

基于 Log-Gabor 小波和二维半监督判别 分析的人脸图像检索^{*}

黄荣兵¹, 郎方年¹, 施 展²

(1. 成都大学 四川省高校模式识别与智能信息处理重点实验室, 成都 610106; 2. 华南理工大学 电子与信息学院, 广州 510640)

摘 要: 针对光照、表情变化给人脸识别造成的影响以及大型人脸图像库的训练样本中只有部分标记的问题, 结合多通道 Log-Gabor 小波和半监督流形学习算法, 提出一种新的人脸图像检索方法。该方法首先使用 Log-Gabor 小波对人脸图像进行滤波获得特征矩阵, 进一步利用提出的二维半监督流形学习算法进行维数约简, 得到低维判别特征。由于该方法直接作用于 Log-Gabor 特征矩阵, 克服了小样本带来的奇异问题; 另外, 通过充分利用标记和未标记信息, 还保留了数据的局部流形结构, 增强了特征匹配的相似性。在 CMU PIE 和 AR 人脸数据库上的实验结果表明, 该方法有效且优于其他方法。

关键词: Log-Gabor 小波; 人脸识别; 图像检索; 半监督学习; 流形学习

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1001-3695(2012)01-0393-04

doi:10.3969/j.issn.1001-3695.2012.01.109

Log-Gabor and 2D semi-supervised discriminant analysis based face image retrieval

HUANG Rong-bing¹, LANG Fang-nian¹, SHI Zhan²

(1. Key Laboratory of PR & IIP, Chengdu University, Chengdu 610106, China; 2. College of Electronic & Information Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract: Face recognition methods are easily affected by variances of lighting and expression, and there are often only part of sample faces labeled in large image database. To address the two problems, this paper presented a face image retrieval method using multi-channel Log-Gabor wavelet based semi-supervised manifold learning algorithm. Firstly, used Log-Gabor to filter the face image to obtain feature matrix. Then proposed two-dimension semi-supervised manifold learning algorithm to extract the discriminative submanifolds. As the proposed method operated on Log-Gabor feature matrix directly, small sample problem was overcome. Furthermore, by making use of labeled and unlabeled information, preserved the local manifold structure of image data. So enhanced the similarity of feature matching. Experimental results on CMU PIE and AR databases show that the proposed approach is effective and superior to others.

Key words: Log-Gabor wavelet; face recognition; image retrieval; semi-supervised learning; manifold learning

纹理作为一种图像的低层特征,在图像检索中具有重要作用。心理生物学的研究表明,人往往对在视网膜上形成的像进行多通道、多频率和多方向分析。Daugman^[1]的研究表明, Gabor 小波与人类视觉系统中简单细胞的视觉刺激响应非常相似,它在提取目标的局部空间和频率信息方面具有良好的特性,因此被广泛应用于图像的特征提取^[2]。

将 Gabor 小波用于人脸的特征提取也得到了广泛的应用^[3,4]。但它存在直流分量(direct current component, DC)及过度表达低频分量而对高频分量表达不足的缺点,影响了它对图像特征的表达^[5]。为此,本文引入它的变种即 Log-Gabor 小波来克服 Gabor 小波的缺陷,而且 Log-Gabor 小波比 Gabor 小波更符合人类的视觉特性^[6]。针对大型人脸图像库的特征高维问题,研究者引入了流形学习算法,以便能够充分利用未标

记和已标记数据信息,如流形排序^[7]、半监督判别分析^[8]、局部保持投影^[9]等方法。但是,如果将 Cai 等人^[8]提出的半监督判别分析方法直接应用于 Log-Gabor 特征形成的列向量,会导致奇异问题。为解决此问题,近年来,研究者又提出了一些直接处理矩阵数据的方法,如 2D-PCA^[10]、2D-LPP^[11]等。基于 Song 等人^[12]提出的半监督判别分析(semi-supervised discriminant analysis, SSDA)框架,本文对 SSDA 进行扩展,提出了 2D-SSDA 直接对 Log-Gabor 矩阵进行特征选择的人脸图像检索方法。

1 Log-Gabor 小波及图像特征提取

1.1 Log-Gabor 小波

Log-Gabor 小波是一种在对数频率尺度上传递函数为高斯函数的滤波器,不受带宽的影响,可以构造出任意带宽和零 DC

收稿日期: 2011-05-09; **修回日期:** 2011-06-14 **基金项目:** 四川省杰出青年基金资助项目(2010JQ0032);成都大学校青年基金资助项目(2010XJZ07)

