

人脸识别中的“误配准灾难”问题研究

山世光¹⁾, 高文^{1),2)}, 唱轶钲²⁾, 曹波¹⁾, 陈熙霖^{1),2)}

¹⁾(中科院计算所-银晨科技面像识别联合实验室, 北京, 100080)

²⁾(哈尔滨工业大学计算机学院, 哈尔滨, 150001)

摘要: 现有的多数人脸识别系统都要依赖于面部特征(比如眼睛中心位置)的严格配准来归一化人脸以便提取人脸描述特征, 但面部特征配准的准确度如何影响人脸识别算法的性能却没有得到足够的重视。本文首次针对这一问题进行了系统的研究, 并提出了一种基于误配准学习的解决方案。为了揭示现有典型识别算法的识别性能对特征配准准确度的敏感程度, 通过对眼睛位置人为加扰, 我们对 Fisherface 算法的识别性能随平移、旋转和尺度改变而变化的情况进行了实验评估, 结果表明: Fisherface 的识别性能随着误配准的增大而急剧下降——本文称这一现象为“误配准灾难”问题。针对该问题, 本文提出了一种基于扰动学习的“误配准灾难”解决方案, 该方法通过在模型训练阶段加入扰动配准偏差来提高判别分析方法对误配准的鲁棒性。在 FERET 人脸图像数据库和 CAS-PEAL-R1 人脸库上的实验表明该方法可以有效的提高识别算法对误配准的鲁棒性。

关键词 人脸识别, 特征配准, 线性判别分析, 误配准灾难, 扰动学习
中图分类号 TP391

“Curse of Mis-alignment” Problem in Face Recognition

SHAN Shi-Guang¹⁾, GAO Wen^{1),2)}, CHANG Yi-Zheng²⁾, CAO Bo¹⁾, CHEN Xi-Lin^{1),2)}

¹⁾(ICT-ISVISION JDL for Face Recognition, ICT, Chinese Academy of Sciences)

²⁾(School of Computer, Harbin Institute of Technology)

Abstract: In this paper, we investigate the rarely concerned curse of mis-alignment problem in face recognition, and propose a novel perturbation learning solution. Mis-alignment problem is firstly empirically investigated through evaluating systematically the Fisherface's sensitivity to mis-alignment on the FERET face database by perturbing the eye coordinates, which reveals that the imprecise localization of the facial landmarks abruptly degrades the Fisherface system. We explicitly define this problem as *curse of mis-alignment* for highlighting its significance. Aiming at this problem, we propose a set of measurement combining the recognition rate with the alignment error distribution to evaluate the overall performance of specific face recognition approach with its robustness against the mis-alignment considered. Finally, a perturbation learning method, named E-Fisherface, is proposed to reinforce the recognizer to model the mis-alignment variations in the training stage. Experimental results on FERET and CAS-PEAL-R1 have impressively indicated the effectiveness of the proposed E-Fisherface to tackle the curse of mis-alignment problem.

Keywords: Face Recognition, Feature Alignment, Linear Discriminant Analysis, Curse of Mis-alignment, Perturbation Learning

收稿日期: 2004-04; 修改稿收到日期2005-01。本课题得到国家自然科学基金项目(60332010), 国家863 计划(2002AA118010), 同时还得到了银晨智能识别科技有限公司的大力资助。山世光, 男, 1975年3月生, 博士, 助理研究员, 研究方向为模式分析, 机器学习, 计算机视觉, 尤其专注于人脸识别问题。高文, 男, 1956 年生, 博士, 教授, 博士生导师, 主要研究领域: 多媒体数据压缩, 图像处理, 计算机视觉, 多模式接口, 人工智能, 虚拟现实等。唱轶钲, 女, 1979年生, 硕士生, 主要研究领域: 人脸识别。曹波, 男, 1977年生, 博士生, 主要研究领域: 模式分析与机器学习。陈熙霖, 男, 1965 年生, 博士, 教授, 博士生导师, 主要研究领域: 计算机视觉, 模式识别, 机器学习, 智能人机交互等

1. 引言

人脸识别在公安、金融等领域具有潜在的重要应用价值，得到了国内外研究人员越来越多的重视。尤其是二十世纪九十年代以来，人脸识别技术更得到了长足的发展，取得了巨大的进步，也出现了若干商业人脸识别系统[1,2,3]。对实用的人脸识别系统而言，人脸检测、面部特征配准和人脸建模与比对是三个不可或缺步骤。针对这三个相对独立的问题，研究人员已经开展了大量的研究。但这三个步骤之间的无缝集成问题却没有得到足够的重视。特别是研究人脸建模与比对问题的研究人员，他们的实验中往往假定面部特征是精确配准的（实验中采用手工标定的双眼位置）。问题是：面部特征配准是否已经非常好地解决了？答案是否定的，尤其是对用户不配合的非限定条件应用系统中，自动提取的面部特征位置往往会有很大的偏差。例如，对一般大小（比如 92×112 像素）的人脸而言，面部特征点的误配准可能高达 5 个像素。另一方面，对研究面部特征提取的研究人员而言，面部特征点的定位精度在 5 个像素以内可能往往都被认为是“正确”的。更为严重的是：在输入人脸图像质量较差时，面部特征点精确位置的定义往往是歧义的，图 1 以眼睛中心为例说明了这一问题，在瞳孔中心与眼睛中心不重合、图像模糊、眼镜反光或者遮挡的情况下，这一问题更为突出。

上述分析表明：一方面需要进一步研究更加精确的面部特征自动定位算法，另一方面，人脸建模与比对方法不应该过分要求面部特征配准的精度，而必须对误配准具有较强的鲁棒性，方能实现稳定可靠的人脸识别。

实际上，在模式识别领域内，对不改变模式类别属性的各种变换（如平移等相似变换）的不变性研究，很早就引起了人们的重视，尤其在数字和字符识别领域，采用切距离（Tangent Distance）度量[4]可以提高识别性能已经成为共识，相关的文献较多，而在人脸识别领域，该问题尚未得到足够的重视，更没有得到系统的量化研究。而且，即使在关于字符识别的文献中，也没有给出识别算法性能随着误配准程度不同而变化的曲线（如本文图 5 所示），也没有给出在考虑误配准鲁棒性的条件下如何评价不同算法的性能，本文首次给出了这些方面的系统研究。我们针对人脸识别中双眼特征的误配准问题进行实验分析和理论研究，揭示了眼睛位置的准确度如何影响识别算法性能，研究了如何提高识别算法对误配准鲁棒性的问题。

