

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2013 2
Vol.18 No.

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2013 年 2 月 第 18 卷 第 2 期(总第 202 期)

目次

计算机视觉前沿论坛

- 计算机视觉——探索行为理解,认知内心世界 徐光祐,刘允才,章毓晋(131)
- 跨学科前沿与应用的交汇点:日常活动理解 徐光祐,陶霖密,邸慧军(132)
- 时空行为理解 章毓晋(141)
- 视觉机制研究对机器视觉的启发示例 李雄,刘允才(152)
- 视点无关的行为识别综述 冯家更,肖俊(157)

图像处理和编码

- 基于演化算法的卷曲 DCT 图像压缩 李康顺,韦蕴珊,张文生(169)
- 基于欧氏距离图的图像边缘检测 张闯,王婷婷,孙冬娇,葛益娴,常建华(176)
- 基于暗原色先验模型的快速去雾算法 张冰冰,戴声奎,孙万源(184)

图像分析和识别

- Gabor 特征集结合判别式字典学习的稀疏表示图像识别 胡正平,徐波,白洋(189)
- 半监督 k 近邻分类方法 陈日新,朱明早(195)
- 最小距离鉴别投影及其在人脸识别中的应用 黄璞,唐振民(201)
- 图像 LSB 匹配隐藏的预分类检测模型 曹卫权,韩杰思,王宏霞(207)
- 具有旋转鲁棒性的纹理谱描述子 陈刚,陈晓云(214)

计算机图形学

- 利用形体特征的铅笔素描画生成 莫晓斐,丁友东(219)

医学图像处理

- 统计相似度特征的医学图像分割 郭艳蓉,蒋建国,郝世杰,詹曙,李鸿(225)

遥感图像处理

- 结合遗传算法和蚁群算法的高光谱图像波段选择 王立国,魏芳洁(235)
- 相似度自适应样本块的高分辨 SAR 图像阴影修复 赵昊,张弓(243)

- 第八届国际数字地球会议征稿通知 封 3

Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 18 No. 2 February 2013

Contents

Forum: Forefront of Computer Vision

Confluence of interdisciplinary frontiers and real-world application: understanding activities of daily living

..... Xu Guangyou, Tao Linmi, Di Huijun (132)

Understanding spatial-temporal behaviors Zhang Yujin (141)

The motivation of visual mechanisms to machine vision: examples Li Xiong, Liu Yuncai (152)

View-invariant action recognition: a survey Feng Jiageng, Xiao jun (157)

Image Processing and Coding

Warped DCT image compression based on an evolutionary algorithm Li Kangshun, Wei Yunshan, Zhang Wensheng (169)

Image edge detection based on the Euclidean distance graph

..... Zhang Chuang, Wang Tingting, Sun Dongjiao, Ge Yixian, Chang Jianhua (176)

Fast image haze-removal algorithm based on the prior dark-channel

..... Zhang Bingbing, Dai Shengkui, Sun Wanyuan (184)

Image Analysis and Recognition

Sparse representation for image recognition based on Gabor feature set and discriminative dictionary learning

..... Hu Zhengping, Xu Bo, Bai Yang (189)

Semi-supervised k -nearest neighbor classification method Chen Rixing, Zhu Minghan (195)

Minimum-distance discriminant projection and its application to face recognition Huang Pu, Tang Zhenmin (201)

Using pre-classification to attack LSB matching Cao Wei-quan, Han Jiesi, Wang Hongxia (207)

Robust rotation-invariant texture spectrum descriptor Chen Gang, Chen Xiaoyun (214)

Computer Graphics

Pencil sketch generation with shape features Mo Xiaofei, Ding Youdong (219)

Medical Image Processing

Medical image segmentation based on statistical similarity feature

..... Guo Yanrong, Jiang Jianguo, Hao Shijie, Zhan Shu, Li Hong (225)

Remote Sensing Image Processing

Band selection for hyperspectral imagery based on combination of genetic algorithm and ant colony algorithm

..... Wang Liguang, Wei Fangjie (235)

Adaptive shadow inpainting from high-resolution SAR image based on exemplar's similarity Zhao Hao, Zhang Gong (243)

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2013)02-189-06

论文引用格式: 胡正平, 徐波, 白洋. Gabor 特征集结合判别式字典学习的稀疏表示图像识别[J]. 中国图象图形学报, 2013, 18(2): 189-194.

