

一种核最大散度差判别分析人脸识别方法

杜海顺^{1,2} 李玉玲¹ 汪凤泉² 张帆¹

(河南大学先进控制与智能信息处理研究所 开封 475004)¹

(东南大学仪器科学与工程学院 南京 210096)²

摘要 提出一种有效的非线性子空间学习方法——核最大散度差判别分析(KMSD),并将其用于人脸识别。核最大散度差判别分析首先把输入空间的样本非线性映射到特征空间,然后通过核方法的技巧,采用最大散度差判别分析(MSD)方法在特征空间里求解。在 Yale 和 ORL 人脸数据库上的实验结果表明,提出的核最大散度差判别分析方法用于人脸识别具有较高的识别率。

关键词 核最大散度差判别分析,子空间学习,人脸识别

中图分类号 TP391.41 **文献标识码** A

Face Recognition Using Kernel Maximum Scatter Difference Discriminant Analysis

DU Hai-shun^{1,2} LI Yu-ling¹ WANG Feng-quan² ZHANG Fan¹

(Institute of Advanced Control and Intelligent Information Processing, Henan University, Kaifeng 475004, China)¹

(College of Instrument Science & Engineering, Southeast University, Nanjing 210096, China)²

Abstract An efficient nonlinear subspace learning method, kernel maximum scatter difference discriminant analysis (KMSD), was proposed for face recognition in this paper. The main idea of KMSD is to map the input sample data into feature space by nonlinear function, and then adopt maximum scatter difference discriminant analysis (MSD) to find the solution in feature space by kernel trick. The experimental results on the Yale and ORL face image database show that the proposed KMSD method for face recognition has higher recognition rate and more effective.

Keywords Kernel maximum scatter difference discriminant analysis (KMSD), Subspace learning, Face recognition

1 引言

基于 Fisher 准则函数的线性判别分析(LDA)^[1]是一种线性子空间学习方法。它作为特征提取和数据描述的有力工具,在模式识别领域得到了广泛的应用和扩展^[2,3]。但是在应用 LDA 处理高维多类模式如人脸图像时,样本的维数很高且训练样本数很少,这就是所谓的“小样本问题”。小样本问题往往会造成 Fisher 线性判别准则由于类内散度矩阵奇异而无法直接求解最优投影轴,即此时的 Fisher 判别准则是“病态的”。这就使得多类高维小样本情况下,直接应用 LDA 是非常困难的。目前,“小样本问题”已成为 LDA 中公认的难题之一,引起了人们广泛的研究兴趣。近年来对其做了一系列的研究和改进^[4,5]。但这些方法大都是在原始的 Fisher 判别准则下所作的改进,该准则由于本身存在的缺陷,因此并没有从根本上解决这一问题。

最近,宋枫溪等^[6]提出了一种最大散度差判别分析(MSD)。与 LDA 不同的是,它是利用类间散度矩阵与类内散度矩阵之差作为线性判别分析准则。这样,从理论上根本消除了 Fisher 判别准则中存在的因类内散度矩阵奇异而无

法求解的问题。随着核函数在支持向量机中的成功应用,关于核方法的研究受到了重视,出现了核主成分分析(KPCA)^[7]等一些非线性子空间学习方法。本文提出一种核最大散度差判别分析(KMSD)方法,并将其用于人脸识别。该方法首先把输入空间的样本非线性映射到特征空间,然后通过核方法的技巧采用 MSD 在特征空间里求解。实验结果表明,该方法具有较高的识别率。

2 最大散度差判别分析(MSD)

令样本数据集为 $X = (x_1, x_2, \dots, x_N)$, 其中 $x_i \in R^D, i = 1, 2, \dots, N$ 。设样本数据可分为 C 类,用 X_i 表示第 i 类样本数据集。第 i 类样本数据个数为 N_i 且满足 $\sum_{i=1}^C N_i = N$ 。MSD 的目的是寻找最优的投影向量 w ,使得下式定义的散度差准则函数最大:

$$J_{\Delta}(w) = w^T (S_B - S_W) w \quad (1)$$

式中, T 表示转置(下同), S_B 和 S_W 分别表示训练样本图像的类间散度和类内散度矩阵,定义如下:

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

到稿日期:2009-07-24 返修日期:2009-10-14 本文受河南省自然科学研究资助计划项目(2008A520003)资助。

杜海顺(1977-),男,博士,讲师,主要研究方向为图像处理、模式识别等,E-mail:jddhs@henu.edu.cn;李玉玲(1973-),女,博士,副教授,主要研究方向为数据挖掘与知识发现;汪凤泉(1940-),男,教授,主要研究方向为动态测试技术与仪器、模式识别等;张帆(1967-),男,博士,副教授,主要研究方向为图像处理、模式识别等。

$$S_W = \sum_{i=1}^C \sum_{x_j \in X_i} (x_j - \mu_i)(x_j - \mu_i)^T \quad (3)$$

式(2)和式(3)中, μ_i 和 μ 分别表示第 i 类训练样本数据集和总的训练样本数据集的均值,即:

$$\mu_i = \frac{1}{N_i} \sum_{x_j \in X_i} x_j, i=1,2,\dots,C; \mu = \frac{1}{N} \sum_{x_j \in X} x_j \quad (4)$$

如前所述,MSD的目的就是寻找使 $J_d(w)$ 最大的最优投影方向 w 。事实上,该最优投影方向为 $S_B - S_W$ 的最大特征值对应的归一化特征向量。一般来说,在样本类别数较多的情况下,单一的最优投影方向是不够的。因此,需要寻找一组满足标准正交条件且极大化散度差准则函数式(1)的最优投影向量组 w_1, w_2, \dots, w_d 。实际上, w_1, w_2, \dots, w_d 是 $S_B - S_W$ 的前 d 个最大特征值对应的特征向量。令 $W = [w_1, w_2, \dots, w_d]$, 称 W 为最优投影矩阵。对任一给定的数据点 $x \in R^D$, 向 W 投影就可以得到维特征向量 $y \in R^d$: $y = W^T x$ 。

3 核最大散度差判别分析(KMSD)

根据式(2)和式(3),类内散度矩阵 S_W 和类间散度矩阵 S_B 可以写成如下形式^[8]:

$$S_W = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N R_W(i, j)(x_i - x_j)(x_i - x_j)^T \quad (5)$$

$$S_B = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N R_B(i, j)(x_i - x_j)(x_i - x_j)^T \quad (6)$$

式(5)和式(6)中, R_W 为类内散度权重矩阵, R_B 为类间散度权重矩阵,分别定义如下:

$$R_W(i, j) = \begin{cases} 1/N_k, & x_i \in X_k \text{ 且 } x_j \in X_k \\ 0, & \text{其他} \end{cases} \quad (7)$$

$$R_B(i, j) = \begin{cases} 1/N - 1/N_k, & x_i \in X_k \text{ 且 } x_j \in X_k \\ 1/N, & \text{其他} \end{cases} \quad (8)$$

给定一非线性函数 ϕ , 输入向量 x 可以映射到特征空间 F :

$$\phi: R^D \rightarrow F, x \mapsto \phi(x) \quad (9)$$

通常非线性变换函数 $\phi(\cdot)$ 相当复杂,无法给出其明确的解析表达式。事实上,在实际应用中只需用核函数计算特征空间 F 中两个向量的内积即可,而核函数则相对简单得多。目前应用比较广泛的核函数有多项式函数、高斯径向函数和多层感知机函数。本文采用的是多项式函数,即:

$$k(x_i, x_j) = (x_i \cdot x_j)^d \quad (10)$$

式中, d 为函数的参数。

令 $\Phi = [\phi(x_1), \phi(x_2), \dots, \phi(x_N)]^T$, 内积矩阵 $K = \Phi\Phi^T$, 则 K 的元素为:

$$K_{ij} = (\phi(x_i) \cdot \phi(x_j)) = k(x_i, x_j) \quad (11)$$

式中, $k(x_i, x_j)$ 为核函数。本文记 $\phi(x_i)$ 为 φ_i , $l(\varphi_i) \in \{1, 2, \dots, C\}$ 为 φ_i 的类别标签,则定义特征空间 F 的类内散度权重矩阵 \tilde{R}_W 和类间散度权重矩阵 \tilde{R}_B 为:

$$\tilde{R}_W(i, j) = \begin{cases} 1/N_k, & l(\varphi_i) = l(\varphi_j) = k \\ 0, & \text{其他} \end{cases} \quad (12)$$

$$\tilde{R}_B(i, j) = \begin{cases} 1/N - 1/N_k, & l(\varphi_i) = l(\varphi_j) = k \\ 1/N, & \text{其他} \end{cases} \quad (13)$$