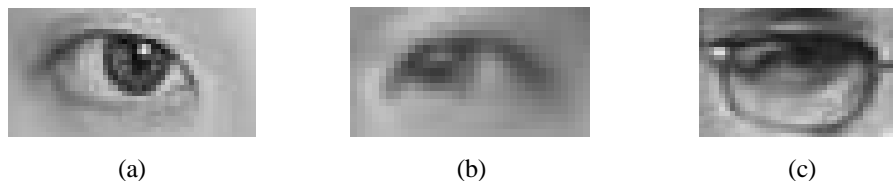


图 1 某些特征点的精确位置本身可能是歧义的，以眼睛中心的定义为例：(a) 眼睛中心与瞳孔中心不重合的情况；(b) 图像模糊的情况；(c) 眼睛遮挡且图像质量很差的情况。眼睛中心被几乎所有的人脸识别系统作为控制点，但其定义往往是主观的而非客观的。

本文的主要贡献包括：(1) 首次系统研究并提出了人脸识别中的“误配准灾难”问题；(2) 提出了一套考虑了误配准分布的系统识别性能评价方法，以弥补识别率变化曲线难以量化比较不同算法性能的不足；(3) 提出了一种增强的判别分析方法 E-Fisherface，通过在模型训练阶段进行扰动学习，有效的降低了原有 Fisherface 方法对误配准的敏感度。上述三方面的工作分别在后续的第 2, 3, 4 节进行了详细介绍，第 5 小节给出了我们提出的增强 Fisherface 在 FERET 人脸库上的实验结果与分析，第 6 节则对全文进行了总结，并讨论了可能的后续工作。

2. “误配准灾难”问题的提出

Fisherface 是最重要的人脸识别方法之一[5]，我们在国际公共人脸库上的实验也表明其在眼睛位置精确定位的条件下具有良好的识别效果。然而，基于 Fisherface 方法的人脸识别应用系统的实际效果却往往会出人意料之差[6]。为了分析这一现象，我们通过记录错误识别案例的实验方法收集了大量的错误样本进行分析，结果表明：大多数被错误识别的样本完全可以通过精确调整眼睛的位置得到正确的识别结果，也就是说，识别性能的下降很大程度上是由于自动定位的眼睛位置不够准确造成的。这使得我们在进一步研究眼睛的精确定位算法的同时，开始认真对待识别算法本身对于眼睛位置定位误差的鲁棒性问题，考虑如何降低识别算法本身对于特

征误配准的敏感性以弥补特征定位不准确带来的系统性能下降。为了系统的分析误配准对识别性能的影响，本文设计了评测实验来显式的揭示该问题的严重性。

2.1 实验设计——在 FERET 人脸库上测试 Fisherface 方法

我们采用的 Fisherface 方法是 Belhumeur[5]提出的方法的一种实现，具体方案是首先采用主成分分析(PCA)进行降维，然后采用线性判别分析(LDA)对降维后的特征进行判别分析，最后采用最近邻方法进行分类识别。为了使得评测结果是统计显著的，我们选择了人脸图像规模在 1000 人以上的人脸库，该人脸库是由美国 DARPA 资助的 FERET 计划构建[2]。方便起见，本文后面的实验只采用了 FERET 中最大的标准测试集 FB，其基本构成如表 1 所示。为了有效的验证算法的性能，本文按照 FERET 标准测试方法[2]的规定将 Fisherface 方法的训练数据集 L (training set) 与测试数据集 T (testing set) 作了严格区分，其中后者包括已知人脸原型图像集合 G (gallery)和探测图像集合 P (probe set)[2]。我们的 Fisherface 实现中，PCA 和 LDA 的变换矩阵 W_{pca} 和 W_{lda} 都是在训练集合 L 上学习得到，然后对 G 和 P 中的所有人脸图像通过降维变换到 LDA 特征空间， P 中任意图像的识别则通过在 LDA 特征空间中与 G 中所有已知样本比对来完成。

表 1 本文所用 FERET 测试标准人脸库集合情况表

数据库	人数	图像数	说明	
训练集 L (Training set)	429	1002	部分人的图像来自参考集 G 和测试集 P	
测试集 T	参考集 G (Gallery)	1196	1196	准正面，均匀光照条件
	探测集 P (Probes) FB	1195	1195	同一天采集，准正面均匀光照，表情不同

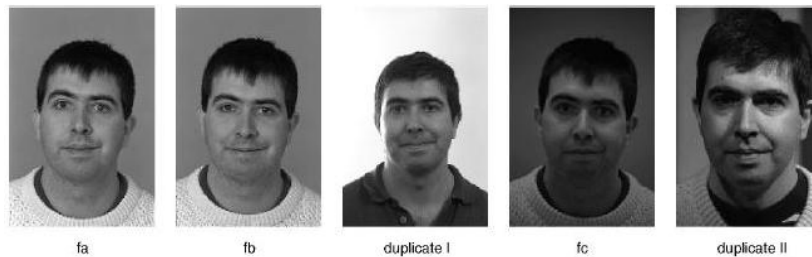


图 2 FERET 人脸图像库示例图像

FERET 人脸库中典型的示例人脸图像如图 2 所示，其中 FB 中人脸图像与 Gallery 图像的主要差别是表情的变化。FERET 提供了所有这些人脸图像的眼睛位置(手工标定)。为了进一步消除背景、头发等非人脸本质因素对识别性能的影响，根据给定的双眼位置，本文采用如下图 3 所示的方式对人脸图像进行几何归一化，其中，图 3(a)中的裁剪参数 h , t , b 分别为 0.64, 0.43 和 1.85。这样，在归一化的 64×64 像素人脸图像中，双眼距离 D 为 28 像素。我们进一步对裁剪出的人脸区域叠加掩模(Mask)以便减小背景和发型变化对人脸识别的不利影响，并进行直方图归一化消除光照影响。最终所有的图像归一化为 64×64 像素。需要指出的是，本文以下讨论所涉及的误配准均指在该尺寸大小图像上的误差情况。