Gabor 特征集结合判别式字典学习的 稀疏表示图像识别

胡正平, 徐波, 白洋

燕山大学信息科学与工程学院, 秦皇岛 066004

摘要: 稀疏编码中字典的选择无论对图像重建还是模式分类都有重要影响, 为此提出 Gabor 特征集结合判别式字典学习的稀疏表示图像识别算法。考虑到 Gabor 局部特征对光照、表情和姿态等变化的鲁棒性, 首先提取图像对应不同方向、不同尺度的多个 Gabor 特征; 然后将降维的增广 Gabor 特征矩阵作为初始特征字典, 通过对该字典的学习得到字典原子对应类别标签的新结构化字典, 新字典中特定类的子字典对相关的类具有好的表示能力, 同时应用 Fisher 判别约束编码系数, 使它们具有小的类内散度和大的类间散度; 最后同时用具有判别性的重构误差和编码系数来进行模式分类。基于 3 个数据库的实验结果表明本文方法具有可行性和有效性。

关键词: 稀疏表示; 稀疏模式分类; Gabor 特征; Fisher 字典学习

Sparse representation for image recognition based on Gabor feature set and discriminative dictionary learning

Hu Zhengping, Xu Bo, Bai Yang

Institute of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, China

Abstract: Choosing the right dictionary used for sparse coding has an important effect on image reconstruction and pattern classification. Therefore, a new sparse representation algorithm based on Gabor Feature Set Discriminative Dictionary Learning is proposed for image recognition. Considering that Gabor feature is robust to variations of illumination, expression, and pose, the proposed method first extracts the image Gabor features with multi-scale and multi-orientation. Then it uses the augmented Gabor local feature matrix whose dimension has been reduced to construct the initial feature dictionary. This reduction is based on the Fisher discrimination criterion. A structural dictionary, whose atoms correspond to the class labels, is learned so that each sub-dictionary of the learned new dictionary is a good representation of the samples from the corresponding class. Meanwhile, the Fisher discrimination criterion is imposed on the coding coefficients so that they have small within-class scatter but big between-class scatter. Consequently, a new classification scheme associated with the proposed method is then presented by using the discriminative information and sparse coding coefficients. Experiments on three types of databases show that the proposed method is valid and efficient.

Key words: sparse representation; sparse pattern classification (SPC); Gabor feature; Fisher dictionary learning

收稿日期: 2012-06-18; 修回日期: 2012-08-04

基金项目: 国家自然科学基金项目(61071199); 河北省自然科学基金项目(F2010001297); 中国博士后自然科学基金项目(20080440124); 第二批中国博士后科学基金项目(200902356)

第一作者简介: 胡正平(1970—)男, 教授, 2007年于哈尔滨工业大学获信息与通信工程专业博士学位, 主要研究方向为稀疏模式识别。

E-mail: hzp@ysu.edu.cn

0 引言

寻求客观事物的稀疏表示方法,一直是计算机视觉、数学、图像处理、模式识别领域的研究目标。特别是近来以 $L1$ 、 $L1/2$ 优化为核心的信号稀疏表示思路已经蔓延到模式识别界。可以预见,基于稀疏表示的模式识别研究理论将越来越受到重视,一方面它提供了高维空间的鲁棒低维表示,为低维空间建立新的模式识别算法提供巨大空间,同时还可以利用表示的稀疏度量准则来衡量模式之间的相似度。

基于信号稀疏表示的图像识别方法是基于过完备字典对测试样本的表示具有稀疏性这一先验知识,将识别问题看作是多个线性回归模型的分类问题,从稀疏的角度得到对图像最紧致的表示。文献[1]提出稀疏表示的分类算法(SRC),该算法将测试样本表示为训练样本的加权线性组合,通过 $L1$ 范数优化求取稀疏解,使得除测试样本所属类别外的其他训练样本的系数均为接近零,进而利用系数中所包含的判别信息,实现对测试图像的分类,对噪声和遮挡等表现出好的鲁棒性。稀疏表示方法本身具有判别性:自适应的选择所有表示中最稀疏的一种来表示待识别样本,并拒绝所有其他不稀疏的表示方式。将稀疏表示应用于图像识别领域,正是利用其本身所具有的这种判别性信息^[1]。在稀疏表示的图像识别方法中,过完备字典由所有训练样本构成,字典中的每一个基向量即代表一个训练样本,而且一般认为同一类训练样本位于同一个子空间中。在文献[1]的基础上,近两年国内外学者已提出不少稀疏表示的图像分类方法。具有代表性的思路是对用于稀疏表示的过完备字典进行学习,使其更利于模式分类。字典学习旨在从训练样本学习一个能更好地表示或编码给定信号的一组基。例如文献[2]为每一类样本学习一个更加紧凑的字典,在人脸识别上取得比 SRC 更好的效果。如何增加字典的约束条件,使得学习的字典更适应分类问题逐渐得到关注。为此文献[3]通过增加不相干限制条件使每一类对应的子字典尽可能的相互独立,且在稀疏编码方式下能更好地实现信号重构与分类;一些研究者提出针对人脸识别的判别 KSVD 算法,通过寻找梯度下降的最优步长达到全局最小

