

# 动态视频人脸考勤方法研究与实现<sup>\*</sup>

王 刚<sup>1,2</sup> 周激流<sup>2</sup> 何 坤<sup>2</sup> 杨柱中<sup>3</sup> 吴俊强<sup>3</sup> 朗方年<sup>3</sup>

(四川警察学院 省公安厅警官培训基地 公安部警务实战训练四川基地 泸州 646000)<sup>1</sup>

(四川大学计算机学院 成都 610065)<sup>2</sup> (四川大学电子信息学院 成都 610065)<sup>3</sup>

**摘要** 研究并实现了利用 Gabor 滤波器和 Fisher 线性鉴别分析(FLDA)方法的动态人脸识别考勤系统。系统实现的基本思想是运用 Gabor 变换提取人脸的局部特征和经过 Gabor 处理后使得人脸对光照变化不敏感;进一步利用 FLDA 来降维和隐含地提取最有利于分类的最佳鉴别特征;最后将视频采集的考勤图像与训练库中的图像通过比对,得出识别结果。实验结果表明,利用该方法开发的动态人脸识别考勤系统具有识别率高、实用性好、可靠性强等特点。  
**关键词** 考勤系统,视频序列,人脸识别,Gabor 滤波器,FLDA,特征提取文献标识码

## Study on the Approaches of Dynamic Attendance in Video Sequence

WANG Gang<sup>1,2</sup> ZHOU Ji-Liu<sup>2</sup> HE Kun<sup>2</sup> YANG Zhu-Zhong<sup>3</sup> WU Jun-Qiang<sup>3</sup> LANG Fang-Nian<sup>3</sup>

(Sichuan Police College, Ministry of Public Security Police Service Live Operational Training Sichuan Base, Luzhou 646000)<sup>1</sup>

(College of Computer Science, Sichuan University, Chengdu 610065)<sup>2</sup>

(College of Electronic Information, Sichuan University, Chengdu 610065)<sup>3</sup>

**Abstract** Dynamic Face Recognition based Attendance System(DFRAS) by using Gabor filter and Fisher Linear Discriminant Analysis(FLDA) is realized successfully. In this paper, the algorithm's idea for DFRAS is that human local feature is extracted by using Gabor conversion, furthermore, the best discriminant features conducive to classification are extracted by FLDA, and then recognition outcomes are come from comparing video face image with trained image base. Experimental results indicate that this method for Face Recognition based Attendance System has the advantages of high recognition rate, good practicability, robustness and so on.

**Keywords** Attendance system, Video sequence, Face recognition, Gabor filter, Fisher linear discriminant analysis, Feature extraction

近年来,特别是美国“9.11”事件以来,生物特征识别技术得到了快速的发展。与此同时,人脸检测和识别技术也在不断的发展和进步,与指纹、掌纹、虹膜等其他人体生物特征识别系统相比,人脸识别具有更加直接、友好、便捷、使用者无心理障碍、非侵犯性等特点,因而容易为用户所接受,并且其在身份验证、公安刑侦、安全验证、档案管理、智能视频监控、视频监控、智能人机交互系统等方面都有着广泛的应用。

目前将人脸识别技术应用于考勤系统的报道很少。张婷等人<sup>[1]</sup>提出了一种用于考勤的人脸识别算法,但存在着考勤前需要为具有访问权限的每人制作一张 IC 卡的不足,如果 IC 卡丢失或损坏,则给考勤带来一定的麻烦。本文将运用 Gabor 滤波器和 Fisher 线性鉴别分析(FLDA)的人脸识别技术应用于单位考勤系统,采用建立人脸图像训练库代替文<sup>[1]</sup>提出的 IC 卡,解决了 IC 卡在丢失或损坏情况下存在的问题,并实现了动态人脸识别考勤系统。动态人脸识别考勤系统(Dynamic Face Recognition-based Attendance System, DFRAS),简称人脸考勤系统,是指利用人体面相所具有的唯一性、排他性而研制的考勤系统。它克服了传统打卡中磁卡、IC 卡等考勤方式存在的代打卡、卡丢失等不足和缺陷,最为有效地杜绝了考勤管理中的人为因素,充分体现了考勤管理的公正,避免不必要的人事纠纷。