图 3 本文实验中采用的人脸图像归一化方法

为了系统而定量地评估识别方法对于误配准的敏感性，本文采用对探测图像(Probe)眼睛位置人为加入误

配准的实验方法。容易理解,在不考虑 3D 姿态变化导致的误配准的条件下,实际上眼睛位置偏差导致的误配准最终反映为归一化后的人脸图像与真实图像之间的平移 (T_x, T_y)、尺度 (s) 和旋转 (q) 等相似变换参数的变化。为简明起见,本文仅针对这些参数分别进行实验。图 4 给出了这些人为误配准对归一化人脸图像模板的影响情况,可以看出这些误配准在较大程度上影响了人脸图像的外观(Appearance)。



图 4 误配准对归一化人脸图像模板的影响示例

2.2 实验结果——凸显的“误配准灾难”

采用上节所描述的测试方法,对 Fisherface 方法在 FERET-FB 探测集上测试由于误配准而导致一定平移、旋转和尺度变化时其首选识别率的变化情况,结果如图 5 所示。

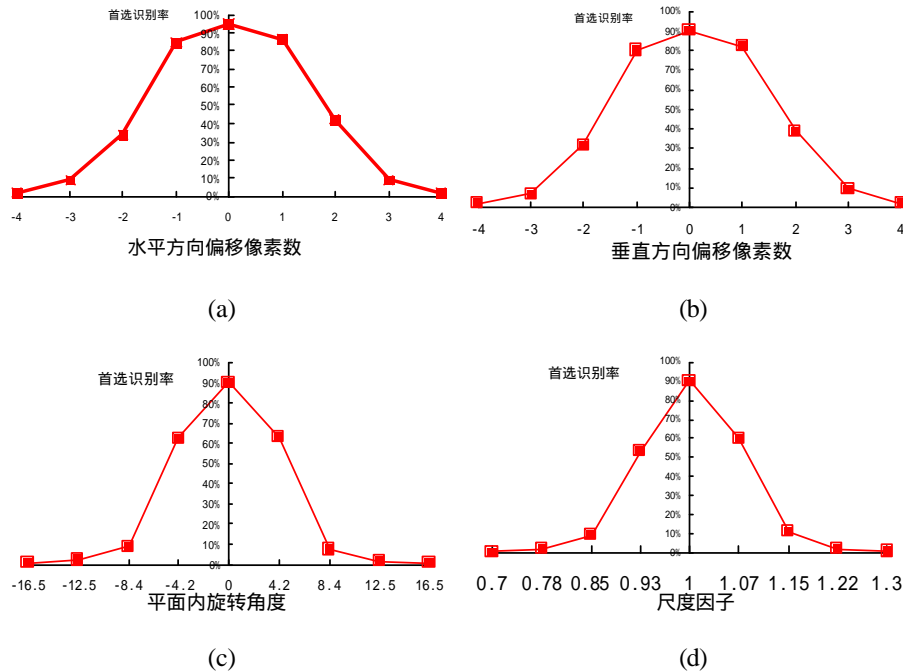


图 5 在 FERET-FB 测试集上, Fisherface 方法的识别率与误配准的关系

其中,图 5(a)是双眼位置在水平方向向相同方向平移若干像素的情况;图 5(b)是双眼位置在垂直方向向相同方向平移若干像素的情况;图 5(c)为双眼位置垂直方向错位从而导致一定角度平面内旋转时识别率的变化情况;图 5(d)为双眼位置在水平方向向相反方向平移导致尺度变化时的情况。不难看出, Fisherface 的识别率会随着误配准的增大而急剧下降,即使在只有一个像素平移偏差的情况下,识别性能也下降了 10 个百分点;而平面内旋转 4.2 度(左、右眼在垂直方向朝相反方向各移动 1 个像素)时,识别性能则下降了 20 多个百分点;尺度发生 0.07 的变化(左、右眼在水平方向朝相反方向各移动 1 个像素)时,首选识别率陡然下降了 30 个百分点。显然,在发生如此小的误配准时,如此剧烈的性能下降是实用的人脸识别系统根本不能接受的,尤其是在全自动的人脸识别系统中,面部特征定位算法往往很难达到百分之百的准确度,偏差一两个像素往往是非常正常的,这就意味着实用识别系统的识别性能会很不稳定。基于此,本文把人脸识别算法的识别性能随误配准增大而急剧下降的现象称为“误配准灾难”,以强调该问题的重要性。

2.3 “误配准灾难”产生原因分析

特征配准的目的本质上是要建立两幅图像各像素在“高层语义”上的对应关系，亦即使得待匹配的两个像素具有相同的高层语义。误配准就意味着图像之间像素高层语义不能严格相同，这对基于外观的识别方法来说，显然会造成“牛唇不对马嘴”的错误匹配。图 6(a-c)以一个极端的情形作例子，生动的阐述了“误配准灾难”问题产生的原因，假设图中实线矩形为原始图像模式（二值边缘图像），虚线为由于误配准而产生的图像模式，很显然，误配准将使得待匹配的两个像素具有完全不同的“高层语义”，如果采用基于外观的方法进行二者的匹配，对应像素灰度值之间的相似度根本上就不能反映输入模式之间的真实相似度，因而也就不可能做到良好的分类。对于基于外观匹配的人脸识别方法而言，情形尽管没有如此直观和极端，但也可以通过图 6(d-f)窥见一斑，其中 d, e, f 分别是图像发生平移、旋转和尺度变化后图像作差（取绝对值）的结果。不难看出，误配准的出现，导致了很大程度的无效匹配，因而计算出来的距离或者相似度就很难反应实际的人脸之间的相似度，并最终导致了识别性能的急剧下降。

从分类的角度讲，误配准显然增大了归一化的图像空间中各人脸类别的类内散度，也在一定程度上减小了不同人脸之间的类间散度，从而导致分类困难。特别的，对 Fisherface 方法而言，由于训练过程并没有考虑对误配准的学习，因而 Fisherface 方法得到的判别向量只在精确配准条件下满足类间散度与类内散度比值的最小化，而不能适应误配准条件，因而导致了识别性能的急剧退化。

