

# 基于 Fisher 投影的多光谱人脸融合识别

刘典婷, 华顺刚, 苏铁明, 欧宗瑛, 张建新

(大连理工大学机械工程学院 CAD 与网络技术研究所, 大连 116024)

**摘要:** 根据可见光图像与红外图像的信息互补性, 分析在决策层融合识别中的归一化法和融合算法, 提出一种基于统计的 Fisher 投影融合法, 利用 Fisher 线性判别准则在二维分数空间寻找最优投影方向, 使不同类样本投影后能最佳分离。在多光谱人脸融合识别中的应用结果表明, 与其他融合算法相比, 该算法具有更好的识别效果。

**关键词:** 红外图像; 可见光图像; Fisher 线性判别; 人脸识别; 融合

## Multi-spectral Face Fusion Recognition Based on Fisher Projection

LIU Dian-ting, HUA Shun-gang, SU Tie-ming, OU Zong-ying, ZHANG Jian-xin

(Institute of CAD & Network Technology, School of Mechanical Engineering, Dalian University of Technology, Dalian 116024)

**【Abstract】** According to the information complementarity for visible image and infrared image, this paper analyzes normalization method and fusion algorithm for fusion recognition at decision level, proposes a Fisher projection fusion algorithm based on statistics. This algorithm employs Fisher linear discriminate criterion to find out the optimal projection in score space that can best separate projected samples from different classes. Application of multi-spectral face fusion recognition results show that this algorithm has higher recognition effect compared with other fusion algorithm.

**【Key words】** infrared image; visible image; Fisher linear discrimination; face recognition; fusion

### 1 概述

随着人脸识别技术的快速发展, 在控制光照条件的情况下, 现有人脸识别算法在可见光人脸图像识别方面已取得较大进展, 但是识别效果受光照条件变化影响较大。红外人脸图像具有成像几乎不受可见光照变化影响的特点, 近几年, 受到学者们的广泛重视。实验证明, 根据红外与可见光人脸图像各自的特性及其在人脸识别中的互补性, 将 2 种传感器图像在识别过程中进行融合可以进一步提高人脸识别的正确率<sup>[1]</sup>。

识别处理正常包括以下 3 个阶层: 数据预处理, 特征提取和识别决策。融合处理对应在这 3 个层次进行: 数据层, 特征层和决策层。数据层融合直接对原始数据进行综合分析。其优点是保持了尽可能多的客体信息, 基本不发生信息丢失或遗漏; 缺点是处理数据多、耗费时间长、实时性差。特征层在特征提取的基础上执行综合分析处理。其优点是能保持足够数量的重要信息, 又较大稀释了数据量。缺点是涉及不同信息的物理和分布规律, 较难把握。决策层融合是根据一定准则和每个传感器的决策执行综合评判, 给出统一的最终决策。它可以根据真正的分布进行统计分析, 易于核实并取得可信的结果, 近年来在这方面的相关研究已取得一定进展。

文献[2]使用双曲线正切法进行分数归一化, 使用加权几何平均法对语音和人脸信息进行融合, 同时为一个多模态认证系统提出了一种层叠的组合认证策略。文献[3]对多种融合策略(求和、乘积、最大、最小、中值)进行比较试验后发现求和策略的效果优于其他策略。文献[1]使用求和策略<sup>[3]</sup>对红外和可见光人脸图像在决策层上进行融合, 实验证明融合后的识别率高于采用单一光谱图像的识别率。上述研究成果表明, 通过决策层融合可提高识别率, 但是如何选择最优的融

合方法仍是一个研究的热点。

本文在列举和分析目前决策层常用融合方法的基础上, 提出一种基于统计的 Fisher 投影融合算法。

### 2 融合前的归一化法

不同模态生物特征的得分在进行融合前, 对原始评价分数进行归一化是很必要的。例如, 一个匹配器产生的评价得分区间为[100, 1 000], 而另一个匹配器产生的评价得分区间为[0, 1], 如果在 2 项得分进行融合前不进行得分归一化, 那么决策融合后, 第 2 个匹配器对识别的贡献将被忽略。