## 1 人脸考勤系统概述

人脸考勤系统,是将先进的人脸识别技术应用到考勤领域的一个成功产品。该产品首先录员工的文字资料和动态采集人脸图像并训练,建立人脸图像库。考勤时,通过与计算机相连的摄像头动态捕捉员工的脸部照片,同时把捕捉到的人脸照片与预先采集的照片进行比对,如果是员工本人,则考勤成功,同时根据考勤时间判断人员出勤情况(正常、迟到、早退或加班等),并记录考勤信息,如考勤日期、班次、上班時間等信息。如果不是员工本人考勤,则系统给出友好提示,并进行相应处理。所有考勤数据通过主机程序处理汇总后,生成相应管理报表或数据文本,可提供给人事部门进行人事考核或财务部门进行工资结算。

## 2 人脸考勤系统的关键技术

人脸考勤系统由人脸识别子系统和考勤管理子系统组成,其核心部分在于人脸识别子系统,下面就这一部分的关键内容加以论述。

### 2.1 视频采集

动态采集视频图像是系统开发工作的基础。视频采集使用微软公司 Windows 操作系统提供的 VFM(Video for Windows)库函数<sup>[2]</sup>来对 USB 摄像头进行操作。VFM 是 Win-

<sup>\*</sup> 基金项目:国家自然科学基金(60272095)和教育部博士点基金(20020610013)。王 刚 讲师,专业技术一级警司,主要研究方向为网络安全、生物特征识别、电子证据检验鉴定;周激流 教授,博士生导师,主要研究方向为数字图像处理、生物特征识别、分数阶微积分。

dows 环境下实现实时视频捕获的视频模块库,其中 AVI-CAP. DLL 模块是 VFW 的一个重要组成部分,它的主要作用是实现视频捕获。AVICAP 为应用程序提供了一个简单的、基于消息的接口,通过该接口,程序可以访问视频和波形音频硬件并控制视频流到硬件的捕获。AVICAP 支持实时视频流捕获和单帧捕获并提供对视频流的控制。它能直接访问视频缓冲区,而不需要生成中间文件,实时性很强,效率高。在进行视频捕获之前需要创建一个捕获窗类,它是所有捕获操作及其设置的基础。用 AVICAP 窗口类实现实时视频捕获的一般流程:(1)创建视频捕获窗口;(2)注册系统回调函数;(3)获取有关视频捕获窗口的缺省设置;(4)设置捕获窗口相关参数;(5)与视频采集设备连接;(6)设置捕获窗口显示模式;(7)捕获视频到缓存或文件并进行相应处理;(8)终止视频捕获并断开与视频采集设备的连接。

## 2.2 人脸图像的 Gabor 变换

Gabor 变换最早是由英国物理学家 Gabor 在 1946 年提出来的<sup>[3]</sup>。Gabor 函数与哺乳动物的视觉感受相当吻合,这一点对于研究图像特征检测或空间频率滤波非常有用。Gabor 滤波器是窄带通滤波器,有明显的方向选择和频率选择特性,而这些特性对图像纹理分析特别有效。对于任何一幅图像,通过一组完备且正交的 Gabor 滤波器滤波后都可以由其系数完全地重建。但是 Gabor 变换得到的不同尺度和纹理信息有选择的 Gabor 函数参数决定,恰当地选择参数,Gabor 变换可以出色地应用于图像分割、识别和理解。Gabor 滤波器可以看作被高斯包络调制的谐波函数,Gabor 函数二维形式表示为<sup>[4,5]</sup>:

$$g(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x'^2 + y'^2}{2\sigma^2}\right) \times \exp\{j[u(x-x_0) + v(y-y_0)]\} \quad (1)$$