收敛,得到用于稀疏表示的最佳字典^[4-6];文献[7]提出标签一致 K-SVD 算法(LC-KSVD),该算法在字典学习模型中增加一个判别性稀疏编码误差约束来获得判别能力用于图像识别;另外一些研究者将最小描述长度(MDL)原则以及区分性限制约束条件引入稀疏表示模型,构建出新的字典学习模型^[8-9]。上述研究出发点都是在原始训练样本上进行字典学习,虽然这些基于字典学习的稀疏表示方法取得较好的模式分类性能,但由于这些方法都是基于全局信息进行处理,难以充分利用局部特征分布信息。

文献[1]指出基于样本稀疏表示的模式分类,特征空间的维数或最佳特征变换的选择对于分类性能提高影响不大,这里主要是针对整体特征而言的,而实际中由于样本数目通常是有限的,整体特征不能有效地处理光照、表情、姿态和局部形变等变化。为此,基于局部特征的描述分类识别思路得到国内外学者的关注^[10-14],例如 Gabor 是一种强大的特征提取器,Gabor 小波可以模拟大脑皮层中单细胞感受野的轮廓,捕捉突出的视觉属性,特别是 Gabor 小波可以提取图像特定区域内的多尺度多方向空间频率特征,像显微镜一样放大灰度的变化。由于 Gabor 特征是从图像局部提取的,所以它对光照、表情和姿态的变化相对于整体特征更鲁棒。

考虑到 Gabor 特征对光照、表情和姿态等变化的鲁棒性和判别性字典学习在模式分类的作用,可以用在 Gabor 特征集基础上进行字典学习代替直接样本的字典学习吗?正是基于这个思路,提出基于 Gabor 特征集结合判别式字典学习的稀疏表示图像识别算法。首先提取图像多尺度多方向的 Gabor 特征,将增广 Gabor 特征矩阵作为初始特征字典,对该特征字典进行学习,确保新字典有更强的表示能力,以减小重构误差,同时将 Fisher 判别准则应用于编码系数,使得稀疏编码系数的类内离散度尽可能小而类间离散度尽可能的大,每类的特征差异变大,从而使每个特定类的子字典对来自与其同一类的训练样本的特征具有更强的表示能力,而对其他类样本的表示却相反,这样稀疏编码系数的判别性增强,更利于提高模式分类性能。

1 稀疏表示分类原理

稀疏表示分类方法的核心思想是将测试图像

表示为训练样本的线性组合,假设有 c 类样本,用 $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_c]$ 表示原始训练样本集, $\mathbf{A}_i$ 是第 i 类的训练样本子集, $\mathbf{A}_i = [\mathbf{s}_{i,1}, \mathbf{s}_{i,2}, \dots, \mathbf{s}_{i,n_i}] \in \mathbf{R}^{m \times n_i}$, $\mathbf{s}_{i,j}, j=1, 2, \dots, n_i$ 是第 i 类中第 j 个样本张成的 m 维列向量。用 $\mathbf{y}$ 代表测试样本。SRC 分类方法流程如下:

1) 通过 l_1 范数最小化在字典 $\mathbf{A}$ 上对 $\mathbf{y}$ 进行稀疏编码

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} \{ \|\mathbf{y} - \mathbf{A}\boldsymbol{\alpha}\|_2^2 + \gamma \|\boldsymbol{\alpha}\|_1 \} \quad (1)$$

式中, γ 是一个标量。

2) 通过式(2)来分类

$$\text{identity}(\mathbf{y}) = \arg \min_i \{e_i\} \quad (2)$$

式中, $e_i = \|\mathbf{y} - \mathbf{A}\hat{\boldsymbol{\alpha}}_i\|_2$, $\hat{\boldsymbol{\alpha}} = [\hat{\boldsymbol{\alpha}}_1; \hat{\boldsymbol{\alpha}}_2; \dots; \hat{\boldsymbol{\alpha}}_c]$, $\hat{\boldsymbol{\alpha}}_i$ 对应于第 i 类样本的系数向量。基于稀疏表示的分类方法将测试样本判定为具有最小重构误差 e_i 的那一类所属的类别。