本文介绍 3 种常用的归一化方法, 用  $S$  表示红外或可见光匹配器产生的全部匹配得分的集合,  $s$  为  $S$  中的一个原始得分,  $n$  为对  $s$  进行归一化后的得分。

#### (1) Max-Min 法

该方法将原始得分映射至[0, 1]区间。 $\max(S)$ 和  $\min(S)$ 分别为  $S$  集合中的最大值和最小值。

$$n = \frac{s - \min(S)}{\max(S) - \min(S)} \quad (1)$$

#### (2) Z 分法

该方法将原始得分变换为平均值为 0, 标准偏差为 1 的分布。 $\text{mean}()$ 和  $\text{std}()$ 分别为平均值和标准偏差算子。

**基金项目:** 大连理工大学交叉学科基金资助项目(光电子+X2006); 大连理工大学与中国科学院沈阳自动化研究所联合基金资助项目(DUT-SIA2006)

**作者简介:** 刘典婷(1981 - ), 女, 博士研究生, 主研方向: 人脸识别, 模式识别, 图像处理; 华顺刚, 副教授; 苏铁明, 讲师; 欧宗瑛, 教授、博士生导师; 张建新, 博士研究生

**收稿日期:** 2009-10-09 **E-mail:** d.liu4@umiami.edu

$$n = \frac{s - \text{mean}(S)}{\text{std}(S)} \quad (2)$$

### (3) 双曲正切法

该方法被称为最具鲁棒性的统计技术。它映射原始得分至[0,1]区间。 $\text{Tanh}()$ 为双曲正切算子。

$$n = \frac{1}{2} \left[ \tan h \left( 0.01 \frac{(s - \text{mean}(S))}{\text{std}(S)} \right) + 1 \right] \quad (3)$$

上述 3 种方法从不同角度对原始得分进行归一化, 针对分布情况不同的匹配器得分产生不同的归一化结果。

## 3 基于统计的 Fisher 投影融合算法

融合不同模态生物特征的目的是通过结合不同模态的生物特征来提高身份鉴别系统的认证率, 得到更可靠、全面、准确的认证结果。通常的身份认证系统解决的问题可描述为: 在特定样本集中, 用户称自己属于其中的某一个身份, 系统通过比较匹配值与事先设定的门限做出决策: 接受或拒绝。决策存在 2 种可能的错误: (1) 假冒者(impostor)被接受为具有真实身份的用户, 该错误用错误接受率(False Accept Rate, FAR)度量; (2) 具有真实身份的用户(client)被拒绝, 认为其是假冒者, 该错误用错误拒绝率(False Reject Rate, FRR)度量<sup>[4]</sup>。对于既定的识别系统, FAR 与 FRR 是相互联系的。当门限取值增高时, FAR 减小, FRR 增大; 当门限取值降低时, FAR 增大, FRR 减小。FAR 与 FRR 相等时的数值称为系统的等错误率(Equal Error Rate, EER)。

对于可见光归一化后的评价得分用  $n_{VS}$  表示, 红外归一化后的评价得分用  $n_{IR}$  表示,  $f$  为  $n_{VS}$  与  $n_{IR}$  融合后的新得分。常用决策层融合方法有:

#### (1) 求和法(sum)

将 2 个匹配器所得评价分数之和作为新的得分。

$$f = n_{VS} + n_{IR} \quad (4)$$

#### (2) 最大值法(max)

将 2 个匹配器评价得分中最大的作为新的得分。

$$f = \max(n_{VS}, n_{IR}) \quad (5)$$

#### (3) 线性加权法

线性加权(Linear Weighting, LW)融合是使用训练时得到的 EER 计算分配给不同匹配器评价得分的权值。对错误率较小的匹配器赋予较大的权值, 错误率较大的匹配器赋予较小的权值。若用  $e_{VS}$  表示单可见光图像的等错误率,  $e_{IR}$  为单红外图像的等错误率, 则可见光和红外匹配器的权值  $\omega_{VS}$  和  $\omega_{IR}$  分别为