式中,  $u = \omega_k \cos\phi$ ,  $v = \omega_k \sin\phi$ ,  $\omega_k$ ,  $\phi$  分别表示调制的频率和方向。

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} \quad (2)$$

其中,  $(x, y)$  为图像的坐标,  $(x_0, y_0)$  是滤波器在图像中的位置,  $\theta$  是预先设定的方向参数(通常与  $\phi$  相等),  $u, v$  分别表示  $\omega_k$  在  $x$  和  $y$  方向上的频率成分,  $\sigma$  为高斯函数在  $x$  轴和  $y$  轴的标准偏差。  $\frac{1}{2\pi\sigma^2} \exp\left(-\frac{x'^2 + y'^2}{2\sigma^2}\right)$  为用来约束平面波的高斯包络函数。  $\exp\{j[u(x-x_0) + v(y-y_0)]\}$  为复数值平面波,其实部为余弦平面波,虚部为正弦平面波,而正弦平面波关于高斯窗口奇对称,在高斯包络函数的约束范围内,其积分为 0。为了消除图像的直流成分对 Gabor 变化的影响,在复数值平面波的实部减去  $\exp(-\sigma^2/2)$ ,使得 Gabor 变换不受图像灰度绝对值的影响,并且对图像的光照变化不敏感。

本文运用 Gabor 变换主要是提取人脸的局部特征。正面人脸、近似正面人脸的人脸轮廓近似为椭圆,双眼和嘴巴都是基本上在水平方向上,而鼻子位于竖直方向上。因此本文滤波器的波长采用 5 个中心频率和滤波器方向参数 8 个方向组成的 40 个 Gabor 滤波器  $k_j$ , 参数  $Kv$ (滤波器的波长)和  $\Phi_u$ (滤波器纹理的方向)的取值如下:

$$\begin{cases} k_v = 2^{-\frac{v+1}{2}} \pi & v=0, 1, \dots, 4 \\ \phi_u = \frac{u}{8} \pi & u=0, 1, \dots, 7 \end{cases}$$

## 2.3 动态人脸采集与训练

动态人脸采集是通过摄像头将每个员工的  $N$  幅正面或接近正面的照片进行训练并保存到人脸训练库中,该库有  $M$

类人脸(即  $M$  个不同的受考勤人员),每类人脸均有  $N$  个样本。对人脸训练库每个样本采用由 5 个中心频率和 8 方向的 Gabor 滤波器进行变换得到 40 幅不同频率尺度和方向的该人脸样本的纹理信息。算出人脸训练库中的总体平均纹理信息为  $\bar{x}g(m, n)$  和人脸训练库中纹理信息的方差  $Sg(m, n)$ 。DFRAS 系统采集每人的 10 幅有一定差别的人脸正面图像,保存到训练库。在对同一个人脸采集 10 个样本的过程中,如果后面采集的人脸与前面采集的相似度很高(设定为 0.95~1.00 之间),则我们认为此次采集的人脸已经在训练库中存在,不保存结果并继续采集,直到采集相似度较高的图像为止,然后保存到训练库中。最后在训练库中,同一个人的不同样本归为相同的类,不同的受考勤人员的样本属于不同的类。

## 2.4 人脸特征提取与 FLDA

特征提取的任务是求出一组对分类最有效的特征。线性鉴别分析是一个著名的特征提取和维数约减的方法<sup>[6~11]</sup>。Fisher 线性鉴别分析(FLDA)<sup>[7~8, 12~22]</sup>的实质是使样本在特征空间的类间离散度( $S_b$ )与类内离散度( $S_w$ )比例达到最大,并提供了一个取得低维分类数据的有效方法。类间离散度与类内离散度分别用矩阵  $S_b, S_w$  表示。则类间离散度与类内离散度之比为:

$$J_F(W) = \frac{\sum_{i=1}^m W_i^T S_b W_i}{\sum_{i=1}^m W_i^T S_w W_i} \quad (3)$$