2 Gabor 特征集结合判别式字典学习的稀疏表示图像识别算法

2.1 Gabor 特征集构造

使用的 Gabor 滤波器为

$$\psi_{\mu,\nu}(z) = \frac{\|k_{\mu,\nu}\|^2}{\sigma^2} e^{-\frac{\|k_{\mu,\nu}\|^2 \|z\|^2}{2\sigma^2}} \times (e^{ik_{\mu,\nu}z} - e^{-\sigma^2/2}) \quad (3)$$

式中, $z = (x, y)$ 代表像素; $\|\cdot\|$ 表示范数运算; $k_{\mu,\nu} = k_v e^{i\phi_\mu}$, $k_v = \frac{k_{\max}}{f^v}$, $\phi_\mu = \frac{\pi\mu}{8}$, $k_{\max}$ 是最大频率, f 是频域中核函数的间隔因数, μ 和 ν 分别表示 Gabor 滤波器的方向和尺度。

Gabor 特征提取是通过 5 个尺度和 8 个方向的 Gabor 滤波器分别与图像进行卷积操作来完成。令 $I(x, y)$ 表示人脸图像的灰度分布, 那么 $I(x, y)$ 和 Gabor 滤波器的卷积可定义为

$$G_{\mu,\nu}(z) = I(x, y) * \psi_{\mu,\nu}(z) \quad (4)$$

式中, $G_{\mu,\nu}(z)$ 是复数, 可以写为 $G_{\mu,\nu}(z) = M_{\mu,\nu}(z) \cdot \exp(i\theta_{\mu,\nu}(z))$, $M_{\mu,\nu}(z)$ 是幅值部分, $\theta_{\mu,\nu}(z)$ 是相位。幅值信息包含图像局部能量的变化, 因而将变换后的幅值作为图像特征的描述。图像与 40 个 Gabor 滤波器进行卷积, Gabor 特征的维数很高, 所以需把降维后的 Gabor 特征矩阵作为稀疏表示的过完备字典。对于得到的幅值 $M_{\mu,\nu}(z)$ 先通过均衡下采样, 并

归一化到零均值和单位方差, 然后将其按列连接起来, 最后将这 40 个列向量连结在一起构成增广 Gabor 特征矢量

$$\boldsymbol{\chi} = [\boldsymbol{\alpha}_{0,0}^{(\rho)\top}, \boldsymbol{\alpha}_{0,1}^{(\rho)\top}, \dots, \boldsymbol{\alpha}_{4,7}^{(\rho)\top}]^\top \quad (5)$$

式中, $\boldsymbol{\alpha}_{u,v}^{(\rho)}$ 是将经过抽样因子 ρ 下采样的幅值矩阵 $\mathbf{M}_{u,v}^{(\rho)}$ 的各列连接得到的列向量。

增广 Gabor 特征矢量 $\boldsymbol{\chi}$ 是一个局部特征描述子, 在稀疏表示框架内, 用 $\boldsymbol{\chi}$ 替代 SRC 中的整体特征, 则 Gabor 特征字典 $\mathbf{G} = \boldsymbol{\chi}(\mathbf{A}) = [\boldsymbol{\chi}(\mathbf{A}_1), \boldsymbol{\chi}(\mathbf{A}_2), \dots, \boldsymbol{\chi}(\mathbf{A}_c)]$, 其中, $\boldsymbol{\chi}(\mathbf{A}_i) = [\boldsymbol{\chi}(\mathbf{s}_{i,1}), \boldsymbol{\chi}(\mathbf{s}_{i,2}), \dots, \boldsymbol{\chi}(\mathbf{s}_{i,n_i})]$, 测试样本 $\mathbf{y}$ 相应的变为 $\boldsymbol{\chi}(\mathbf{y})$ 。