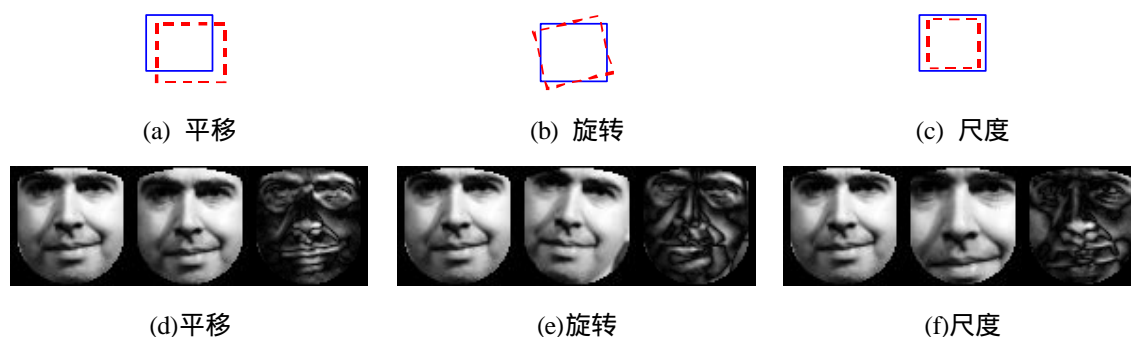


图 6 导致“误配准灾难”产生的根源：无效匹配图示

2.4 “误配准灾难”解决方法回顾与分类

即使是微小的特征配准误差都引起了人脸图像外观模式的非线性变化[4]，从而导致了模式类内变化不够紧凑，类间模式重叠增多，因而带来了分类困难。参考光学字符识别（OCR）等传统模式识别领域对变换不变分类问题的研究，本文将“误配准灾难”问题的解决方案划分为四大类，即：精确再配准方法、误配准不变的特征方法、配准误差学习方法、变换不变距离度量方法。当然，这四类方法实际上还可以进行结合，得到对误配准具有更强的鲁棒性的识别方法。这四类方法的基本思路及其相关的前人工作分别简述如下。

1、精确再配准方法

这类方法通常在特征配准算法的基础上，利用从分类器反馈回来的匹配信度或识别结果的信度进一步调整特征的配准，以便得到更高的匹配/识别信度，最终完成更加鲁棒的识别。这类方法由于引入反馈导致了算法时间复杂度的升高。

2、误配准不变的鲁棒特征方法

这类方法试图从输入人脸图像模式中提取对误配准不敏感的人脸描述特征，即提取的特征不会随误配准的增大而发生大的漂移，因而模式类别的表达更加紧凑（compact），从而使得后续的分类算法能够容易地实现鲁棒的识别。Gabor 小波特征使具有方向和尺度选择性的局部特征，也已经被证明具有更佳的误配准鲁棒性能[7,8,9,10,11]。弹性图匹配技术[8, 9]则在采用 Gabor 特征方面更进了一步，通过图匹配技术在精确定关键特征点的基础上进一步增强识别算法对配准误差的鲁棒性。Lampinen[12]则采用多层的自组织映射神经网络进一步分

类 Gabor 特征。高斯函数导数 (Gaussian Derivatives) 特征具有与 Gabor 特征类似的特性, 同样得到了在变换不变的模式识别中得到了应用[13]。另外一种常被采用的是所谓的高阶自相关特征(High-Order Autocorrelation, HOA) 或者矩特征[14]等, 这些特征被认为具有一定的仿射变换(平移、尺度和旋转) 不变性。

3、误配准学习方法

这类方法试图在分类器的训练阶段学习可能的配准误差[15, 16, 17], 从而使得最终得到的分类器不受误配准的影响。Martinez 的扰动(Perturbation)[16]方法就属于这一类, 该方法通过扰动扩大原型图像集 (Gallery), 从而增加了每个待识别人脸的图像数目, 然后采用高斯混合模型对每个人的图像进行建模, 因此, 该方法每个人脸都需要较大的存储空间来建模, 时间和空间复杂度都比较高。本文提出的方法则是在训练阶段通过加扰扩大训练集, 而并不需要扩大原型图像集, 因此其时间和空间复杂度都很低, 更加实用有效。

4、变换不变距离度量方法

模式识别最终在本质上要计算模式描述之间的距离或者相似度, 对外部变换不变的距离度量是一种自然的思路。切距离(Tangent Distance) 及其相关技术[4, 18]是这类方法中最典型的, 也是 OCR 领域广泛采用的技术。切距离方法认为一个原型模式经过各种仿射变换产生的模式在空间中扩展为一个流形结构, 该流形在原型模式点处可以用一个线性切空间来近似, 所谓切距离就是输入模式到该原型模式线性切空间的距离。鲁棒匹配(Robust Correlation)[19]和全局仿射变换(Global Affine Transform, GAT)[20]也可以归类为变换不变的距离度量方法, 其基本思路都是试图在进行相关量计算之前进行图像之间的仿射变换校准。

3. 考虑误配准鲁棒性的识别性能评价方法

不同算法的识别性能对误配准会有不同的鲁棒性, 对误配准过分敏感的识别算法是难以实用化的, 因此, 单纯采用单一的识别率指标来评价某一识别算法的性能是不够的, 我们需要一种新的评价体系来评价识别算法的性能。例如, 考虑三种人脸识别算法 A 和 B, 在相同的测试条件下 (已知眼睛标准位置), 其识别性能随误配准变化的情况如图 7 所示: A 算法在没有误配准的情况下识别率为 92%, B 算法在没有误配准的情况下识别率为 100%, C 算法则是将 A 算法性能曲线人为下调 10 个百分点后的结果。直观上, B 算法似乎明显优于 A 算法。但事实真得如此吗? 该如何合理地评价这两种识别算法的实际性能呢? 本文对此进行了深入探讨, 考虑了存在配准误差的情况下如何量化比较它们的系统识别性能[21]。

