

基于 Gabor 特征和字典学习的高斯混合 稀疏表示图像识别

詹 曙¹, 王 俊¹, 杨福猛², 方 琪¹

(1. 合肥工业大学计算机与信息学院, 安徽合肥 230009; 2. 三江学院电子信息工程学院, 江苏南京 210012)

摘 要: 为了克服图像识别中光照、姿态等变化带来的识别困难, 同时提高稀疏表示图像识别的鲁棒性, 本文提出了一种基于 Gabor 特征和字典学习的高斯混合稀疏表示图像识别算法. 高斯混合稀疏表示是基于最大似然估计准则, 将稀疏保真度表示为余项的最大似然函数, 最终识别问题转化为求解加权范数的优化逼近问题. 本文算法首先提取图像的 Gabor 特征; 然后对 Gabor 特征集进行字典学习, 由于在学习过程中引入了 Fisher 准则作为约束, 学习得到具有类别标签的新字典; 最后使用高斯混合稀疏表示识别方法进行分类识别. 在 3 个公开数据库(人脸数据库 AR 库和 FERET 库以及 USPS 手写数字库)上的实验结果验证了该算法的有效性和鲁棒性.

关键词: Gabor 特征; 稀疏表示; fisher 字典学习; 最大似然估计

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2015)03-0523-06

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2015.03.017

Gaussian Mixture Sparse Representation for Image Recognition Based on Gabor Features and Dictionary Learning

ZHAN Shu¹, WANG Jun¹, YANG Fu-meng², FANG Qi¹

(1. School of Computer & Information, Hefei University of Technology, Hefei, Anhui 230009, China;

2. School of Electronic Information Engineering, Sanjiang University, Nanjing, Jiangsu 210012, China)

Abstract: To overcome the problems of the illumination and pose variations in image recognition, the algorithm of Gaussian mixture sparse representation for image recognition based on dictionary learning and Gabor features is proposed. Based on the maximum likelihood estimation principle, a mixture Gaussian sparse coding model is proposed to express the discriminating items to the maximum likelihood function of residuals, so the problem of identification is converted to the optimal weighted norm approximation problem. This approach extracts the Gabor features of the images by the Gabor filter, and then uses the Gabor features to learn a new dictionary. As the Fisher criterion is added in the learning process as a constraint, a new dictionary with category labels can be obtained. Finally, the method of Gaussian mixture sparse representation is used for classification and identification. The experimental results in three public databases demonstrate that the algorithm proposed is effective and robust.

Key words: Gabor features; sparse representation; fisher dictionary learning; maximum likelihood estimation

1 引言

近年来, 图像识别目前已成为计算机视觉和模式识别领域的热门课题, 具有广阔的应用前景. 然而, 图像识别仍有许多问题远未得到解决, 例如光照、表情和姿态等变化都在一定程度上限制了该研究的发展, 因此如何处理这些问题是目前该领域研究的关键和难点.

在图像识别的众多研究方法中, 稀疏表示的分类思想^[1]已成功在图像识别领取得了重要的地位. 基于稀疏

表示的图像分类是可以用少量同一类的低维图像编码或者表示高维图像; 主要有两个阶段: 稀疏表示和分类识别. 首先, 通过字典原子与一些稀疏性约束对测试图像进行表示, 然后在稀疏表示系数和字典的基础上进行分类识别. 人脸识别方面, 2009 年, Wright^[2]等提出一个基于稀疏表示的分类器 (Sparse Representation Classification, SRC), 将原始的训练人脸图像作为字典, 通过范数求解测试样本的稀疏系数, 通过该系数对测试人脸进行重构, 进而求出其残差, 将其归为残差最小类, 取得了

不错的分类效果.