作者简介: 黄荣兵(1977-),男,四川隆昌人,副教授,博士,主要研究方向为模式识别、基于内容的图像检索(huangrb2006@126.com);郎方年(1974-),男,副教授,博士(后),主要研究方向为模式识别、智能算法;施展(1984-),男,博士,主要研究方向为图像处理、模式识别。

分量的滤波器。一维 Log-Gabor 函数的传递函数定义为^[13]

$$G(\omega) = \exp\left(-\frac{(\log(\omega/\omega_0))^2}{2(\log(\beta/\omega_0))^2}\right) \quad (1)$$

其中: ω_0 为滤波器的中心频率, 为了保证滤波器的形状恒定, 对于不同的 ω_0 应选择合适的 β 使 β/ω_0 保持不变。例如, 当 β/ω_0 取值 0.74 时, 大致相当于一倍频的带宽, 取值 0.55 时相当于两倍频的带宽, 取值 0.41 时相当于三倍频的带宽。

为了提取图像的特征, 需要将一维的 Log-Gabor 小波扩展到二维小波, 方法就是在一维函数垂直方向上用高斯函数平滑, 在极坐标上表示为在 θ 方向上的高斯平滑, 定义为

$$G(\theta) = \exp(-(\theta - \theta_0)^2/2\varphi_0) \quad (2)$$

其中: θ_0 是滤波器的角度; $\varphi_0 = \lambda\Delta\theta$, λ 为设定的系数, $\Delta\theta$ 为相邻滤波器角度的偏差。最后, 得到二维 Log-Gabor 小波的传递函数为

$$G(\omega_0, \theta_0) = \exp\left(-\frac{(\log(\omega/\omega_0))^2}{2(\log(\beta/\omega_0))^2}\right) \exp\left(-\frac{(\theta - \theta_0)^2}{2\varphi_0}\right) \quad (3)$$

1.2 人脸图像的 Log-Gabor 特征提取

假设一幅人脸图像用 $I(x, y)$ 表示, 要求取 Log-Gabor 特征, 采用的方法和求 Gabor 特征类似, 将 Log-Gabor 小波与人脸图像进行卷积, 即 $G_{u,v}(x, y) = I(x, y) * \psi_{u,v}(x, y)$ 。其中: $\psi_{u,v}(x, y)$ 是尺度为 u 、方向为 v 的 Log-Gabor 小波; $G_{u,v}(x, y)$ 是相应的卷积结果。为了方便在后面与 Gabor 小波进行比较, 此处仍选择 $v = \{0, 1, 2, 3, 4\}$, $u = \mu\pi/8$, $\mu \in \{0, 1, 2, \dots, 7\}$, 即五个尺度和八个方向, 共有 40 种组合, 即图像的每一像素位置可以获得 40 个 Log-Gabor 展开系数。图 1 给出了一幅人脸经过 Log-Gabor 小波变换之后的实部和幅值图像。

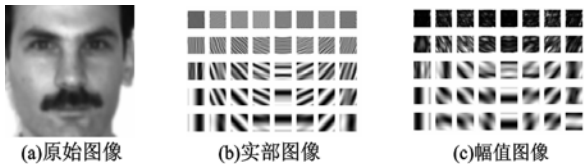


图1 图像的Log-Gabor变换

2 基于 2D-SSDA 的 Log-Gabor 特征选择及分类

2.1 2D-SSDA 算法

如果以一个大小为 32×32 的人脸图像经 Log-Gabor 小波变换后的幅值谱作为特征, 会得到 40 个与原始图像同样大小的特征矩阵, 通过行堆叠的方式将每个特征矩阵转换成一个 1024 维实向量, 然后再把所有实向量连接成一个 $1024 \times 40 = 40960$ 维实向量作为人脸的表征, 直接用于人脸图检索, 势必降低整个系统的性能。因此, 有必要对这个高维的特征进行选择, 选出更具判别能力的 Log-Gabor 特征。

本文采用二维 Log-Gaborface 矩阵来表征人脸, 即图像由 40 个 $m \times n$ 的特征矩阵来表征, 而非采用一个一维列向量。它的定义如下:

$$R = \begin{pmatrix} G_{0,0} & G_{0,1} & \cdots & G_{0,7} \\ G_{1,0} & G_{1,1} & \cdots & G_{1,7} \\ \cdots & \cdots & \cdots & \cdots \\ G_{4,0} & G_{4,1} & \cdots & G_{4,7} \end{pmatrix} \quad (4)$$