式中  $W$  是变换矩阵,则最佳鉴别矢量集  $W^*$  为:  $W^* = \text{MAX}(J_F(W))$

Fisher 线性鉴别技术是经过投影变换使得同类模式分布更集中,而不同类模式相距更远,求取得最佳鉴别矢量集为对应矩阵  $S_w^{-1} S_b$  较大特征值得特征矢量,该矢量集之间不不正交,但样本在该矢量集上得投影线性无关。

设待识别人脸属于  $k$  个候选人脸覆盖范围,即有  $k$  个总体  $C_1, C_2, \dots, C_k$ , 各类总体的样本  $n_1, n_2, \dots, n_k$ , 令  $n = \sum_{i=1}^k n_i$ 。第  $i$  总体的第  $a$  个样本为  $x_a^{(i)} = (x_{a1}^{(i)}, x_{a2}^{(i)}, \dots, x_{ap}^{(i)})$ 。设要建立的判别函数为:

$$Y = w_1 x_1 + w_2 x_2 + \dots + w_p x_p \triangleq W^T X \quad (4)$$

其中  $W = (w_1, w_2, \dots, w_n)^T$ ,  $X = (x_1, x_2, \dots, x_p)$   $\bar{x}^{(i)}$  和  $S^{(i)}$  分别为总体  $C_i$  内样本的均值向量和样本协方差阵:

$$\bar{x}^{(i)} = \frac{1}{n_i} \sum_{j=1}^{n_i} x_j^{(i)} \quad (5)$$

$$S^{(i)} = \frac{1}{n_i} \sum_{j=1}^{n_i} (x_j^{(i)} - \bar{x}^{(i)})(x_j^{(i)} - \bar{x}^{(i)})^T \quad (6)$$

$k$  类总体均值向量  $\bar{x}$ :

$$\bar{x} = \frac{1}{k} \sum_{i=1}^k \bar{x}^{(i)} \quad (7)$$

根据随机变量线性组合的均值和方差的性质可知,  $Y$  在  $C_i$  上样本的均值向量、样本方差和  $k$  类总体均值分别为:

$$\bar{y}^{(i)} = W^T \bar{x}^{(i)} \quad (8)$$

$$\sigma_i^2 = W^T S^{(i)} W \quad (9)$$

$$\bar{y} = W^T \bar{x} \quad (10)$$

Fisher 准则就是要选取系数向量  $W$ , 使得:

$$\lambda = \frac{\sum_{i=1}^k n_i (\bar{y}^{(i)} - \bar{y})^2}{\sum_{i=1}^k (n_i - 1) \sigma_i^2} \quad (11)$$

达到最大,将式(8)、(9)、(10)分别代入式(11)可化为:

$$\lambda = \frac{W^T A W}{W^T E W} \quad (12)$$

其中  $E$  为组内离散阵,  $A$  为总体之间样本协方差阵。为求  $\lambda$

的最大值,根据极值存在的必要条件,

$$\frac{\partial \lambda}{\partial W} = 0 \Rightarrow AW = \lambda EW \quad (13)$$

$\lambda$  和  $W$  是  $A, E$  矩阵的广义特征根及其对应的特征向量。由于一般都要求加权协方差  $E$  是正定的,上述非零特征根的个数  $m$  不超过  $\min(k-1, p)$  又因为  $A$  非负定的,所以非零特征根必为正根,记为  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m > 0, \lambda_1, \lambda_2, \dots, \lambda_m$  相应的特征向量分别为  $c^{(1)}, c^{(2)}, \dots, c^{(m)}$ 。特征向量可构成  $m$  个判别函数:

$$y = [c^{(1)}, c^{(2)}, \dots, c^{(m)}]^T (x - \bar{x}) \quad (14)$$

将候选人脸第  $i$  总体的第  $a$  个样本  $x_a^{(i)} = (x_{a1}^{(i)}, x_{a2}^{(i)}, \dots, x_{ap}^{(i)})$ , 得到  $m$  维特征向量  $y_a^{(i)} = (y_{a1}^{(i)}, y_{a2}^{(i)}, \dots, y_{am}^{(i)})$ , 即为训练样本中第  $i$  个总体 ( $C_i$ ) 第  $a$  个样本的特征向量。