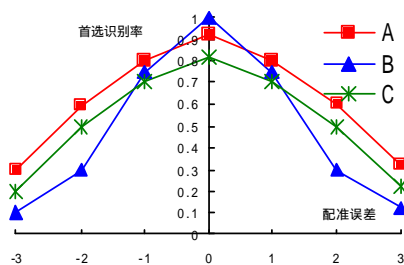


图 7 三种识别算法 A,B,C 的识别率与误配准的关系

我们进一步假设对 A, B, C 算法采用同样的眼睛定位算法进行系统集成。如前所述, 人脸特征对齐算法往往很难完美的定位面部的关键特征点, 为此, 本文考虑对特征对齐算法的精度以概率分布的方式进行建模。设特征点 (本文特指眼睛中心) 实际位置为 P^* , 而自动特征定位算法给出特征点位置为 P , 则可以合理的假定 P 偏离 P^* 的程度满足正态分布, 即:

$$p(\partial) \sim N(\mathbf{m}, \mathbf{S}^2),$$

其中 $\partial = d(P, P^*)$ 。

下面给出本文提出的考虑了误配准敏感性的识别算法性能评价方法：

定义 1 考虑误配准鲁棒性的正确识别率 r^* 定义为：

$$r^* = \int_{\Omega} P(\partial)r(\partial)d\partial, \quad (1)$$

其中： ∂ 为配准误差， Ω 为可能的配准误差值域范围， $P(\partial)$ 为眼睛定位算法的配准误差分布概率密度函数， $r(\partial)$ 为识别算法在误配准为 ∂ 时的识别率。

定义 r^* 实际是把误配准的分布作为权值计算识别算法的加权平均识别率，因而可以很好地反映识别算法在与特征对齐算法集成后的实际系统的识别性能，比单一的识别率具有更大的应用价值。

r^* 实际上同时考虑了识别性能与鲁棒性，但有时候单独考虑鲁棒性也是非常必要的。例如，对图 7 所示的算法 A 和 C 而言，C 算法的识别性能曲线是 A 曲线简单向下平移 10 个百分点而得到的，尽管 C 算法的平均识别性能下降了 10 个百分点，但直觉上二者对误配准的敏感性/鲁棒性应该是没有差别的（因为它们的下降速度一致），为了度量这一点，本文进一步提出下面的鲁棒性及其敏感性度量准则。

定义 2：识别算法对误配准的鲁棒性度量 R 定义为：

$$R = \int_{\Omega} P(\partial) \frac{r(\partial)}{r_0} d\partial = \frac{r^*}{r_0}. \quad (2)$$

其中， r_0 表示识别算法在没有误配准时的识别率。

R 反映了识别算法的识别性能受误配准的影响程度，其值域范围为 $(0, 1)$ ，该值越大表示识别算法越鲁棒，等于 1 是鲁棒性的理想情况：识别算法的性能不受误配准的影响。为了叙述的方便，本文还同时引入了识别算法对误配准的敏感性度量，取 R 的倒数是一个自然的选择。

定义 3：识别算法对误配准的敏感性度量 S 定义为：

$$S = 1/R = r_0/r^*. \quad (3)$$

S 的值域范围 $(1, \infty)$ ，该值越大表示该算法对误配准越敏感，相反，该值越接近 1，则表示识别算法的鲁棒性更强！

上述三个度量的引入，使得我们能够有效的对识别算法的总体识别性能进行量化的评估。例如，对图 7 所示的 A, B 算法，假设它们所使用的眼睛定位算法的定位误差满足 0 均值 1 方差的高斯分布，即： $p(\partial) \sim N(0,1)$

（值得指出的是：这意味着如果规定定位误差在 3 个像素以内为正确定位，定位正确率应为 99.7%），则根据公式 1~公式 3 计算出来的 A, B 算法的识别性能及其对误配准的鲁棒性和敏感性度量如表 2 所示。其结果表明：在误配准满足 0 均值 1 方差正态分布的条件下，算法 A 的实际识别性能要优于算法 B，算法 A 也具有更强的鲁棒性，算法 B 则对误配准更敏感——这一结论应该是更为合理的，也是合乎对图 7 的直觉感受的。

另外， R 和 S 的引入使得我们能够有效的抛开识别率的高低而单独评估识别算法对误配准的鲁棒性，例如对图 7 所示的算法 A 和 C，计算得到的 R, S 是相同的，这非常吻合我们对他们鲁棒性能的直观感觉。

表 2 考虑误配准敏感性后的 A, B, C 算法识别性能比较

算法	$r_o(\%)$	$r^*(\%)$	R	S
A	92	82.3	0.895	1.12
B	100	79.5	0.795	1.26
C	82	72.3	0.895	1.12

4. 基于配准误差学习的增强判别分析方法 (E-Fisherface)

误配准并不改变人脸图像类别属性,但却增大了归一化的图像空间中人脸类别的类内散度,也在一定程度上减小了不同人脸之间的类间散度,从而导致分类困难。因此,一种直观的解决方案就是将可能的误配准加入到训练集中,通过加强训练的方法来对误配准进行建模。基于此,本文提出了增强的 Fisherface 人脸识别方法,简称 E-Fisherface 方法。简单的说,该方法通过人工扰动给定的眼睛位置来产生若干存在误配准的“虚拟”人脸样本,并把这些样本作为正常样本用来训练 LDA,以使得 LDA 能够学习到误配准导致的类内/类间变化,从而增强 Fisherface 方法对误配准的鲁棒性。该方法的具体过程如下:

(1) PCA 变换矩阵的计算

与标准 Fisherface 方法一样,首先采用 PCA 进行高维图像特征的降维处理。需要指出的是,计算 PCA 的训练集有两种可能的选择:一种是采用“扰动”扩展后的训练集,另一种是只采用原始的未扩展训练集。考虑到计算的复杂度问题,本文采用了后一种策略来产生 W_{pca} ;

(2) 增强的 LDA 训练过程

对训练集中的任意图像,获取其眼睛特征位置后,采用图 8 所示的 8 邻域扰动方法扩展 LDA 的训练集合。也就是说,左右眼睛分别有 9 种不同的位置选择(包括精确位置),这样二者组合可以产生 81 种不同的误配准扰动图像,也就是说,训练集合中的每一幅图像都最终产生 81 幅归一化的人脸图像加入到训练集中,图 9 给出了若干图像的示例,这些扰动图像进一步用前面计算得到的 PCA 变换矩阵 W_{pca} 进行降维,然后训练 LDA 变换矩阵得到加强训练后的 W_{lda} 。由于这些存在配准误差的训练图像的加入,学习得到的线性判别分析模型可以在很大程度上将误配准造成的模式差别学习为类内变化而不是类间变化,从而提高识别的精度;