那么,特征空间 F 中的类内散度矩阵 \tilde{S}_W 和类间散度矩阵 \tilde{S}_B 为:

$$\tilde{S}_W = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \tilde{R}_W(i, j)(\varphi_i - \varphi_j)(\varphi_i - \varphi_j)^T$$

$$= \sum_{i=1}^N \varphi_i \left(\sum_{j=1}^N \tilde{R}_W(i, j) \right) \varphi_i^T - \sum_{i=1}^N \sum_{j=1}^N \varphi_i \tilde{R}_W(i, j) \varphi_j^T \\ = \Phi^T (D_W - \tilde{R}_W) \Phi \quad (14)$$

$$\tilde{S}_B = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \tilde{R}_B(i, j)(\varphi_i - \varphi_j)(\varphi_i - \varphi_j)^T \\ = \sum_{i=1}^N \varphi_i \left(\sum_{j=1}^N \tilde{R}_B(i, j) \right) \varphi_i^T - \sum_{i=1}^N \sum_{j=1}^N \varphi_i \tilde{R}_B(i, j) \varphi_j^T \\ = \Phi^T (D_B - \tilde{R}_B) \Phi \quad (15)$$

式(14)和式(15)中, D_W 和 D_B 为 $N \times N$ 维对角阵,且

$$D_W(i, i) = \sum_{j=1}^N \tilde{R}_W(i, j) \quad (16)$$

$$D_B(i, i) = \sum_{j=1}^N \tilde{R}_B(i, j) \quad (17)$$

在特征空间 F 中采用 \tilde{S}_W 和 \tilde{S}_B 进行最大散度差判别分析,即求解一组最优的投影向量 $W = [w_1, w_2, \dots, w_d]$, 使得下式定义的散度差准则函数最大:

$$\tilde{J}_d(w) = w^T (\tilde{S}_B - \tilde{S}_W) w \quad (18)$$

由于特征空间 F 中的任意向量 w 可以表示成 $\varphi_1, \varphi_2, \dots, \varphi_N$ 的线性组合,即:

$$w = \sum_{i=1}^N \alpha_i \varphi_i = \Phi^T \alpha \quad (19)$$

式中, $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_N]^T$, 则:

$$w^T \tilde{S}_W w = \alpha^T \Phi \tilde{S}_W \Phi^T \alpha = \alpha^T P_W \alpha \quad (20)$$

$$w^T \tilde{S}_B w = \alpha^T \Phi \tilde{S}_B \Phi^T \alpha = \alpha^T P_B \alpha \quad (21)$$

式中,

$$P_W = \Phi \tilde{S}_W \Phi^T = K(D_W - \tilde{R}_W)K \quad (22)$$

$$P_B = \Phi \tilde{S}_B \Phi^T = K(D_B - \tilde{R}_B)K \quad (23)$$

则使得式(18)最大化的解为:

$$\alpha = \underset{\alpha}{\operatorname{argmax}} (\alpha^T (P_B - P_W) \alpha) \quad (24)$$

由于矩阵 $P_B - P_W$ 是对称和半正定的,通过简单的代数运算,式(24)可以很容易地转换成下面的广义特征值求解问题:

$$(P_B - P_W) \alpha = \lambda \alpha \quad (25)$$

令 $A = [\alpha_1, \alpha_2, \dots, \alpha_d]$ 是 $P_B - P_W$ 的前 d 个最大特征值对应的特征向量。

对任一输入向量 $x \in R^D$ 经过非线性变换,向由 $W = [w_1, w_2, \dots, w_d]$ 张成的特征空间投影可提取特征向量 $y \in R^d$ ($d \leq D$):

$$y = W^T \phi(x) = [w_1, w_2, \dots, w_d]^T \phi(x) = [\alpha_1, \alpha_2, \dots, \alpha_d]^T \Phi \phi(x) = A^T Q \quad (26)$$

式中, Q 为 $N \times 1$ 的列向量,其元素由 $Q(i) = (\phi(x_i) \cdot \phi(x)) = k(x_i, x)$, $i=1, 2, \dots, N$ 定义。

4 实验与分析

本文采用两个国际著名的人脸数据库(Yale库、ORL库)来测试所提出的 KMSD 算法的性能。实验中比较了 KMSD 与 MSD, PCA+LDA, KPCA 的识别率。值得指出的是,4 种人脸识别方法均采用最近邻分类器完成最终分类。