## 2.5 人脸识别

属于  $k$  个候选人脸覆盖范围的待识别人脸图像采用由 5 个中心频率和 8 方向的 Gabor 滤波器进行变换得到 40 幅不同频率尺度和方向的该人脸样本的纹理信息  $N_{gk}(m, n)$ , 将 40 幅该人脸样本的纹理信息取其均值作为该样本的纹理信息  $\bar{N}_g(m, n)$ :

$$\bar{N}_g(m, n) = \frac{1}{40} \sum_{k=1}^{40} N_{gk}(m, n) \quad (15)$$

对识别具有较大贡献的纹理信息为  $\tilde{N}_g(m, n)$ :

$$\tilde{N}_g(m, n) = \begin{cases} |\bar{N}_g(m, n) - \bar{x}_g(m, n)|, & |\bar{N}_g(m, n) - \bar{x}_g(m, n)| > S_g(m, n) \\ 0, & \text{Otherwise} \end{cases} \quad (16)$$

式中,  $\bar{x}_g(m, n)$  为人脸训练库中的总体平均纹理信息,  $S_g(m, n)$  为人脸训练库中纹理信息的方差。

人脸局部信息按  $\tilde{N}_g(m, n)$  从大到小的选取  $number$  个变换系数所对应的  $\tilde{N}_g(m, n)$  的值,按图像扫描顺序排列构成的向量为训练样本对识别贡献的局部信息  $N_x = (N_{x1}, N_{x2}, \dots, N_{x_{number}})$ , 局部信息  $N_x$  运用式(14)得到  $m$  维最佳分类特征向量  $N_y = (N_{y1}, N_{y2}, \dots, N_{ym})$ 。

得到待识别人脸的特征向量  $N_y$  以后,通过以下几个步骤来确定待识别人脸是不是属于训练库中的人脸:

(1) 利用欧氏距离公式,查找和待识别图像最接近的训练库中的人脸类,并设  $d1 = \min_{1 \leq i \leq k} \|N_y - y^{(i)}\|$ ; 同理,计算待识别图像和重建图像之间的距离,并设其值为  $d2$ 。(2) 计算每两个人脸类之间的距离并取其最大值的一半作为距离的阈值  $\xi_1$ , 该距离阈值表示两个人脸类之间的最大允许距离。(3) 计算变换后的训练库中的每个人脸类内的距离并取其最大值作为距离阈值  $\xi_2$ , 该距离阈值表示属于某个人脸类的最大允许距离。(4) 考勤系统识别分类: ① 如果  $d2 \geq \xi_1$ , 则表明待识别图像不是人脸图像; ②  $d2 < \xi_1$  且  $d1 < \xi_2$ , 则待识别图像归类为训练库中的  $C_i$  类人脸; ③  $d2 < \xi_1$  且  $d1 \geq \xi_2$ , 则待识别图像不是训练库中的人脸, 则考勤系统给出提示, 考勤失败。

## 3 实验结果及分析

该系统的开发环境为: CPU 为 Intel C4 2.4GHz, 内存条为 DDR 512MB, 硬盘为 80G, Panasonic 摄像头, PV980m 采集卡; 客户端 Windows XP professional 2002 SP2, VC++6.0 和 Matlab7.0; 服务器端 Windows Sever 2003 + SQL Server2000。

由四川大学天智智能研究所生物特征识别小组开发的动态人脸考勤系统采集了实验室 50 位博、硕士生的人脸图像, 且每位的 10 幅人脸图像均有一定的差别, 图 1 给出了人脸库