其中: 每个 Log-Gaborface: $G_{u,v}(x, y)$ 是一个二维矩阵, 它体现了人脸图像在某个方向 u 和尺度 v 下的特征。本文将所有 Log-

Gaborface 以矩阵 R 的形式进行整合, 这样 R 就包含了所有方向和尺度下的全部特征信息, 称 R 为人脸图像的整体 Log-Gaborface 图像矩阵(entity Log-Gaborface, ELoG)。最后, 利用本文提出的 2D-SSDA 来直接处理 R , 这样既可以克服小样本带来的奇异问题, 又可大大减少整个检索系统的运行时间。

传统的有监督学习的 LDA 方法, 其实质是寻求一个最佳投影矩阵 $A_{opt} = \arg \max_a |a^T S_b a / a^T S_w a|$ 。为防止小样本带来的奇异问题, 可在判别式中增加一个规则项, 即

$$A_{opt} = \arg \max_a \left| \frac{a^T S_b a}{a^T (S_w + \lambda I) a} \right| \quad (5)$$

为了进行半监督学习, 本文基于 Song 等人^[12] 提出的规则化框架, 对式(5)进行修改:

$$\arg \max_a \frac{a^T S_b a}{a^T S_l a + \alpha J(a)} \quad (6)$$

其中: S_l 为整体散布矩阵, $J(a)$ 为一规则化项, 其定义为

$$J(a) = \|a\|^2 \quad (7)$$

进一步利用 K 近邻或者 ε 邻域构建邻近图 $G(N, V)$, 并对 G 的边赋予权值, 建立加权邻接矩阵 S :

$$S_{ij} = \begin{cases} 1 & \text{如果 } x_i \text{ 是 } x_j \text{ 的 } k \text{ 邻居或 } \varepsilon \text{ 邻域, 反之亦然} \\ 0 & \text{otherwise} \end{cases}$$

这样, 式(7)可以转换为

$$J(a) = \sum_{i,j} (a^T x_i - a^T x_j)^2 S_{ij} = 2 \sum_i a^T x_i D_{ii} x_i^T a - 2 \sum_{ij} a^T x_i S_{ij} x_j^T a = 2a^T XDX^T a - 2a^T XLX^T a = 2a^T X(D - S)X^T a = 2a^T XLX^T a \quad (8)$$

其中: D 为对角阵, $D_{ii} = \sum_j S_{ij}$, $L = D - S$ 为 Laplacian 矩阵。这样利用数据之间的依赖性关系, 半监督判别分析的目标函数为

$$\max_a \frac{a^T S_b a}{a^T (S_l + \alpha XLX^T) a} \quad (9)$$

通过求解式(8)的最优化问题, 即对应求解式(10)的最大特征值所对应的特征向量。

$$S_b a = \lambda (S_l + \alpha XLX^T) a \quad (10)$$

下面给出 2D-SSDA 算法的具体步骤:

假定有一组标记样本集为 $\{(X_i, Y_i)\}_{i=1}^l$ 属于 c 个类别, 非标记样本集为 $\{(X_i)\}_{i=1}^m$, X_i 为 $m \times n$ 的矩阵。其中第 k 类有 l_k 个标记样本, 则共有 $\sum_{k=1}^c l_k = l$ 。不失一般性, 假设标记样本数已按照各自的标记排序好, 即 $\{X_1, X_2, \dots, X_l\}$ 。

a) 利用 K 近邻或者 ε 邻域构建邻近图 G , 并建立加权邻接矩阵 S 。

b) 构建标记图, 为标记图建立权值矩阵 $W \in R^{m \times m}$, 即 $W = \begin{pmatrix} W_{l \times l} & 0 \\ 0 & 0 \end{pmatrix}$, 其中 $W_{l \times l} \in R^{l \times l}$, 定义为

$$W_{l \times l} = \begin{pmatrix} W^{(1)} & 0 & \cdots & 0 \\ 0 & W^{(2)} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & W^{(c)} \end{pmatrix}$$