(3) 最近邻分类过程

得到 W_{pca} 和 W_{lda} 之后,识别过程与标准 Fisherface 方法一样,就变得相对简单了。即首先计算 Gallery 中的所有图像(眼睛位置相对精确已知)的 LDA 特征。然后对探测集中的任意图像,同样计算其 LDA 特征并与 Gallery 中所有已知 LDA 特征进行相似度计算排序以得到识别结果。

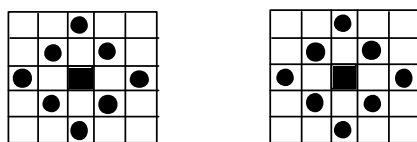


图 8 眼睛中心位置配准扰动的 8 邻域选择

需要指出的是,本文提出的 E-Fisherface 方法与 Martinez 的扰动(Perturbation)方法[16]是很不相同的: Martinez 的扰动方法扩大的是参考集(Gallery)而不是训练集,参考集中的每个原型人脸图像都被加扰,从而增加了每个待识别人脸的原型图像数目,然后采用高斯混合模型对这些原型图像进行建模,因此,该方法每个人脸都需要较大的存储空间来建模,时间和空间复杂度都比较高。而本文的方法并不增大 Gallery 集,而只是增大训练集,因此空间复杂度更低,时间耗费也主要是在训练过程,一旦训练完成,可以实现非常快速的特征提取和特征比对。





图9 根据单一输入图像生成的多个虚拟样本

5. 实验评估与分析

为了量化比较和分析上述识别方法对误配准的鲁棒性，本文分别在 FERET-FB 测试库和 CAS-PEAL-R1 正面人脸测试库上进行了对比实验。

5.1 在 FERET 上的实验

为了验证该方法的有效性，我们首先在 FERET 人脸数据库的 FB 测试集上进行了实验，实验方法同 2.1 节之描述。图 10 给出了 E-Fisherface 与 Fisherface 方法的识别性能随误配准大小变化而变化的对比情况，表 3 则给出了这两种算法在考虑了误配准后的综合识别性能和误配准鲁棒性/敏感性的对比情况。需要注意的是这些图表都是在双眼距离为 28 像素的条件下得出的。从图 10 不难看出，与 Fisherface 方法相比，E-Fisherface 方法的识别性能随配准误差增大而下降的趋势更加平缓，尤其是对于尺度和旋转情况。从表 3 则可以更方便的看出二者性能的量化比较情况，尤其对于旋转和尺度变化，E-Fisherface 的鲁棒性评测指标远远好于 Fisherface 方法。这充分说明本文提出的 E-Fisherface 方法具有更强的误配准鲁棒性。

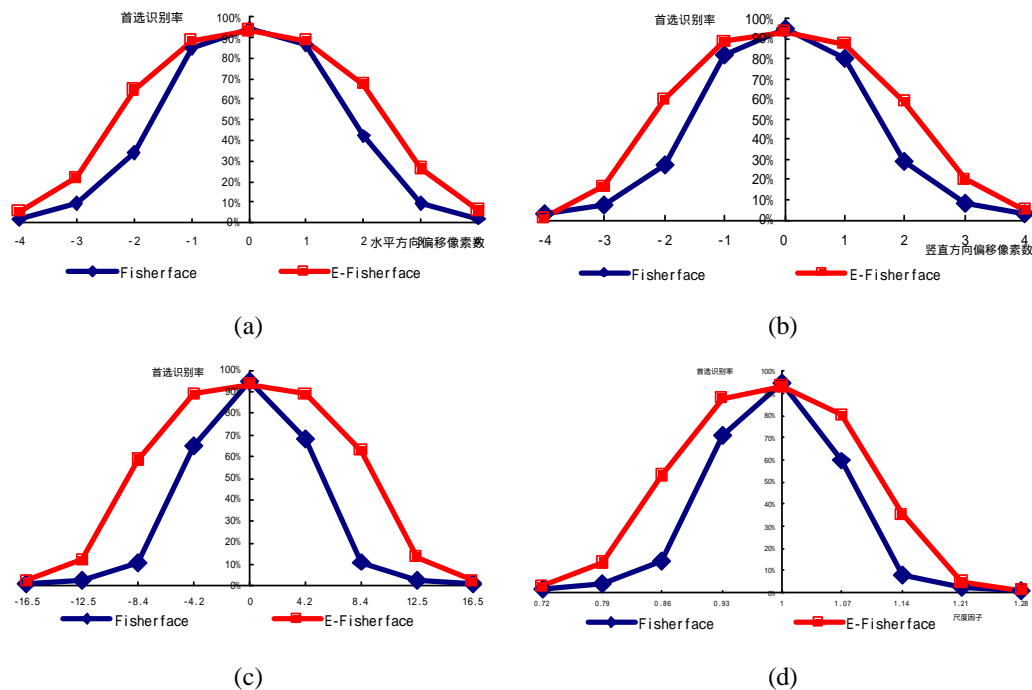


图 10 E-Fisherface 和 Fisherface 方法识别性能随配准误差变化情况比较

表 3 考虑了误配准后的综合识别性能和误配准鲁棒性/敏感性的对比

误配准	算法	r_0 (%)	r^* (%)	R	S
X 平移	Fisherface	94.8	83.6	0.881	1.13
	E-Fisherface	93.4	87.5	0.937	1.07
Y 平移	Fisherface	94.8	80.2	0.846	1.18
	E-Fisherface	93.4	86.4	0.925	1.08
旋转	Fisherface	94.8	71.2	0.751	1.33
	E-Fisherface	93.4	87.0	0.931	1.07
尺度	Fisherface	94.8	70.8	0.747	1.34
	E-Fisherface	93.4	82.9	0.887	1.13

值得指出的是，E-Fisherface 方法在没有配准误差的情况下识别性能 (r_0) 略有下降，这主要是因为加入扰动数据导致类内模式覆盖扩大，不同类别模式之间重叠的可能性也因此增大。尽管判别分析技术可以在一定程度上降低类内散度、增大类间散度，但还不够理想，因此导致了较精确配准条件下识别性能的略微下降。但这与方法对配准误差鲁棒性的优点相比是可以大体忽略不计的，下一个实验的结果也进一步表明了这一点。