$$\begin{cases} \omega_{VS} = \frac{1}{\frac{1}{e_{VS}} + \frac{1}{e_{IR}}} = \frac{e_{IR}}{e_{VS} + e_{IR}} \\ \omega_{IR} = \frac{1}{\frac{1}{e_{VS}} + \frac{1}{e_{IR}}} = \frac{e_{VS}}{e_{VS} + e_{IR}} \end{cases} \quad (6)$$

由式(6)可见, 匹配器的权值与其本身的等错误率呈反比。融合后的得分  $f$  按下式计算:

$$f = \omega_{VS}n_{VS} + \omega_{IR}n_{IR} \quad (7)$$

除以上 3 种最常用的融合方法外, 本文提出一种基于统计的 Fisher 投影法。该方法利用 Fisher 线性判别准则<sup>[5]</sup>在二维分数空间寻找一个最优的投影方向  $w$ , 使投影在这个方向

上的 impostor 类及 client 类样本能最好地分开, 如图 1 所示。

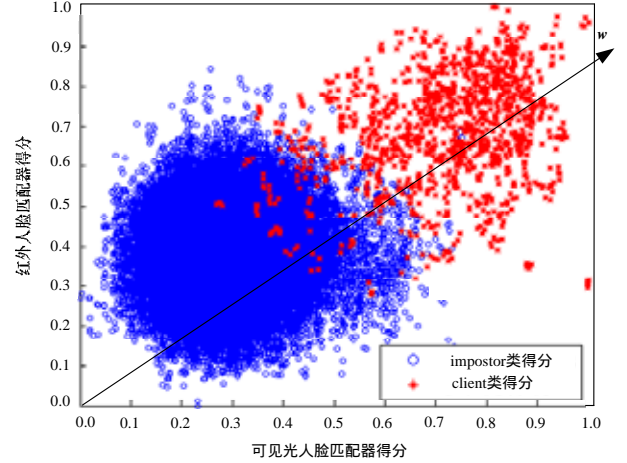


图 1 impostor 类和 client 类得分分布及投影方向  $w$  的示意图

具体地, 对一个红外与可见光评价得分样本集  $X = \{x/x \in \mathcal{R}^2\}$ , 每一个  $x$  属于 impostor 类或 client 类。定义类间散度矩阵为

$$S_B = \sum_{i=1}^2 N_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (8)$$

其中,  $N_i$  为第  $i$  类  $X_i$  的样本数;  $\mu_i \in \mathcal{R}^2$  为第  $i$  类的样本平均值;  $\mu$  为样本集  $X$  中所有样本的平均值。定义类内散度矩阵为

$$S_W = \sum_{i=1}^2 \sum_{x \in X_i} N_i (x - \mu_i)(x - \mu_i)^T \quad (9)$$

如果  $S_W$  是非奇异阵, 依据 Fisher 线性判别准则, 最优投影为

$$w = \arg \max_w \frac{|w^T S_B w|}{|w^T S_W w|} = [w_1 \ w_2] \quad (10)$$

其中,  $w_1$  和  $w_2$  是  $S_B$  和  $S_W$  的广义特征值, 分别对应广义特征向量  $\{\lambda_i, i=1,2\}$ , 使得

$$S_B w_i = \lambda_i S_W w_i \quad i=1,2 \quad (11)$$

对应最大特征值  $\lambda_1$  的投影向量  $w_1 \in \mathcal{R}^2$  即为本文所求取的最佳投影方向  $w$ 。若取  $n = [n_{VS} \ n_{IR}]^T$ , 则融合后的得分为