2.2 基于 Fisher 准则的特征字典学习

令 $\mathbf{G} = [\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_c]$ 表示 c 类样本的 Gabor 特征字典, $\mathbf{G}_i$ 是来自第 i 类训练样本的 Gabor 特征子集。不同于文献[6]中各类共享字典方法, 学习一个结构化的字典 $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_c]$, $\mathbf{D}_i$ 是第 i 类的子字典。用 $\mathbf{X}$ 表示 $\mathbf{G}$ 在 $\mathbf{D}$ 上的编码系数矩阵, 即 $\mathbf{G} \approx \mathbf{D}\mathbf{X}$, $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_c]$, 其中, $\mathbf{X}_i$ 是包含 $\mathbf{G}_i$ 在字典 $\mathbf{D}$ 上编码系数的子矩阵。要获得比 $\mathbf{G}$ 更适于分类的字典, 需要对 $\mathbf{D}$ 进行优化约束, 除了要求 $\mathbf{D}$ 具有较强的重构能力, 还希望具有较强的区分能力, 为此基于 Fisher 准则的特征字典学习模型为

$$J_{(\mathbf{D}, \mathbf{X})} = \arg \min_{(\mathbf{D}, \mathbf{X})} \{ r(\mathbf{G}, \mathbf{D}, \mathbf{X}) + \lambda_1 \|\mathbf{X}\|_1 + \lambda_2 f(\mathbf{X}) \} \quad (6)$$

式中, $r(\mathbf{G}, \mathbf{D}, \mathbf{X})$ 是判别保真项; $\|\mathbf{X}\|_1$ 是稀疏性约束; $f(\mathbf{X})$ 是加在系数矩阵 $\mathbf{X}$ 上的判别性约束; λ_1 和 λ_2 是平衡因子。

假设 $\mathbf{G}_i$ 在字典 $\mathbf{D}$ 上编码系数 $\mathbf{X}_i$ 表示为 $\mathbf{X}_i = [\mathbf{X}_i^1; \dots; \mathbf{X}_i^j; \dots; \mathbf{X}_i^c]$, 其中, $\mathbf{X}_i^j$ 是 $\mathbf{G}_i$ 在子字典 $\mathbf{D}_j$ 上的编码系数。假设字典 $\mathbf{D}_k$ 对 $\mathbf{G}_i$ 的表示记为 $\mathbf{R}_k = \mathbf{D}_k \mathbf{X}_i^k$ 。首先, 字典 $\mathbf{D}$ 可以表示 $\mathbf{G}_i$, 则 $\mathbf{G}_i \approx \mathbf{D}\mathbf{X}_i = \mathbf{D}_1 \mathbf{X}_i^1 + \dots + \mathbf{D}_i \mathbf{X}_i^i + \dots + \mathbf{D}_c \mathbf{X}_i^c = \mathbf{R}_1 + \dots + \mathbf{R}_i + \dots + \mathbf{R}_c$ 。其次, 由于字典 $\mathbf{D}_i$ 是与第 i 类相关联的子字典, 因此期望 $\mathbf{G}_i$ 能被子字典 $\mathbf{D}_i$ 稀疏表示, 而不是 $\mathbf{D}_j, j \neq i$, 这意味着 $\mathbf{X}_i^j$ 中的系数应满足使式子 $\|\mathbf{G}_i - \mathbf{D}_i \mathbf{X}_i^i\|_F^2$ 较小, 同时系数矩阵 $\mathbf{X}_i^j$ 为稀疏矩阵, 从而使 $\|\mathbf{D}_j \mathbf{X}_i^j\|_F^2$ 很小。因此判别保真项为

$$r(\mathbf{G}_i, \mathbf{D}, \mathbf{X}_i) = \|\mathbf{G}_i - \mathbf{D}\mathbf{X}_i\|_F^2 + \|\mathbf{G}_i - \mathbf{D}_i \mathbf{X}_i^i\|_F^2 + \sum_{j=1, j \neq i}^c \|\mathbf{D}_j \mathbf{X}_i^j\|_F^2 \quad (7)$$

$r(\mathbf{G}_i, \mathbf{D}, \mathbf{X}_i)$ 中 3 项的一个直观的解释如图 1 所示,

图 1(a) 仅要求 D 能逼近表示 G_i , 虽然第 1 项确保 D 能逼近表示 G_i , 但是 R_i 可能会严重偏离, 增加一个约束 $\|G_i - D_i X_i\|_F^2$, 如图 1(b) 限定 D 和 D_i 都能逼近表示 G_i , 但 G_i 仍可能被其他的子字典较好地表示, 比如 D_{i-1} , 从而会降低字典 D 的表示能力。图 1(c) 表示通过 $D_j, j \neq i$ 对 G_i 的表示的多个约束, 从而克服上述问题。判别保真项使学习出的新字典中特定类的子字典对来自与其同属一类的样本特征具有更强的表示能力。