$\{W^{(k)}\}_{k=1}^c$ 是一个 $l_k \times l_k$ 的矩阵, 并且里面的每个元素为 $1/l_k$ 。

另外, 还定义 $\tilde{I} = \begin{pmatrix} I_{l \times l} & 0 \\ 0 & 0 \end{pmatrix}$, $I_{l \times l}$ 为单位矩阵。

c) 通过求解式(11)的广义特征向量问题获得特征值所对应的特征向量:

$$XWX^T a = \lambda X(\tilde{I} + \alpha L)X^T a \quad (11)$$

其中: $X=(X_1,X_2,\cdots,X_l,X_{l+1},\cdots,X_m)$ 。从矩阵 W 观察可知,其秩为 c ,即为类别数,因此,式(11)相应应有 c 个非零特征值并对应 c 个特征向量,组成大小为 $n\times c$ 的变换矩阵 $A=(a_1,a_2,\cdots,a_c)$ 。

d)求取 2D-SSDA 嵌入。对任意的一幅样本图像矩阵 X ,通过在变换矩阵 A 的投影,求得在低维子空间的嵌入为

$$Z_i=A^T X_i$$

(12)

其中: Z_i 是一个 n 维行向量, X_i 是一个 $m\times n$ 的矩阵。最后,获得的低维嵌入为

$$Z_i=\begin{bmatrix} Z_i^1 \\ \cdots \\ Z_i^c \end{bmatrix}$$

2.2 特征提取与分类

任意给定一幅要查询的图像 q ,其所对应的 Log-Gaborface 矩阵表示为 R_q 。其降维后的特征表示为

$$Z_q=\begin{bmatrix} Z_q^1 \\ \cdots \\ Z_q^c \end{bmatrix}$$

通过计算 Z_q 与图像库中的 Z_i 之间的相似度,其定义为

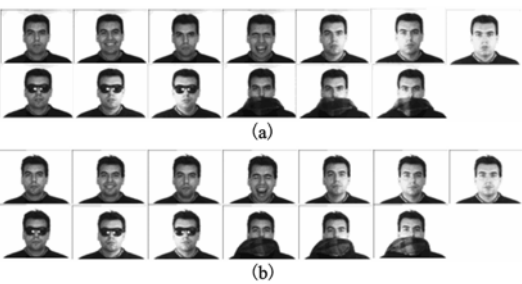
$$d(Z_q,Z_i)=\sum_{v=1}^c\|Z_q^v-Z_i^v\|$$

(13)

其中: $\|\cdot\|$ 表示马氏距离。

3 实验结果与分析

为比较本文算法与其他算法的性能,选用了 CMU PIE 数据库和 AR 人脸数据库进行实验。CMU PIE 人脸库包括 68 个对象的人脸图像,共计 41 368 幅,其中每个人有 13 种姿态、43 种光照和 4 种表情。AR 人脸库共有 126 人(70 名男士,56 名女士),每人 26 张图像。该数据库中的人脸图像全部为正面直立人脸,并且在严格控制的环境中采集。数据库中的人脸图像包含了表情、光照变化及部分遮挡等情况,但是对服饰、化妆、发型没有严格的限制。整个图像采集过程分为两个阶段进行,前后间隔两星期,其余的采集条件保持不变。其人脸图像的命名形式为“m-001-1”,其中第一个字母表示人的性别分别为“m”和“w”;第二个字段“-001”表示该人的编号;第三个字段表示人脸图像的状态包括自然表情、微笑、愤怒等共计 13 种不同状态。图 2 给出了一个在 13 种不同状态下前后两次采集的图像。由于部分人脸图像丢失,实验中,本文选择了 110 个人(64 名男士,46 名女士)的人脸图像,共计 2 860 幅。



(a)(b)两组图像的采集时间间隔两星期。从左到右,从上到下依次为自然表情、微笑、愤怒、尖叫、左侧光、右侧光、双侧光、太阳镜、太阳镜+左侧光、太阳镜+右侧光、围巾、围巾+左侧光、围巾+右侧光

图2 AR库中一个人采集的26张人脸图像

3.1 在 CMU PIE 数据库上的实验

在本实验中,选用了近似正面姿态(C05, C07, C09, C27 和 C29)的所有光照和表情条件下的人脸图像,这样每个对象

获得 170 幅人脸,共计 10 560 幅。实验中为减少计算量,在进行预处理时,将人脸图像的大小变为 32×32 像素大小。图 3 给出了其中一人经过预处理后的部分样本人脸。下面将进行两个实验:一是在部分标记样本情况下,比较各种方法的检索性能;二是进行 Gabor 和 Log-Gabor 的性能比较。



图3 来自CMU PIE的人脸库剪切后的部分人脸图像

在第一个实验里,任意选取每个对象的 30 幅人脸组成训练集,其余人脸图像形成查询库。为了比较本文方法与其他方法如 PCA、LPP、LDA、SSDA 等的检索性能。对于半监督学习方法,仅对训练集中每个对象的其中 5 幅人脸进行标记;对于 PCA 和 LPP,则使用全部的训练集(每个对象的 30 幅人脸都未标记)获得投影向量;对于 LDA 则使用全部经过手工标记的训练样本集(每个对象的 30 幅人脸都标记)来获得投影向量。