5.2 在 CAS-PEAL-R1 上的实验结果

CAS-PEAL 共享人脸数据库包含 1,040 人的 30,900 幅图像，分为两个子库：正面图像子库和姿态图像子库。本文仅使用了其中的正面图像子库，包含 1,040 人的共 9,060 幅图像，其中涉及饰物、表情、光照等各种复杂的成像条件，该人脸库已经完全共享，详细的信息请参考 [22] 或者数据库发布网站 <http://www.jdl.ac.cn/peal/index.html>

为了进一步验证本文提出的问题及其解决方案的有效性，我们在 CAS-PEAL-R1 人脸库上评测了上述两种算法在与同一人脸检测、眼睛定位模块集成得到全自动的人脸识别系统时候的性能，其中的人脸检测、眼睛定位模块分别采用了文献[23]和[24]的方法，人脸图像归一化方法同 2.1 节所述，最终的实验结果见图 11 和表 4。

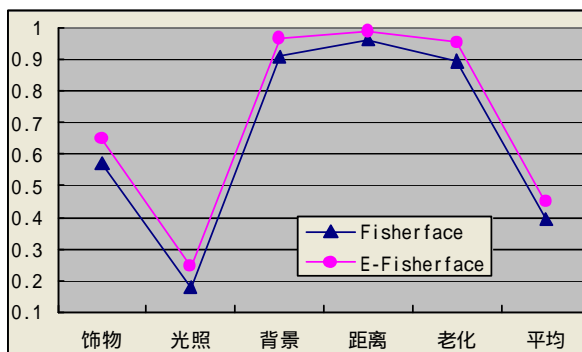


图 11 在 CAS-PEAL-R1 人脸库上的全自动测试性能比较

表 4 两种算法在 CAS-PEAL-R1 人脸库上的全自动测试性能比较

变化 方法	表情	饰物	光照	背景	距离	老化	加权平均
Fisherface	0.669	0.571	0.179	0.91	0.963	0.895	0.396
E-Fisherface	0.745	0.648	0.244	0.967	0.989	0.955	0.449

从图 11 和表 4 不难看出，由于全自动系统中眼睛的定位存在一定的误差，造成了一定程度的配准误差。对比全自动的 Fisherface 和 E-Fisherface 可以看出：由于对配准误差进行了学习，E-Fisherface 的识别率比

Fisherface 提高了 5.3个百分点。这个实验进一步说明了误配准灾难问题的重要性以及本文所研究的方法的有效性。

6. 结论及后续工作

为了揭示现有典型识别算法的识别性能对特征配准确度的敏感程度,本文通过对眼睛位置人为加扰的方法对 Fisherface 算法的识别性能随平移、旋转和尺度改变而变化的情况进行了实验评估,结果表明:Fisherface 的识别性能随着误配准的增大而急剧下降——我们称这一现象为“误配准灾难”。本文进一步提出了一种考虑了误配准鲁棒性的识别性能评估方法,给出了识别算法对误配准的鲁棒性度量策略;最后提出了一种对误配准进行强化学习的“误配准灾难”解决方案 E-Fisherface。在 FERET 人脸图像数据库上的实验表明该方法可以有效的提高识别算法对误配准的鲁棒性。本文的主要贡献包括:

1. 首次系统研究了人脸识别中的“误配准灾难”问题。我们首先通过实验评估了眼睛位置在偏离手工标定真实位置的情况下 Fisherface 方法性能的变化情况,结果发现:Fisherface 方法的识别性能随着误配准(反映为平移、旋转和尺度等仿射变换参数的变化)的增大而急剧下降,本文称这一现象为“误配准灾难”问题;
2. 提出了一种考虑了误配准分布的人脸识别系统识别性能评价方法,以弥补识别率变化曲线难以量化比较不同算法性能的不足,并给出了衡量识别算法对误配准鲁棒性和敏感度的相应度量;
3. 针对 Fisherface 方法的“误配准灾难”问题,提出了一种增强判别分析方法 E-Fisherface,该方法在模型训练阶段通过对误配准进行学习,可以有效的降低原有判别分析方法得到的判别特征对误配准的敏感性,大大提高了 Fisherface 算法的实用程度;

我们将在以后的工作中考虑变换不变鲁棒特征与配准误差学习的融合,例如 Gabor 特征与误配准学习策略的融合,以期能够得到更为鲁棒的系统识别性能。

参考文献

- [1] Chellappa R., Wilson C.L., Sirohey S., Human and machine recognition of faces: a survey, Proceedings of the IEEE, 1995, 83(5): 705-740
- [2] Phillips P.J., Moon H., etc. "The FERET evaluation methodology for face-recognition algorithms," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(10):1090-1104
- [3] Philips P.J., Grother P.J., Micheals R.J., Blackburn D.M., E.Tabassi, and J.M.Bone, Face Recognition Vendor Test 2002: evaluation report, Technical Report, NISTIR 6965, National Institute of Standards and Technology, 2003 (<http://www.frvt.org>)
- [4] Simard P.Y., LeCun Y.A., Denker J.S., and Victorri B, Transformation invariance in pattern recognition-tangent distance and tangent propagation. International Journal of Imaging System and Technology, 2001, 11(3):181-194
- [5] P.N.Belhumeur, J.P.Hespanha etc. Eigenfaces vs Fisherfaces: recognition using class specific linear projection, 1997, IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(7):711-720
- [6] Shan S., Chang Y., Gao W., Cao B., Curse of mis-alignment in face recognition: problem and a novel mis-alignment learning solution, in Proceeding of The 6th International Conference on Face and Gesture Recognition, Jeju, Korea, 2004, 314-320
- [7] Buhman J., Lades M., and Malsburg C. von der, Size and distortion invariant object recognition by hierarchical graph matching, in Proc. of IEEE International Joint Conference on Neural Network, San Diego, CA, Jan. 1989, vol.2:411-416.
- [8] Lades M, Vorbruggen J.C., etc, Distortion invariant object recognition in the dynamic link architecture, IEEE Trans. On Computers, 1993, 42(3):300-311
- [9] Wiskott L, Fellous J.M, Kruger N., Malsburg C. von der, Face recognition by elastic bunch graph matching, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7):775-779