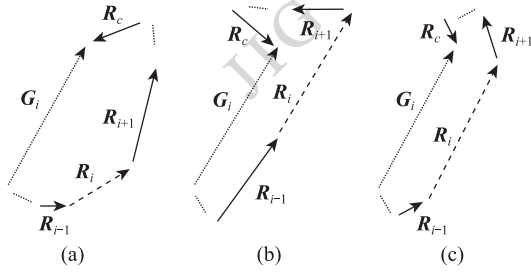


图 1 判别保真项图解

Fig. 1 Illustration of the discriminative fidelity term

为使字典 D 对图像特征具有判别性, 对图像特征 G 在字典 D 上的编码系数 X 施加 Fisher 判别准则, 通过最小化 X 的类内散度 $S_w(X)$ 和最大化 X 的类间散度 $S_B(X)$ 来实现字典 D 的这种判别性。 $S_w(X)$ 和 $S_B(X)$ 定义为

$$S_w(X) = \sum_{i=1}^c \sum_{x_k \in X_i} (x_k - m_i)(x_k - m_i)^T$$

$$S_B(X) = \sum_{i=1}^c n_i (m_i - m)(m_i - m)^T$$

式中, m_i 和 m 分别是 X_i 和 X 的均值向量, n_i 是第 i 类中样本的数目。

直观上, 可将判别性约束 $f(X)$ 定义为 $\text{tr}(S_w(X)) - \text{tr}(S_B(X))$ 。然而, 这样的 $f(X)$ 是非凸且不稳定的。为解决这个问题, 添加一个松弛项 $\|X\|_F^2$ 到 $f(X)$ 中, 则 $f(X)$ 的形式为

$$f(X) = \text{tr}(S_w(X)) - \text{tr}(S_B(X)) + \eta \|X\|_F^2 \quad (8)$$

式中, η 是平衡因子。将式(7)(8)代入式(6)中可以得到 GFDDL 模型(gabor feature discriminative dictionary learning)

$$J_{(D,X)} = \underset{(D,X)}{\text{argmin}} \left\{ \sum_{i=1}^c r(G_i, D, X_i) + \lambda_1 \|X\|_1 + \lambda_2 (\text{tr}(S_w(X)) - \text{tr}(S_B(X)) + \eta \|X\|_F^2) \right\} \quad (9)$$

式中, 目标函数 J 对 (D, X) 不是联合凸的, 但当其中一个固定时, 它关于 D 或 X 是凸的。目标函数

可分为如下两个子问题: 1) 固定 D , 更新 X ; 2) 固定 X , 更新 D 。通过迭代获得所需判别字典 D 和判别系数 X 。

首先, 假设 D 固定不变, 目标函数简化为计算 $X = [X_1, X_2, \dots, X_c]$ 的一个稀疏编码问题。分类计算 X_i , 当计算 X_i 时, 所有 $X_j, j \neq i$, 固定不变。因此, 式(9)中的目标函数可简化为

$$J_{(X_i)} = \underset{(X_i)}{\text{arg min}} \{ r(G_i, D, X_i) + \lambda_1 \|X_i\|_1 + \lambda_2 f_i(X_i) \} \quad (10)$$

式中, $f_i(X_i) = \|X_i - M_i\|_F^2 - \sum_{k=1}^c \|M_k - M\|_F^2 + \eta \|X_i\|_F^2$, M_k 和 M 分别是第 k 类和所有类的平均向量矩阵, 式(10)可以通过迭代投影算法求解出 X_i 。

当 X 固定时, 再分类更新 D_i 。更新 D_i 时, 所有 $D_j, j \neq i$, 是固定的。此时, 式(9)目标函数简化为

$$J_{(D_i)} = \underset{(D_i)}{\text{arg min}} \left\{ \|G - D_i X_i - \sum_{j=1, j \neq i}^c D_j X_j\|_F^2 + \|G_i - D_i X_i\|_F^2 + \sum_{j=1, j \neq i}^c \|D_i X_j\|_F^2 \right\} \quad (11)$$

式中, X_i 是 G 在 D_i 上的编码系数。

2.3 分类策略

经过上述方法得到特征字典 D , 在字典 D 上对测试样本的特征进行稀疏编码, 重构误差和编码系数都具有判别性。因此, 同时利用两者的判别性对测试样本的特征进行编码, 确定了样本特征的类别即得到了该样本所属的类别。

在特征字典 D 上对测试样本特征 $\chi(y)$ 进行编码, 令 $g = \chi(y)$ 。则稀疏编码系数

$$\hat{\alpha} = \underset{\alpha}{\text{arg min}} \{ \|g - D\alpha\|_2^2 + \gamma \|\alpha\|_1 \} \quad (12)$$

