别有两点，其一，原始 LBP 定义于方形邻域上，而修改后的 LBP 定义于圆形邻域中，从而更容易定义具有旋转不变性的纹理描述算子；其二，修改后的算子 $LBP_{P,R}$ 中，其邻域中对角线上的像素值是通过线性内插得到的[13]。

3.2 特征层特征融合算法原理

特征层融合特征的提取包含两个步骤，首先对人耳特征和人脸特征进行加权融合，然后利用PCA进行降维，形成识别特征[14]。

3.2.1 加权融合算法

将应用LBP提取的人耳和人脸两种特征串联构成单模态融合向量，由于不同的生物模式的特征有不同的鉴别能力，因此将不同模式的融合特征向量赋予不同的权值，并串联形成一个更高维的融合向量。设人耳特征为 $feature_ear$ ，人脸特征为 $feature_face$ ，融合特征为 $feature_fusion$ ，则有

$$x_{fusion} = \begin{pmatrix} \alpha x_{ear} \\ \beta x_{face} \end{pmatrix} \quad (3-8)$$

式中：使用约束条件 $\alpha + \beta = 1$ 。

在实验过程中权值可以通过训练得到，具体是：将实验用的图像库分成权值训练子库和识别子库两大部分。在权值训练子库中，将 α 以 0.1 为间隔从 0.1 调整到 0.9，在每个对应的 α 和 β 下进行特征的串联，并完成整个识别过程，其中最高识别率对应的一对权值就是所需的最好权值。然后将其直接用于识别子库完成特征的加权串联。可以看出这种权值获取方法是最优的，可以适应不同特征的串联融合问题，而且结果可靠稳定，然而这是以牺牲一部分数据库规模为代价的[14]。

3.2.2 PCA 降维

PCA 是将分散在一组变量上的信息，集中到某几个综合指标(主成分)上的一种探索性统计分析方法。它利用降维的思想，将多个变量化为少数几个互不相关的主成分，从而描述数据集的内部结构。

先假定只有二维，即只有两个变量，它们由横坐标和纵坐标所代表。因此每个观测值都有相应于这两个坐标轴的两个坐标值。如果这些数据形成一个椭圆形状的点阵（这在变量的二维正态的假定下是可能的），那么这个椭圆有一个长轴和一个短轴。在短轴方向上，数据变化很少。在极端的情况，短轴如果退化成一点，那只有在长轴的方向才能够解释这些点的变化了。这样，由二维到一维的降维就自然完成了。

当坐标轴和椭圆的长短轴平行，那么代表长轴的变量就描述了数据的主要变化，而代表短轴的变量就描述了数据的次要变化。但是，坐标轴通常并不和椭圆的长短轴平行。因此，需要寻找椭圆的长短轴，并进行变换，使得新变量和椭圆的长短轴平行。如果长轴变量代表了数据包含的大部分信息，就用该变量代替原先的两个变量（舍去次要的一维），降维就完成了。椭圆（球）的长短轴相差得越大，降维也越有道理。

3.3 决策层融合算法

对于仅两个分类器投票的问题，因为对待识别样本，若两个分类器将其分入不同类别，则会出现票数相等均为1的问题，系统无法决策。对普通投票法进行改进，以避免上述问题。

首先，不同的分类器相当于决策过程中不同的专家，应有不同的说话份量，即权值。本文采用最简单的方法，即人耳和人脸这两个分类权值是一样的，并且只投票一次，若相同则结果确定，若结果不同则显示结果不能确定。

4 模板匹配

4.1 理论基础

本文在数据层、特征层、决策层识别时都采用了模板匹配的方法。将待识别图像的特征向量与已经训练和建立在模板中的各样本的特征向量或者一类模板的特征评价价值进行对比和计算，得到最相似模板的样本，进行分类的方法就是模板匹配分类法。在图像识别中，最简单和易于理解的识别方法就是模板匹配法。模板匹配分类法把未知图像和一个标准图像对比，看它们是否相同或者计算其相似度。模板匹配分类器将训练样本集合的每个样本都作为标准模板，用待测样本与每个模板进行比较，找出最相似、最近邻的标准模板，将标准模板中的最近类别作为识别结果的类别[2]。