- [10] Rolf P. Wurtz, Object recognition robust under translations, deformations, and changes in background, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7):pp769-775
- [11] Shan S., Gao W., Chang Y., Cao B., Yang P., Review the strength of Gabor features for face recognition from the angle of its robustness to mis-alignment, in Proceeding of International Conference on Pattern Recognition, Cambridge, UK, 2004, vol.1, pp338-341
- [12] Lampinen J. and Oja E.. Distortion tolerant pattern recognition based on self-organizing feature extraction. IEEE Transactions on Neural Networks, 1995, 6(3):539-547
- [13] Schiele B, Crowley J. L., Recognition without correspondence using multidimensional receptive field histograms, International Journal of Computer Vision, 2000, 36(1):31-50
- [14] Perantonis S. J. and Lisboa P. J. G., "Translation, rotation, and scale invariant pattern recognition by high-order neural networks and moment classifiers," IEEE Trans. Neural Network, Mar. 1992 vol.3:241-251,
- [15] Thien M. Ha and Horst Bunke, Off-Line, Handwritten numeral recognition by perturbation method, IEEE Transactions On Pattern Analysis and Machine Intelligence, 1997, 19(5):535-539,
- [16] A.M. Martinez. Recognizing imprecisely localized, partially occluded and expression variant faces from a single sample per class, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(6):748-763
- [17] Shan S.G., Gao W., Zhao D.B., Face Identification Based On Face-Specific Subspace, International Journal of Image and System Technology, 2003, 13(1):23-32
- [18] Keysers D., Macherey W., Ney H., Dahmen J., Adaptation in statistical pattern recognition using tangent vectors. In IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(2):269-274
- [19] Jonsson K., Kittler J., Li Y., Matas J: Support vector machines for face authentication. Image and Vision Computing, 2002, 20(5-6): 369-375
- [20] Wakahara T., Kimura Y., Tomono A., Affine-invariant recognition of gray-scale characters using global affine transformation correlation, IEEE Trans. on Patten Analysis and Machine Intelligence, 2001, 23(4):384-394
- [21] Shan S., Study on Some Key Issues in Face Recognition [Ph.D dissertation], Institute of Computing Technology, Chinese Academy of Sciences, Beijing, 2004 (In Chinese)
(山世光, 人脸识别中若干关键问题的研究 [博士学位论文], 中科院计算所, 北京, 2004)
- [22] Cao B., Shan S., Zhang X., Gao W., Baseline Evaluations On The CAS-PEAL-R1 Face Database, in Lecture Notes in Computer Science (LNCS3338), Advances in Biometric Person Authentication, Guangzhou, 2004, 370-378
- [23] Viola P, Jones M., Rapid Object Detection using a Boosted Cascade of Simple Features, in Proc. of International Conference on Computer Vision and Pattern Recognition, Kawaii, 2001, vol. I:511-518
- [24] Cao B., Shan S., Gao W., Zhao D., "Localizing the Iris Center by Region Growing Search," Proceedings of IEEE International Conference on Multimedia and Expo, Vol. 2, pp129-132, 2002.

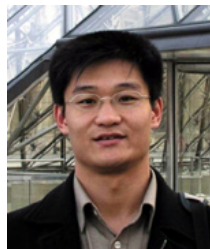
资助情况：

本课题得到国家自然科学基金项目(60332010)，国家十五863计划(2002AA118010)，同时还得到了银晨智能识别科技有限公司的大力资助。

中文作者简介：

山世光，男，1975年3月生，博士，助理研究员，研究方向为模式分析，机器学习，计算机视觉，尤其专注于人脸识别问题。高文，男，1956年生，博士，教授，博士生导师，主要研究领域：多媒体数据压缩，图像处理，计算机视觉，多模式接口，人工智能，虚拟现实等。唱轶钲，女，1979年生，硕士生，主要研究领域：人脸识别。曹波，男，1977年生，博士生，主要研究领域：模式分析与机器学习。陈熙霖，男，1965年生，博士，教授，博士生导师，主要研究领域：计算机视觉，模式识别，机器学习，智能人机交互等

英文作者简介：



SHAN Shiguang, male, born in March 1975, Ph.D., Assistant Researcher in Institute of Computing Technology, Chinese Academy of Sciences. His research interests include pattern analysis, machine learning, and computer vision. He is especially absorbed in face recognition related research topics.

GAO Wen, male, born in 1956, Ph.D., Professor in Institute of Computing Technology and Graduated School of Chinese Academy of Sciences. He is also a professor in Harbin Institute of Technology. His research interests include multimedia data compression, image processing, computer vision, multimodal interface, artificial intelligence and virtual reality.

CHANG Yizheng, female, born in 1979, Master candidate in School of Computer, Harbin Institute of Technology. Her research is focused on face recognition.

CAO Bo, male, born in 1977, Ph.D Candidate in Institute of Computing Technology, Chinese Academy of Sciences. He is currently working on face recognition.

CHEN Xilin, male, born in 1965, Ph.D., Professor in Institute of Computing Technology, Chinese Academy of Sciences and School of Computer, Harbin Institute of Technology. His research interests include computer vision, pattern recognition, machine learning, and intelligent human-computer interface.

联系人：山世光；电话：010-58858300-314；Email: sgshan@ict.ac.cn

Background of the article:

This work is conducted in the ICT-ISVISION Joint R&D Lab for face recognition (hereinafter named by FRJDL), Institute of Computing Technology Chinese Academy of Sciences, which has been working on face recognition related researches for more than 9 years as the first specialized lab on face recognition in China. FRJDL has always been aiming at inventing fully innovative kernel technologies and dealing with key issues in Biometrics, especially in face recognition area, based on which, multiple disciplines such as pattern recognition, machine learning, and computer vision, are theoretically studied. In the last 9 years, plenty of research work has been done on the basic theories in face recognition, as well as the practical engineering techniques. About 100 research papers had been published by the FRJDL members. And many of our research fruits have been successfully applied to practical applications such as bank surveillance, access control, and time-card systems.

The work presented in this paper is just motivated by one of our practical fully automatic face recognition systems. In the system, we noticed that a large percent of false identification events are caused actually by the inaccurate eye localization results. To improve the robustness of the system to this kind of errors, we have conducted the work presented in this paper.