$$f = w^T n \quad (12)$$

## 4 Fisher 投影在多光谱人脸融合识别中的应用

### 4.1 实验数据

本文实验中的图像于 2002 年在美国 Notre Dame 大学<sup>[6]</sup>采集, 以在不同时间段分多次采集的方式对 63 人进行 1 次~10 次不等的图像采集。在拍照时, 在拍照对象的左、中、右各放置一盏照明灯。中间灯熄灭, 两边灯亮的光照设置简称为“LF”; 3 盏灯都亮的光照设置简称为“LM”。在每种光照设置下, 对人物拍摄 2 幅图像: 一幅为中性表情简称为“FA”; 另一幅微笑表情简称为“FB”。因此, 拍照对象每人每次采集含 2 种表情和 2 种光照条件的红外与可见光图像各 4 幅。本文将实验图像按照采集时光照和表情的不同平均分为 4 组, 每组各含 494 对红外与可见光图像, 分别标记为 FA|LF, FA|LM, FB|LF 和 FB|LM。图 2 为一组在 Notre Dame 大学采集的可见光与红外人脸图像。在实验前, 本文先对所有实验图像进行光照补偿并将图像手工旋转缩放至  $100 \times 128$  像素大小, 两瞳孔间距统一调整至 50 像素大小。训练时, 从每组图像集中抽取 63 对红外和可见光图像(每人一对), 共计 252 对

图像,采用先主成分分析后线性判别分析算法(PCA+LDA 算法)分别对红外和可见光图像进行训练,生成投影空间。剩余的每组 431 对图像作为测试集,验证本文算法的有效性。

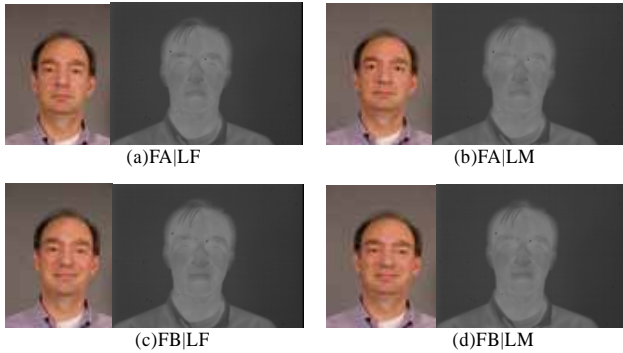


图 2 一组可见光与红外人脸图像

#### 4.2 实验结果与分析

本文首先测试单传感器图像在不同光照和表情情况下的等错误率,如表 1 所示。

表 1 单红外和单可见光图像等错误率比较 (%)

光谱	FA LF	FA LM	FB LF	FB LM
红外	9.501 3	11.096 9	9.753 6	10.758 7
可见光	2.027 5	3.452 4	3.926 3	2.644 5

从表 1 可看出,在所有测试集上单可见光图像比单红外图像取得较低的等错误率。FA|LF 测试集相对于其他 3 组测试集取得更好的识别效果。由于被测试图像是分多次采集的,其时间间隔为 1 周左右,而同一个人在不同时间可能因为某些原因(如环境温度、情绪、心理状态、生理状态)使面部温度产生变化,从而反应在红外成像中,使不同时间段采集的红外图像产生一些差异,影响红外识别结果,因此在光照不同(LF 和 LM)时,红外图像的识别结果有很大差别。

表 2~表 4 分别为采用 Max-Min 法,Z 分法和双曲正切法进行得分归一化后,不同融合算法对 2 类图像在决策层的融合识别结果比较。

表 2 Max-Min 法归一化后 4 种融合算法的等错误率比较 (%)

算法	FA LF	FA LM	FB LF	FB LM
Sum	1.850 5	3.211 8	2.459 1	2.156 3
Max	3.302 1	3.608 5	3.754 8	3.603 2
LW	1.822 3	3.189 1	2.277 9	2.277 9
本文算法	1.669 2	3.189 1	2.277 9	1.740 1

表 3 Z 分法归一化后 4 种融合算法的等错误率比较 (%)