图 4 为两幅查询图像利用本文方法获得的最前面的 10 个检索结果。图 5 给出了基于整体 LoGaborface 表征的 2DSSDA 与直接作用于图像矩阵的 2DSSDA 方法、2DPCA 方法在不同特征矢量个数情况下的 top 10 retrieval rate 结果。可以看出,本文方法采用 Log-Gabor 后,检索率得到了很大的提高。另外还比较了基于一维降采样后的 Log-Gabor 特征矢量的 SSDA 方法以及常规的 PCA、LPP、LDA 方法。从图 5 中可以看出,Log-Gabor + SSDA 方法获得了与 LDA 相当的平均检索率,但单独的 SSDA 方法的检索率要低于 LDA 方法,这主要是因为 LDA 方法利用了全部标记信息。



图4 两幅查询人脸得到的top 10检索结果

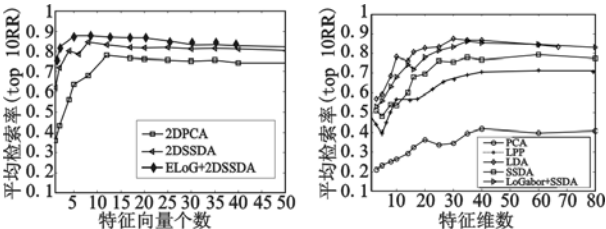


图5 各种方法在不同投影矢量下的平均检索率

表 1 给出了各种算法获得最好 top 10 检索率时所对应的特征向量个数以及检索时间。从表 1 中可以看出,基于 ELoGFace 表征的 2DSSDA 获得了最好的检索率,并且在获得好的检索率时,本文方法仅需要 5 个特征向量,这说明所提出的方法在低维情况下比其他方法更有效;2DSSDA 的检索率优于 PCA、LPP 以及 SSDA 方法,表明 2DSSD 方法不但充分利用所有的标记和非标记样本进行学习,而且还保留了人脸数据的局部流形结构,在

检索性能方面也获得了提高,而基于一维 LoGabor 特征矢量的 SSDA 方法又优于 2DSSDA,表明经过 Log-Gabor 滤波器后的特征对光照和表情具有更好的鲁棒性;另外,基于整体 LogGabor 表征的 2DSSDA 方法比基于一维 LoGabor 特征矢量的 SSDA 方法所用更少的检索时间。但是从表中可以看出,本文所提出的方法占用的存储空间是所有方法中最大的,这是因为二维特征提取算法所需的存储量总是高于一维算法。

表 1 各种方法在 CMU PIE 数据库上的检索性能比较

方法	top 10 retrieval rate/%	维数	检索时间/s
PCA	41.6	104	6.71
LPP	68.6	78	4.63
SSDA	79.2	67	4.22
2D-SSDA	84.5	8×32	2.54
LoG+SSDA	87.2	67	12.12
ELoG+2DSSDA	90.3	5×160	7.16

本文对 SSDA 方法在不同标记训练样本数量下的平均检索率进行了实验对比,实验结果如表 2 所示。从表 2 中可知,随着训练样本标记数量的增加,平均检索率随之增加,逐步接近 LDA 方法的性能,但 SSDA 方法只需要对少量训练样本数据进行标记,因而可以更好地应用于图像检索领域。

表 2 SSDA 在 CMU 库上的平均检索率与类别标记数量关系

	类别标记数量			
	5	10	18	25
平均检索率/%	79.2	83.4	84.7	86.2

第二个实验是验证 Gabor 和 Log-Gabor 的检索性能。在此实验里,两种滤波器都选用相同的 5 个尺度和 8 个方向对人脸图像进行卷积。构造整体 Gaborface 和 Log-Gaborface,然后利用 2DSSDA 进行降维,采用与第一个实验相同的实验数据,它们的比较结果如图 6 所示。从图中可以看出,基于 Log-Gaborface 的方法的性能要优于 Gaborface 方法,这也充分说明 Log-Gabor 小波更符合人类的视觉特性。