4.1 Yale 人脸数据库上的实验

Yale 人脸数据库^[9]由 15 人、每人 11 幅共 165 幅人脸图像组成。每幅图像灰度级为 256,分辨率为 243×320 。Yale 数据库中人脸图像光照条件变化较大,并且人的脸部表情和细节也有一定的变化。为了减少背景成分对识别性能的影响,对每幅图像以人脸为中心提取尺寸为 190×170 的子图像,如图 1 所示。



图1 Yale人脸数据库部分图像

实验中,从Yale人脸库中每人随机选取 $k(k=3,4,5,6,7,8)$ 幅图像构成训练样本集,其余图像构成测试样本集。分别采用MSD,PCA+LDA,KPCA和本文提出的KMSD提取特征向量,用最近邻分类器进行分类识别,比较它们在不同训练样本数下的识别率。对任一训练样本数 k ,为了消除单次选择样本的随机性,独立重复实验10次,最后取平均识别率作为最终识别率。实验结果如图2和表1所示。

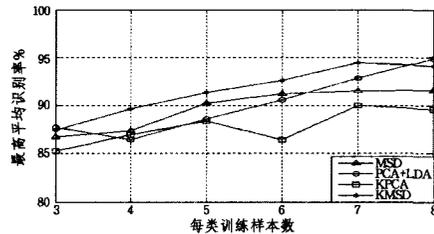


图2 Yale人脸数据库上的实验结果

表1 每类训练样本数不同时4种人脸识别方法在Yale人脸数据库上的最高平均识别率(%)比较

识别方法	每类训练样本数					
	3	4	5	6	7	8
MSD	86.67 (20)	87.33 (24)	90.22 (26)	91.20 (60)	91.50 (18)	91.56 (14)
PCA+LDA	87.67 (14)	86.38 (14)	88.56 (14)	90.53 (14)	92.83 (14)	94.89 (14)
KPCA	85.25 (32)	86.95 (58)	88.33 (34)	86.40 (36)	90.00 (78)	89.56 (66)
KMSD	87.38 (32)	89.62 (56)	91.33 (42)	92.60 (36)	94.50 (76)	94.11 (96)

注:括号中的数字代表每一训练样本数下最高平均识别率所对应的特征向量维数。

4.2 ORL人脸数据库上的实验

ORL人脸数据库(<http://www.cam-orl.co.uk>)是由英国剑桥大学 Olivetti 研究所制作的。该数据库由40人、每人10幅共400幅人脸图像组成。每幅图像为256个灰度级,分辨率为 112×92 。ORL人脸图像是在不同时间、不同视角、各种表情(闭眼、睁眼、微笑、吃惊、生气、愤怒、高兴等)和不同脸部细节(戴眼镜/不戴眼镜、有胡子/没胡子、不同发型)的条件下拍摄的。数据库中的部分人脸图像如图3所示。

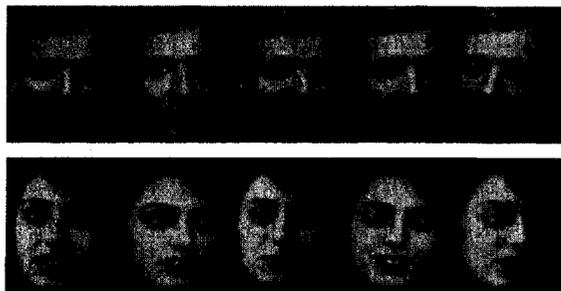


图3 ORL人脸数据库部分人脸图像

实验中,从ORL人脸数据库中每人随机选取5幅人脸图像作为训练样本,其余图像作为测试样本。分别采用MSD,PCA+LDA,KPCA和本文提出的KMSD提取特征向量,用最近邻分类器进行分类识别,比较它们在不同特征向量维数下的识别率。同样为了消除单次选择样本的随机性,对任一特征向量维数,独立重复实验20次,最后取平均识别率作为最终识别率。实验结果如图4和表2所示。