算法	FA LF	FA LM	FB LF	FB LM
Sum	1.822 3	3.046 3	2.277 9	1.885 5
Max	2.324 3	3.416 9	3.086 1	2.555 0
LW	1.822 3	3.189 1	2.277 9	2.277 9
本文算法	1.672 6	3.189 1	2.277 9	1.921 5

表 4 双曲正切法归一化后 4 种融合算法的等错误率比较 (%)

算法	FA LF	FA LM	FB LF	FB LM
Sum	4.902 0	6.525 5	5.449 5	6.473 6
Max	3.044 3	3.415 6	3.297 4	3.043 0
LW	2.219 8	4.390 8	5.296 1	2.632 9
本文算法	1.679 0	3.189 1	2.277 9	1.929 3

比较表 1~表 4 的结果可看出,红外与可见光图像在决策层的融合,可充分利用 2 类图像的互补性,在大多数情况下,识别率比单红外和单可见光图像有较大提高。

在所有比较试验中,本文算法仅在微笑表情的条件下,采用 Z 分法归一化时,等错误率略高于求和法(见表 3)。本文算法法与其他融合方法相比均取得较高的识别率。另一方面,从表 2~表 4 可见,采用不同的归一化算法时,本文算法在相同测试集上取得的识别率较稳定。其他 3 种融合算法对归一化方法比较敏感,采用不同的归一化算法取得的识别率波动较大。

#### 5 结束语

本文研究了已有的关于红外和可见光人脸图像在决策层上的归一化和融合算法,并提出一种基于统计的 Fisher 投影融合算法。应用结果表明,本文算法具有较高的识别率与稳定性。由于 Fisher 投影法是建立在统计分析基础上的,因此训练数据数量越大、与真实测试数据分布越接近,融合算法取得的效果越显著。本文算法可用于红外和可见光图像的两类融合,也可用于更多类生物特征融合。这种新的分类器融合算法对生物识别,以及模式识别的其他领域也具有一定参考价值。

#### 参考文献

- [1] Chen Xin, Flynn P J, Bowyer K W. PCA-based Face Recognition in Infrared Imagery: Baseline and Comparative Studies[C]//Proc. of IEEE International Workshop on Analysis and Modeling of Faces and Gestures. [S. l.]: IEEE Press, 2003.
- [2] Brunelli R, Falavigna D. Person Identification Using Multiple Cues[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(5): 955-966.
- [3] Kittler J, Hatef M, Robert P W, et al. On Combining Classifiers[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(3): 226-239.
- [4] Lin Hong, Jain A. Integrating Faces and Fingerprints for Personal Identification[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(12): 1295-1300.
- [5] Fisher R A. The Use of Multiple Measures in Taxonomic Problems[J]. Annals of Eugenics, 1936, (7): 179-188.
- [6] CVRL Data Sets. Biometrics Data Sets[Z]. [2009-02-26]. [http://cse.nd.edu/~cvrl/CVRL/Data\\_Sets.html](http://cse.nd.edu/~cvrl/CVRL/Data_Sets.html).

编辑 陆燕菲

(上接第 179 页)

- [4] 庄越挺,毛 炜,吴 飞,等.基于隐马尔可夫链的广播新闻分割分类[J].计算机研究与发展,2002,9(9):1057-1063.
- [5] Chen Shaobing, Gopalakrishnam P. Speaker, Environment and Channel Change Detection and Clustering via the Bayesian Information Criterion[C]//Proc. of DARPA Broadcast News Transcription Understanding Workshop. Landsdowne, USA: [s. n.],

1998.

- [6] Cheng Shih-Sian, Wang Hsin-Min, Fu Hsin-Chia. BIC-based Audio Segmentation by Divide-and-conquer[C]//Proc. of ICASSP'08. Las Vegas, USA: CRC Press, 2008.
- [7] 贾 磊,穆向禹,徐 波,等.广播语音的音频分割[J].中文信息学报,2002,16(1):37-42.

编辑 张 帆