中的三位被考勤人员通过 DFRAS 系统采集的图像, 其大小为  $180 \times 240$  pixel。实验结果表明基于 Gabor 滤波器和 FLDA 的人脸识别技术, 其一次性识别率达到为 96%, 主要原因在于: (1) 人脸训练库比较小、人员固定; (2) 要求考勤时人脸正面或近似正面对准摄像头; (3) 通常情况下用户在考勤时候姿态、表情都是正常情况, 没有太大变化, 也没有复杂背景及夸张性的化妆等情况, 所以识别率很高。

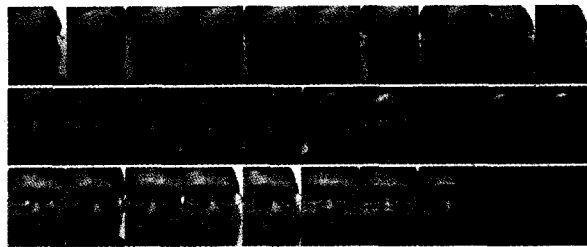


图 1 DFRAS 系统构建的人脸库部分图像示意图

**结论** 本文介绍了一个基于 Gabor 和 Fisher 线性鉴别分析 (FLDA) 的动态人脸考勤系统实现, 描述了系统实现的关键技术。Gabr 变换同时在时域和频域获得最佳的局部化, 并通过合理设置参数, 从而有效提取人脸的局部特征, 并在一定程度上消除了光照对人脸图像的影响。使用 FDLA 从高维特征空间中提取出最具有分类能力的低维特征, 隐含地提取最有利于分类的最佳鉴别特征。系统运行结果表明, 动态人脸识别考勤系统具有人脸识别动态性、识别率高、实用性好、可靠性强等特点。我们下一步的工作是进一步优化人脸特征的选择和提取算法, 提高识别的速度和系统的性能。

## 参考文献

- 张婷, 朱秋煜, 朱军林. 人脸识别考勤系统中有效人脸特征提取 [J]. 上海大学学报 (自然科学版), 2006, 12(3): 244~248
- Microsoft Corporation. Visual Studio MSDN, 2005
- Gabor D. Theory of communication [J]. Journal of the Institute of Electrical Engineers, 1946, 93(26): 429~457
- Daugman J. Uncertainty Relation for Resolution in Space, Spatial Frequency and Orientation Optimized by Two Dimensional Visual Cortical Filters. Journal of the Optical Society of America, 1985, 2(7): 1160~1169
- Daugman J G. Complete discrete 2D Gabor transforms by neural networks for image analysis and compression. IEEE Trans. On Acoustics, speech, and Signal Processing, 1988, 36(7): 1169~1179
- Zhao W, Nandhakumar N. Linear Discriminant Analysis of MPF for Face Recognition [C]. In: Proceedings of Fourteenth International Conference on Pattern Recognition, 1998, (1): 185~188
- Liu Chengjun, Wechsler H. Enhanced Fisher linear discriminant models for face recognition [C]. In: Proceedings. Fourteenth International Conference on Pattern Recognition, 1998 (2): 1368 ~ 1372
- Liu Chengjun, Wechsler H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition [J]. IEEE Transactions on Image Processing, 2002, 11(4): 467 ~ 476
- Lu J, Plataniotis K N, Venetsanopoulos A N. Boosting linear discriminant analysis for face recognition [C]. In: Proceedings. International Conference on Image Processing, 2003(1): I- 657~660
- Dai Guang, Qian Yuntao. A Gabor direct fractional-step LDA algorithm for face recognition [C]. In: IEEE International Conference on Multimedia and Expo, 2004(1): 61~64
- Kim T-K, Kittler J. Locally Linear Discriminant Analysis for Multimodally Distributed Classes for Face Recognition with a Single Model Image [J]. IEEE Trans. On PAMI, 2005, 27(3): 318~327
- Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711 ~ 720
- Li W S, Zhou C L. A novel combined fisherfaces framework [C]. In: 2004 International Conference on Image Processing (ICIP '04), 2004 (3): 2011 ~ 2014
- Li M, Yuan B Z. A novel statistical linear discriminant analysis for

image matrix; two-dimensional Fisherfaces. In: ICSP'04 Proceedings, 2004. 1419~1422