例如 A 类有 20 个训练样本，因此有 20 个模板，B 类有 8 个训练样本，就有 8 个模板。任何一个待测样本在进行分类时都要与这 28 个模板比较其相似度，或者每一个待测样本的特征与各个模板的特征值的平均值相比较，以找出最相似的模板，如果该模板的特征相似 B 类中的一个或平均值，就确定为 B 类，否则就为 A 类。因此，从原理上来讲，模板匹配是最简单的分类方法。但模板法有一个明显的缺点就是计算量大，存储量大，要存储的模板很多，每个待测样本要对模板库中的每个模板计算一次相似度。本文中采用了三种距离测度分类法，分别是近距离模板匹配法、基于 PCA 的模板匹配法，基于 LBP 描述符的模板匹配。在这重点介绍 LBP 模板匹配。

4.2 三种匹配法实现步骤

4.2.1 欧式距离匹配法

- (1) 待测样本 X 与训练集里每个样品 X_i 的距离为 $d(X, X_i) = |X - X_i|^2$ 。
- (2) 循环计算待测样品和训练集中各已知样本之间的距离，找出距离待测样本最近的已知样品，该已知样品的类别就是待测样本的类别[17]。

4.2.2 基于 PCA 的模板匹配法

在使用模板匹配法之前，先对特征进行主成分分析，按照一定贡献值，提取前 m 个主分量，用较低维数的特征来进行分类。实现步骤如下：

- (1) 选取各类全体样本组成的矩阵 $X_n \times N$ ，待测样品为 $X_n \times 1$ 。

- (2) 计算 $X_{n \times N}$ 的协方差矩阵 $S_{n \times n}$ 。
- (3) 计算 $S_{n \times n}$ 的特征值 $L_1 \geq L_2 \geq \dots \geq L_n$ 和特征向量 $C_{n \times n}$ 。
- (4) 根据一定的贡献率，选取 $C_{n \times n}$ 的前 m 列，构成 $C_{n \times m}$ 。
- (5) 计算样本库样本主成分 $X_{m \times N} = C$ 和样本主成分 $X_{m \times 1}$ 。
- (6) 采用模板匹配进行多类别分类。

4.2.3 基于 LBP 描述符的模板匹配法

LBP 最后会把一张图像以一序列空间直方图来表示。两张图像的相似度是空间直方图序列相似度的总和，这个过程的公式如下：

假设一张耳朵图像样本 I 最后表示为 m 个空间直方图，定义如下：

$$F = [H_1 \ H_2 \ \dots \ H_m] \quad (3-9)$$

其中 H_i 是第 i 个空间直方图，然后两副人耳图像样本 I 和 I' 的相似度按如下计算：

$$S(I, I') = S(F, F') = \frac{1}{m} \sum_{i=1}^m S(H_i, H'_i) \quad (3-10)$$

其中 $S(H_i, H'_i)$ 表示直方图的相似度，下面介绍5种通用的直方图匹配方法

一个直方图有 L 个柱条，第 j 个柱条的频率是 $h(j)$ 。

直方图交叉被定义如下：

$$S_{Hi}(H, H') = \sum_{j=1}^L \min(h(j), h'(j)) \quad (3-11)$$

4.3 模板匹配实现过程

距离测度在程序中实现如下：

```

M=(DataClass(j,:,i)-H1).^2;
n=DataClass(j,:,i)+H1;
p=m/n;
temp=temp+p;

```

其中 $\text{DataClass}(j,:,i)$ 代表特征库第 i 个人的第 j 个特征值， $H1$ 代表待识别样本的特征值， m 代表上面两值的差然后再平方， n 代表上面两值的和， p 为 m 和 n 的商， temp 代表特征库中一个人的所有人耳特征值与 $H1$ 计算所得 p 的和^[17]。

```

t=temp/256;
if min(1)>t
min(1)=t;
min(2)=i;

```

t 代表 temp 与特征数的商，后面的 3 条语句是求 t 中的最小值，然后把最小 t 时候的 i 值赋给 $\text{min}(2)$ ，即为类别号。

5 层次融合

5.1 数据层融合

数据层融合的实现使用上述提取特征方法，保存为 `data_whole`，然后分别用以上三种匹配方法匹配结果和统计识别率[9]。

5.2 特征层融合

可先分别对人耳和人脸进行识别，然后用上述方法建立融合特征和降维，之后仍用上述匹配方法进行模板匹配识别。

5.3 决策层融合

决策层就是利用上述模板匹配法分别对人脸和人耳识别，对产生的结果进行融合，本文采用简单的融合，即若对于同一个人的人脸和人耳，若识别的结果一样，则说明识别结果有效，否则就是不能确定识别结果。