式中, γ 是一个平衡调节因子, $\hat{\alpha} = [\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_c]$, $\hat{\alpha}_i$ 是与子字典 D_i 相关联的系数向量。用来评价最终分类的逼近残差为

$$e_i = (1 - w) \cdot \|g - D_i \hat{\alpha}_i\|_2 + w \cdot \|\hat{\alpha} - m_i\|_2 \quad (13)$$

式中, 第 1 项是重构误差, 第 2 项是系数向量 $\hat{\alpha}$ 和第 i 类的均值向量 m_i 之间的距离, $w \geq 0$ 是预先设置的权重。测试样本特征 g 判别为具有最小逼近残差的类别, 即数学形式表示为

$$\text{identity}(g) = \underset{i \in \{1, \dots, c\}}{\text{arg min}} \{ e_i \} \quad (14)$$

3 实验仿真

为检验上述算法的有效性,将该算法应用于人脸识别和数字识别,采用 AR 和 Extended Yale B 人脸数据库以及 USPS 手写体数字库进行实验。实验所用平台为 AMD Athlon (TM) 2 处理器,主频 3.11 GHz,2 G 内存。在实验中,为防止产生过拟合,所有的正则项参数均采用交叉验证选取。为保证结果的客观性,本文方法实验结果为独立运行 20 次后的平均识别率。

3.1 AR 人脸库

AR 人脸库包含 126 个人的 4 000 多幅正面人脸图像,从中选择 50 名男性和 50 名女性各 14 张图像用于实验,存在光照和表情变化的前 7 张用于训练,另外 7 张用于测试。所有图像尺寸归一化为 60×43 像素,经 PCA 变换后特征维数 $d = 300$,参数 $\lambda_1 = 0.005, \lambda_2 = 0.05, \gamma = 0.005, w = 0.05$ 。

表 1 列出最近邻法(NN)、支持向量机(SVM)、判别 KSVD (DKSVD) 算法、稀疏表示分类方法(SRC)和文献[14]中 Gabor 特征稀疏表示算法(GSRC)在 AR 数据库上的识别准确率,可以看出基于 Gabor 特征集结合判别式字典学习的稀疏表示图像识别算法优于其他分类器的识别效果,且由于 AR 人脸库中存在光照、表情变化等,与单纯采用 PCA 降维的 SRC 算法相比,性能提升将近 10%;与单纯采用 Gabor 特征稀疏表示(GSRC)算法相比,Gabor 特征集结合判别式字典学习的稀疏表示识别方法表现出更好的鲁棒性,达到 98.3% 的最高识别率,说明在新字典上的编码对于样本的判别性增强了。

表 1 AR 库上几种方法识别率的比较

Table 1 The recognition results for AR

	NN	SVM	DKSD	SRC	GSRC	本文方法
识别率/%	71.4	87.1	85.4	88.8	96.9	98.3

3.2 Extended Yale B 人脸数据库

Extended Yale B 数据库包含 38 个人取自不同光照条件下的 2 414 张正面人脸图像,每个人大约有 64 张图像,这些图像含有光照和表情的变化。从每个人的图像中随机选取 20 幅作为训练样本,余下的部分用于测试。所有图像尺寸归一化为 54×48 像素,降维后的特征维数 $d = 300$,参数 $\lambda_1 = 0.005,$

$\lambda_2 = 0.005, \gamma = 0.001, w = 0.05$ 。

表 2 列出最近邻法(NN)、支持向量机(SVM)、判别 KSVD (DKSVD) 算法、稀疏表示分类方法(SRC)和 Gabor 特征稀疏表示算法(GSRC)在 Extended Yale B 数据库上的识别率,可以看出相对于其他分类器,本文方法的所获得的性能提升超过了 3%,比经典的 SRC 算法提高了近 5%,优于单一的采用 PCA 降维的 SRC 算法和 Gabor 特征稀疏表示算法,说明 Gabor 特征集结合判别式字典学习稀疏表示思路的合理性。

表 2 Extended Yale B 库上几种方法的识别率

Table 2 The recognition results for Yale B

	NN	SVM	DKSD	SRC	GSRC	本文方法
识别率/%	61.7	88.8	75.3	90.0	91.3	94.6

3.3 USPS 手写体数字库

选择 USPS 手写体数字库作为实验对象,验证 Gabor 特征集结合判别式字典学习的稀疏表示识别算法。该数据库包含 0~9 共 10 个数字的手写体图像,其中训练图像 7 291 幅,测试图像 2 007 幅。实验中分别从每个类别中随机选择 100 幅作为训练样本,100 幅作为测试样本。图像像素大小为 16×16 ,通过 PCA 变换对图像进行降维,降维后维数 $d = 120$,参数 $\lambda_1 = 0.01, \lambda_2 = 0.001, \gamma = 0.001, w = 0.05$ 。