15 Etemad K, Chellapa R. Face recognition using discriminant eigenvectors [C]. In: Proc. of ICASSP, 1996

16 Turk M A, Pentland A P. Face recognition using eigenfaces [C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Los Acanitos, 1991. 586~591

17 Pentland A P, Moghaddam B, Starner T. View-based and modular eigenspaces for face recognition [C]. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, 1994. 84~91

18 Blanz V, Vetter T. Face Recognition Based on Fitting a 3D Morphable Model [J]. IEEE Transactions on PAMI, 2003, 25(9): 1063 ~ 1074

19 Pablo N, Javier R. Analysis and comparison of eigenspace-based face recognition approaches [J]. International Journal of Pattern Recognition and Artificial Intelligence, 2002, 16(7): 817~830

20 Chen Song-can, Liu Jun, Zhou Zhi-hua. Making FLDA applicable

to face recognition with one sample per person [J]. Pattern Recognition, 2004, 37(7): 1553~1555

21 Monwar M, Pau P P, Islam W, et al. A Real-Time Face Recognition Approach from Video Sequence using Skin Color Model and Eigenface Method [C]. In: Canadian Conference on Electrical and Computer Engineering, 2006. 2181 ~ 2185

22 Wang Jie, Plataniotis K N, Venetsanopoulos A N. Selecting discriminant eigenfaces for face recognition [J]. Pattern Recognition Letters. Available online 26 January 2005

23 周激流, 张晔, 郭晶, 等. 基于活动轮廓模型的人脸特征提取方法的研究[J]. 中国图象图形学报, 2000, 5(4): 341~344

24 王刚, 陈世福, 陈兆乾, 等. 基于分布式 workflow 技术的校园应用软件集成模型研究[J]. 计算机科学, 2005, 32(8): 94~96, 108

25 张生亮, 陈伏兵, 杨静宇. 对单训练样本的人脸识别问题的研究[J]. 计算机科学, 2006, 33(2): 225~229

26 周晓飞, 姜文瀚, 杨静宇. 11 范数最近邻凸包分类器在人脸识别中的应用[J]. 计算机科学, 2007, 34(4): 234~235, 238

(上接第 206 页)

该数据集具有 9 大类概念, 数据集中的每篇文档均隶属于若干的概念类别。统计每个类别所包含的文档数, 然后将包含文档数最多的  $l$  个类 ( $l=3, 4, \dots, 9$ ) 组合在一起, 由此可以得到 7 个不同的数据集, 分别记为 num3, num4, num5, num6, num7, num8, num9。对于每一数据集, 统计其中每个词汇的文档频率 (document frequency), 即包含该词汇的文档数目。根据每个词汇的统计结果, 取文档频率排在前 2% 的词汇构成最终的词汇表。数据集中的每篇文档即可采用 Bag-of-Word 的方法表示成为一个数值向量以供多标记学习算法处理。表 1 给出了数据集的描述信息。

表 1 实验数据集的描述信息

数据集	文本数	词汇表大小	每篇文档平均标记数
num3	7258	530	1.0074
num 4	8078	599	1.0140
num 5	8655	651	1.0207
num 6	8817	663	1.0352
num 7	9021	678	1.0375
num 8	9158	683	1.0396
num 9	9190	686	1.0480

### 3.2 实验结果

本文将  $\theta$ -MLkNN 算法与基于空间距离的 ML-kNN 算法进行了实验比较。对于每个数据集而言, 本文采用与文 [13] 相同的实验策略, 在每个数据集上进行 3 倍交叉运算 (3-fold cross-validation) 并给出平均结果。限于时间因素, 本文对两种算法都取近邻数为 10。