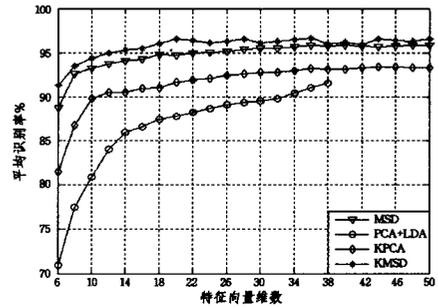


图4 ORL人脸数据库上的实验结果

表2 每类训练样本数为5时4种识别方法在ORL人脸数据库上的性能比较

特征向量维数	识别方法			
	MSD	PCA+LDA	KPCA	KMSD
特征向量维数	50	38	44	36
最高平均识别率(%)	95.82	91.55	93.35	96.60

图2和表1给出了Yale人脸库上,每类训练样本数不同的情况下,4种人脸识别方法的最高平均识别率的比较;图4给出了ORL人脸数据库上,每类训练样本数为5、特征向量维数不同的情况下,4种人脸识别方法的平均识别率的比较;表2给出了ORL人脸数据库上,每类训练样本数为5的情况下,4种人脸识别方法的最高平均识别率及其对应的特征向量维数的比较。由图2和图4、表1和表2可以看出,本文的KMSD人脸识别方法识别率优于其他3种方法,说明所提出的KMSD方法是有效的。

结束语 随着核函数在支持向量机中的成功应用,关于核方法的研究受到了重视,出现了一些非线性子空间学习方法。受KPCA启发,本文在最大散度鉴别分析(MSD)算法的基础上,采用核方法的技巧推导出了核最大散度鉴别分析(KMSD)算法,并将其用于人脸识别。通过Yale和ORL人脸数据库上的实验结果表明,本文提出的KMSD与MSD以及PCA+LDA,KPCA等其他特征提取方法相比,具有更高的人脸识别率。值得指出的是,KMSD不仅适用于人脸识别,还可以推广应用到其他模式识别领域。

参考文献

- [1] Belhumeur P, Hespanha J, Kriegman D. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720
- [2] Tenenbaum J B, Silva V D, Langford J C. A global geometric framework for nonlinear dimensionality reduction [J]. Science, 2000, 290(12): 2319-2322
- [3] 祝磊,马莉,厉力华.一种基于GDLPP的人脸识别算法[J].光电工程,2008,36(6):108-112

- sign for High Performance Memory Systems[A]//IEEE Proceedings of the Second International Symposium on High-Performance Computer Architecture[C]. 1996;254-263
- [3] Lai A, Fide C, Falsafi B. Dead-Block Prediction and Dead-Block Correlating Prefetchers[A]//Proceedings of the 28th International Symposium on Computer Architecture[C]. 2001;144-154
- [4] Schumann R C. Design of the 2 1 1 7 4 Memory Controller for Digital Personal Workstations[J]. Digital Technical Journal, 1997, 9(2);57-69

- [5] Zhao Z, Zhichun Z, Zhang X. A permutation-Based Page Interleaving Scheme to Reduce Row-Buffer Conflicts and Exploit Data Locality[A]//Proceedings of the 33rd Annual International Symposium on Microarchitecture[C]. 2000;32-41
- [6] Burger D C, Austin T M. The SimpleScalar Tool Set [R]. CS-TR-97-1342. 1997
- [7] Samsung Semiconductor. 512 Mb B-die DDR400 SDRAM Specification[R]. Revision 1. 0. 2003

(上接第 285 页)

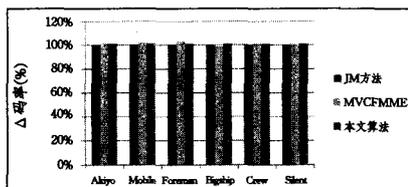


图 7 码率变化比较

速度的对比通过表 1 给出。表 1 中列出了在 16 和 28 两种量化参数下的实验结果,分别对应高码率和低码率两种情况。无论在低码率还是高码率,本文的算法对于全参考帧搜索方法(即 JM 中使用的算法)有至少一倍的提高,而对于快速算法 MVCMMME,也有平均 20% 的减少。在 Akiyo, Silent 和 Crew 这些运动不甚剧烈的序列中,本文算法能提供更快的速度。其原因主要是较静止的序列内周期运动和分数运动出现得较少,本文的算法可以跳过绝大多数对冗余参考帧的搜索。