表 3 列出该算法与其他几种方法识别性能的比较,包括 K 近邻(K_NN)、采用高斯核的支持向量机(SVM-Gauss)、稀疏表示的信号分类算法(SRSC)、文献[3]中的结构化非相干字典学习(DLSI)算法和稀疏表示的分类算法(SRC),可以看出 Gabor 特征集结合判别式字典学习的稀疏表示图像识别算法优于其他算法的识别效果,比 SRC 算法性能提升 2%。

表 3 USPS 手写数字库上几种方法的识别错误率

Table 3 The recognition results for USPS

	K_NN	SVM-Gauss	SRSC	DLSI	SRC	本文方法
错误率/%	5.2	4.2	6.05	3.98	5.28	3.27

4 结 论

考虑到图像的 Gabor 特征对光照、表情和姿态等变化的鲁棒性,提出 Gabor 特征集结合判别式字典学习的稀疏表示图像识别算法。首先提取图像多

尺度多方向的 Gabor 特征作初始字典,然后通过判别保真项和 Fisher 准则使该字典更具判别性。该算法的判别性能主要体现在特征字典中的子字典对与其相对应类别的样本特征具有较强的表示能力,从而可以减小重构误差,通过最小化类内散度和最大化类间散度使编码系数具有判别性,每类的特征差异性更易于区分;同时用具有判别性的重构误差和稀疏编码系数来对输入的待查询图像进行分类的策略,提高了模式分类的准确率,实验结果表明本文方法具有较好的效果。

参考文献 (References)

- [1] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2) : 210-227.
- [2] Yang M, Zhang L, Yang J, et al. Metaface learning for sparse representation based face recognition [C] // Proceedings of the International Conference on Image Processing. Hong Kong, China; IEEE Signal Processing Society, 2010 : 1601-1604.
- [3] Ramírez I, Sprechmann P, Sapiro G. Classification and clustering via dictionary learning with structured incoherence and shared features [C] // Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. California, USA; IEEE Computer Society, 2010 : 3501-3508.
- [4] Zhang Q, Li B X. Discriminative K-SVD for dictionary learning in face recognition [C] // Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. California, USA; IEEE Computer Society, 2010 : 2691-2698.
- [5] Mailhé B, Plumbley M D. Dictionary learning with large step gradient descent for sparse representations [C] // Proceedings of International Conference on Latent Variable Analysis and Signal Separation. Tel-Aviv, Israel; Springer, 2012 : 231-238.
- [6] Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. IEEE Transactions on Signal Processing, 2006, 54 (11) : 4311-4322.
- [7] Jiang Z L, Lin Z, Davis L S. Learning a discriminative dictionary for sparse coding via label consistent K-SVD [C] // Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Colorado Springs, USA; IEEE Computer Society, 2011 : 1697-1704.
- [8] Ramirez I, Sapiro G. Sparse coding and dictionary learning based on the MDL principle [C] // Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. Prague, Czech Republic; IEEE, 2011 : 2160-2163.
- [9] Yang M, Zhang L, Feng X C, et al. Fisher discrimination dictionary learning for sparse representation [C] // Proceedings of IEEE International Conference on Computer Vision. Barcelona, Spain; IEEE, 2011 : 543-550.
- [10] Porat M, Zeevi Y. The generalized Gabor scheme of image representation in biological and machine vision [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1988, 10(4) : 452-468.
- [11] Wiskott L, Fellous J M, Kruger N, et al. Face recognition by elastic bunch graph matching [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1997, 19(7) : 775-779.
- [12] Liu C J, 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.
- [13] Zhang W C, Shan S G, Zhang H M, et al. Histogram sequence of local Gabor Binary Pattern for face description and identification [J]. Journal of Software, 2006, 17(12) : 2508-2517. [张文超, 山世光, 张洪明, 等. 基于局部 Gabor 变化直方图序列的人脸描述与识别 [J]. 软件学报, 2006, 17(12) : 2508-2517.]
- [14] Yang M, Zhang L. Gabor feature based sparse representation for face recognition with Gabor occlusion dictionary [C] // Proceedings of the 11th European Conference on Computer Vision. Crete, Greece; Springer, 2010 : 448-461.