我们对实验结果重点考察标记的准确率, 采用以下公式:

$$avr\_pre = \frac{1}{n} \sum_{i=1}^n \frac{|l| |L_i \cap L'_i|}{|L_i|} \quad (6)$$

公式 (6) 中  $L_i$  为测试样本  $i$  的标记集合,  $L'_i$  为  $i$  经分类器标记后的标记集合。表 2 为我们对上述 Reuters21578 进行实验后的结果。

表 2  $\theta$ -MLkNN 与 ML-kNN 实验结果对比

Avr_pre	$\theta$ -MLkNN	ML-kNN
num3	0.904622	0.849195
num4	0.88479	0.837702
num5	0.881038	0.836508
num6	0.862281	0.825364
num7	0.852382	0.814089
num8	0.846534	0.799531
num9	0.837955	0.784393

由表 (2) 可以明显看出, 基于向量角度  $\theta$  的  $k$  近邻算法在对于文档分类这一多标记学习中的准确度要高于传统的基于 Euclidean 距离的  $k$  近邻算法。

**结语** 本文针对多标记文本学习提出了一种基于向量角度距离的  $k$  近邻算法, 该算法基于传统  $k$  近邻算法思想, 并在此基础上对近邻选取上以向量夹角来代替空间距离, 通过实验可以看出, 得到了较好的准确度。进一步的工作将考虑在标记数量更大的环境中对该算法进行测试, 并将其应用在其他多标记学习中, 与传统的多标记学习进行对比。

### 参考文献

1 Clare A, King R D. Knowledge discovery in multi-label phenotype data. In: De Raedt L, Siebes A, eds. Lecture Notes in Computer Science 2168, Berlin: Springer, 2001. 42~53

2 Comit  F D, Gilleron R, Tommasi M. Learning multi-label alternating decision trees from texts and data. In: Perner P, Rosenfeld A, eds. Lecture Notes in Computer Science 2734, Berlin: Springer, 2003. 35~49

3 Elisseeff A, Weston J. A kernel method for multi-labelled classification. In: Dietterich T G, Becker S, Ghahramani Z, eds. Advances in Neural Information Processing Systems 14, Cambridge, MA: MIT Press, 2002. 681~687

4 Dietterich T G. Ensemble learning. In: Arbib M A, ed. The Handbook of Brain Theory and Neural Networks, 2nd edition. Cambridge, MA: MIT Press, 2002

5 Zhou Z-H, Chen S-F. Neural network ensemble. Journal of Computers, 2003, 25(1): 1~8

6 Schapire R E, Singer Y. BoosTexter: a boosting-based system for text categorization. Machine Learning, 2000, 39(2-3): 135~168

7 Freund Y, Schapire R E. A decision-theoretic generalization of on-line learning and an application to boosting. In: Vit nyi P M B, ed. Lecture Notes in Computer Science 904, Berlin: Springer, 1995. 23~37

8 Dempster A P, Laird N M, Rubin D B. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistics Society - B, 39(1): 1~38

9 McCallum A. Multi-label text classification with a mixture model trained by EM. In: Working Notes of the AAAI'99 Workshop on Text Learning, Orlando, FL, 1999

10 Ueda N, Saito K. Parametric mixture models for multi-label text. In: Becker S, Thrun S, Obermayer K, eds. Advances in Neural Information Processing Systems 15, Cambridge, MA: MIT Press, 2003. 721~728

11 Freund Y, Mason L. The alternating decision tree learning algorithm. In: Proceedings of the 16th International Conference on Machine Learning, Bled, Slovenia, 1999. 124~133

12 Lauser B, Hotho A. Automatic multi-label subject indexing in a multilingual environment. In: Koch T, S lvberg J, eds. Lecture Notes in Computer Science 2769, Berlin: Springer, 2003. 140~151

13 Zhang M L, Zhou Z H. ML-kNN: A lazy learning approach to multi-label learning. Pattern Recognition (PRJ), 2007, 40(7): 2038~2048