3.2 在 AR 数据库上的实验

本实验将在 AR 数据库上对本文所提出来的方法与其他方法进行比较,以表明本文方法对光照、表情变化更具鲁棒性。

实验中,选用第一阶段的 110 个对象的所有表情和光照人脸组成训练集,第二阶段对应的所有表情和光照人脸作为查询集。任意选择训练集中每个对象的两幅人脸进行标记,将本文的算法与 LPP、SSDA 等方法进行了比较实验,图 7 给出了比较结果。

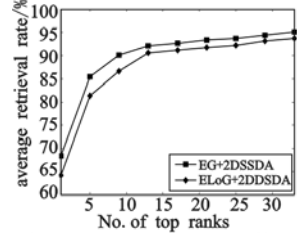


图6 Gabor和Log-Gabor的检索性能比较

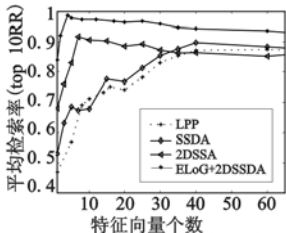


图7 各种方法在不同投影矢量下的平均检索率

从图 7 中可以看出,本文所提方法仍然优于其他方法,并且采用基于 ELoGface 表征的 2DSSDA 方法,取得了 98.7% 的最好平均检索率,所需的特征向量的个数仅为 4 个。直接 2DSSDA 的最好平均检索率为 91.8%,所需特征向量的个数为 7 个。可见,Log-Gabor 特征对具有光照和表情变化的人脸图

像具有很好的表征能力。

4 结束语

针对 Gabor 小波存在直流分量与过度表达低频分量而对高频分量表达不足等缺点,本文引入了更符合人类视觉特性的 Log-Gabor 小波。进一步针对大型人脸图像库的训练样本中只有部分标记问题,本文结合传统的 LDA 方法,对半监督判别分析算法进行扩展,提出了二维半监督判别分析算法。该算法直接作用于二维 Log-Gabor 特征构成的组合矩阵,克服了小样本带来的奇异问题;另外,该方法还充分利用了标记信息和未标记信息,同时保留了数据的局部流形结构,增大了特征匹配的相似性。在人脸数据集上的实验结果证实了本文方法是可行的。

参考文献:

[1] DAUGMAN J. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression[J]. IEEE Trans on Acoustics, Speech, and Signal Processing,1988,36(7):1169-1179.

[2] HUANG Rong-bing, DU Ming-hui. Object region based color image retrieval integrating multi-features[C]//Proc of International Conference on Wavelet Analysis and Pattern Recognition. [S. l.]:IEEE, 2007: 1729-1734.

[3] LIU C, WECHSLER H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition[J]. IEEE Trans on Image Processing, 2002,11(4):467-476.

[4] SHAN S, YANG P, CHEN X, et al. AdaBoost Gabor fisher classifier for face recognition[C]//Proc of the 2nd International Workshop on Analysis and Modeling of Faces and Gestures. [S. l.]: Springer, 2005: 278-291.

[5] KOVESI P. Invariant measures of image features from phase information [D]. Perth: University of Western Australian, 1996.

[6] 黄鸿,李见为,冯海亮. 融合 Log-Gabor 小波和监督保局映射的人脸识别算法[J]. 计算机辅助设计与图形学学报, 2008,20(10): 1332-1337.

[7] HE J R, LI M J, ZHANG H J, et al. Generalized manifold-ranking-based image retrieval [J]. IEEE Trans on Image Processing, 2006,15(10):3170-3177.

[8] CAI Deng, HE Xiao-fei, HAN J. Semi-supervised discriminant analysis[C]//Proc of the 11th IEEE International Conference on Computer Vision. [S. l.]:IEEE,2007:1-7.

[9] HE X F. Incremental semi-supervised subspace learning for image retrieval[C]//Proc of the 12th ACM International Conference on Multimedia. New York: ACM Press,2004:2-8.

[10] YANG J, ZHANG D, FRANGI A F, et al. Two-dimensional PCA: a new approach to appearance-based face representation and recognition [J]. IEEE Trans on PAMI, 2004,26(1):131-137.

[11] CHEN S, ZHAO H, KONG M. 2D-LPP: a two-dimensional extension of locality preserving projections[J]. Neurocomputing,2007, 70(4-6):912-921.

[12] SONG Y, NIE F, ZHANG C, et al. A unified framework for semi-supervised dimensionality reduction[J]. Pattern Recognition,2008, 41(9):2789-2799.

[13] FIELD D J. Relations between the statistics of natural images and the response properties of cortical cells[J]. Journal of Optical of America A,1987, 4(12): 2379-2394.