5.4 实验结果分析

数据层的数据不仅包含人耳还包含人脸，比单独只取人耳或单独取人脸包含的信息多，准确性和可靠性强。但由于图像是自己切割并且不太准确，所以识别率不高。特征层实现了人耳和人脸的数据在特征层面而不仅仅是空间关系层面，因此识别结果有所提高，因此也说明了特征融合识别有着很好的发展前景。决策层增加了识别的正确率，也就是在生活中更不容易产生错误，因为依靠不同的特征或对相同特征应用不同分类器，这样识别更加严谨。

表 1 数据层特征

特征维数	特征值
第 1 维	2.0990
第 2 维	0.1700
第 3 维	0.5558
第 4 维	0.6997
.....
第 256 维	329.62798

表 2 特征层融合特征

特征维数	特征值
第 1 维	0.6418
第 2 维	0.1070
第 3 维	0.1605
第 4 维	0.4279
...	
第 509 维	5.9875
第 510 维	1.6729
第 511 维	9.0643
第 512 维	162.4377

表 3 权值变化时的识别率情况

权值 (x,y)	识别率
(0.1,0.9)	0.05
(0.2,0.8)	0.17
(0.3,0.7)	0.14
(0.4,0.6)	0.11
(0.5,0.5)	0.16
(0.6,0.4)	0.24
(0.7,0.3)	0.33
(0.8,0.2)	0.45
(0.9,0.1)	0.60

6 结论和展望

本课题经过接近半年的努力，完成了部分工作，由于受时间、技术和学识等所限，本文所取得的成果是有局限性的，离实际应用还有较大的距离，需要朝以下几方面发展：

(1)建立更多的样本：由于模式识别很多时候要求有较大数量的样本库，而本实验中的样本数目相当有限，所以效果不佳在所难免，最好是能按自己的意愿去完成数据采集工作，当然，这需要技术和资金上的大力支持。

(2)对图像能够自动定位和识别，本文中还需要将人脸人耳自己分别切割下来，如果数量庞大，这将是件不能忽视的工作量，当然目前也有在这方面的研究，但可操作性不好。

(3)对数据融合和特征融合时采用更有效的算法的，本文对数据层只是简单空间组合，而特征层也采用比较简单的加权串联。融合度不高，效果不是很明显。

(4)可以在决策层设计一个更有效地决策方法，而不是判断一个图片相同与否。

当然基于信息融合的多模态识别越来越发挥其优越性，其研究的方向也越来越深入，我觉得经过研究者的不断努力，一定会实现更好的识别结果。

参考资料

- [1] Kyong Chang, Kevin W. Bowyer, Sudeep Sarkar. Comparison and combination of ear and face images in appearance-based biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25 (9), 2005
- [2] 雷明. 计算机生物特征识别算法与应用发展. 计算机辅助工程, 2001
- [3] 田捷, 杨鑫. 生物特征识别理论与应用. 清华大学出版社, 2009
- [4] 郭颂, 魏立峰. 基于信息融合的身份识别技术. 仪器仪表学报, 2004
- [5] 陈宝峰, 罗白玲, 王元卓, 张大鹏. 基于人脸和人耳的组合生物特征识别方法. 计算机科学, 2005
- [6] 徐晓娜, 穆志纯. 基于 CCA 的人耳和侧面人脸特征融合的身份识别. 计算机应用研究, 24 (11), 2007
- [7] 袁立, 穆志纯, 曾慧. 基于人脸和人耳的多模态生物特征识别. 北京科技大学学报, 2007
- [8] 敦文杰, 穆志纯. 基于特征融合的人脸人耳多生物身份鉴别. 天津大学学报, 42 (7), 2009
- [9] 王润生. 信息融合. 科学出版社, 2007
- [10] 李南南, 吴清, 曹辉林. MATLAB7 简明教程. 清华大学出版社, 2006
- [11] 姚敏. 数字图像处理. 机械工业出版社, 2006
- [12] 张兆礼, 赵春晖等. 现代图像处理技术及 Matlab 实现. 人民邮电出版社, 2001
- [13] 毛秉毅. 基于自适应 LBP 算法的纹理分析. 计算机工程与应用, 43(2), 2007
- [14] 穆志纯, 郭文杰. 人脸人耳特征融合在身份鉴别中的研究. 计算机科学, 2009
- [15] Li Yuan, Zhi-Chun Mu, Xiao-Na Xu . Multimodal Recognition Based on face and ear. *International Conference on Wavelet Analysis and Pattern Recognition*, 2-4 Nov. 2007
- [16] 杨淑莹. 模式识别与智能计算-matlab 技术实现. 电子工业出版社, 2008
- [17] A. K. Jain, A. Ross and S. Prabhakar. An introduction to biometrics recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14 (1), 2001