表 1 本文算法与 H. 264 JM 快速运动估计算法的比较结果

视频序列	量化值	格式	Δ 复杂度(%) 与 JM 方法	Δ 复杂度(%) 与 MVCMMME
Akiyo	16	CIF	-56.31	-27.14
Mobile	16	4CIF	-55.27	-24.72
Foreman	16	4CIF	-45.54	-25.6
Bigship	16	720p	-53.44	-22.48
Crew	16	720p	-58.18	-21.98
Silent	16	CIF	-57.89	-28.15
Akiyo	28	CIF	-51.11	-26.3
Mobile	28	4CIF	-49.2	-25.02
Foreman	28	4CIF	-55.41	-25.03
Bigship	28	720p	-49.89	-23.57
Crew	28	720p	-50.82	-28.8
Silent	28	CIF	-52.9	-27.9

结束语 本文提出了一种联合分数运动估计的多参考帧快速选择算法。首先,通过对多参考帧选择技术和分数运动估计技术的分析发掘了两者之间的相关性;然后,利用合成的运动向量和分数残差曲面模型,确定进行分数运动估计的参考帧。该方法仅在时间相邻、运动相邻和分数预测残差最小

的参考帧上进行分数运动估计,在大幅地加快多参考帧运动估计速度的同时,保持了视频编码的率失真性能。

参考文献

- [1] Shen Liquan, Liu Zhi, Zhang Zhaoyang, et al. An Adaptive and Fast H. 264 Multi-Frame Selection Algorithm Based on Information from Previous Searches [C]//2007 IEEE International Conference on Multimedia and Expo. 2007
- [2] Sun Qichao, Chen Xin-hao, Wu Xiaoyang, et al. A Content-adaptive Fast Multiple Reference Frames Motion Estimation in H. 264 [C]//2007 IEEE International Symposium on Circuits and Systems (ISCAS). 2007
- [3] Liu Zhenyu, Li Lingfeng, Song Yang, et al. Motion Feature and Hadamard Coefficient-Based Fast Multiple Reference Frame Motion Estimation for H. 264 [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(5): 620-632
- [4] Chen Mei-juan, Li Gwo-long, Chiang Yi-yen, et al. Fast multi-frame motion estimation algorithms by motion vector composition for the MPEG-4/AVC/H. 264 standard [J]. IEEE Transactions on Multimedia, 2006, 8(3): 478-487
- [5] Chen Zhibo, Xu Jianfeng, He Yun, et al. Fast integer-pel and fractional-pel motion estimation for H. 264/AVC [J]. Journal of Visual Communication and Image Representation, 2006, 17(2): 264-290
- [6] Yi Xiaoquan, Zhang Jun, Ling Nam, et al. Improved and simplified fast motion estimation for JM, JVT-P021. doc [C]//Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q. 6) 16th Meeting. Poznan, Poland, 2005
- [7] Alexis T, Pankaj T. Fast Subpixel Motion Estimation Support for the Enhanced Predictive Zonal Search Scheme, JVT-Q079 [C]//Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q. 6) 17th Meeting. Nice, France, 2005
- [8] 陆寄远, 张培钊, 朝红阳. 一种面向 H. 264 KTA 的快速分数运动估计方法 [J]. 中国图象图形学报, 2010, 15(3): 367-371

(上接第 288 页)

- [4] Chen L F, Liao H Y, Ko M T, et al. A new LDA-based face recognition system which can solve the small sample size problem [J]. Pattern Recognition, 2000, 33(10): 1713-1726
- [5] Zhuang X S, Dai D Q. Improved discriminant analysis for high dimension data and its application to face recognition [J]. Pattern Recognition, 2007, 40(5): 1570-1578
- [6] 宋枫溪, 程科, 杨静宇. 最大散度差和大间距线性投影与支持向量 [J]. 自动化学报, 2004, 30(6): 890-896
- [7] Yang M H. Kernel Eigenfaces vs. Kernel Fisherfaces face reco-

- gnition using kernel methods [A]//Proceedings of Fifth IEEE International conference on Automatic Face and Gesture Recognition [C]. Washington DC, USA, 2002; 215-220
- [8] Hu H F. Orthogonal neighborhood preserving discriminant analysis for face recognition [J]. Pattern Recognition, 2008, 41(9): 2045-2054
- [9] Georgiades A S, Belhumeur P, Kriegman D, et al. From Few to Many: Illumination Cone Models for Face Recognition Under Variable Lighting and Pose [J]. IEEE Trans. Pattern Anal. Mach. Intelligence, 2001, 23(6): 643-660