由于原始的训练图像中可能存在的不确定性和噪声干扰,使得这样的字典不是非常有效.近几年国内外的学者提出了很多过完备字典学习的方法^[3,4],目的是从训练样本中得到一组基能够更好的对测试样本进行表示或者编码.例如 Elad, M^[5]等提出的 KSVD 算法,是目前用于计算稀疏表示超完备字典效果较好的算法,该算法的字典更新对字典原子逐一更新,综合考虑了字典原子对每一个训练样本的影响,增强了原子的描述能力.Zhang^[6]等提出 D-KSVD(Discriminative K-SVD)算法,将稀疏系数变换到类别间差异更加突出的空间,其分类器在 K-SVD 训练字典的同时得到训练,使得超完备字典和分类器的性能都得到了提高,从而取得了较理想效果.Yang^[7]等提出 FDDL 算法,将 Fisher 判别准则融入到字典学习中,训练得到一个结构性字典,同类样本在此字典下的表示误差较小,而对不同类别样本的表示误差较大,提高了分类的准确性.

上述的研究都是对原始的训练样本直接进行字典学习,虽然这些算法都取得了较好的识别效果,但是这些算法是基于全局特征进行处理,没有考虑使用局部特征信息;在现实中由于受样本数目的限制,全局特征并不能有效的应对光照、表情以及姿态等变化因素.因此,以局部特征为研究方向的分类识别算法引发了国内外学者的关注,例如 Gabor 小波^[8~11]能够很好地提取目标图像的不同空间位置、方向和频率上的特征,由于该特征从图像的局部提取,能够更好的克服光照、姿态和表情等全局干扰对识别效果的影响.同时,传统的稀疏表示算法未能很好地运用判别性的信息,而将判别性的信息应用于稀疏编码系数可以使得训练得到的字典,具有小的类内散度和大的类间散度从而获得很强的鉴别能力.结合上述分析,受文献[7]的算法启发,本文提出基于 Gabor 特征和字典学习的高斯混合稀疏表示图像识别算法;该算法主要是使用 Gabor 特征取代原来的全局特征进行字典学习,并将高斯混合稀疏表示模型作为分类策略.

2 稀疏表示分类方法

2.1 稀疏表示分类

稀疏表示分类是指将测试图像表示为训练样本的线性组合,然后用 l_1 或 l_2 范数对保真度项进行编码.文献[2]提出的传统的稀疏表示算法是假设有 k 类样本,令 $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_k]$ 表示训练样本集,其中 \mathbf{A}_i 是第 i 类的训练样本子集.令 \mathbf{y} 代表测试样本,将测试图像 \mathbf{y} 表示为训练样本的线性组合,即 $\mathbf{y} = \mathbf{A}\boldsymbol{\alpha}$,其中 $\boldsymbol{\alpha} = [\alpha_1; \dots; \alpha_i; \dots; \alpha_k]$. SRC 算法过程如下:

(1)将测试样本表示为字典 \mathbf{A} 的线性组合,由 l_1 范

数最小化求解得到稀疏系数:

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

式中 λ 为一个标量;

(2)计算各类样本对测试样本的逼近残差:

$$e_i = \|\mathbf{y} - \mathbf{A}\delta_i(\hat{\boldsymbol{\alpha}})\|, \quad i = 1, \dots, k \quad (2)$$

其中, $\delta_i(\hat{\boldsymbol{\alpha}})$ 是与第 i 类样本对应的系数向量;

(3)根据最小逼近残差将测试图像分类:

$$\text{identity}(\mathbf{y}) = \arg \min_i (e_i(\mathbf{y})) \quad (3)$$

2.2 高斯混合稀疏表示模型

上述稀疏表示方法中的编码模型是假定残差余项服从高斯分布或拉普拉斯分布,但是这样的假设在当测试图像出现异常像素(如被部分遮挡、受噪声干扰或局部形状变化)时可能并不成立,所以稀疏表示模型在上述情况下的有效性和鲁棒性将降低.因此,文献[12]中提出了基于最大似然估计理论的高斯混合稀疏表示模型.

首先将训练样本字典改写为 $\mathbf{A} = [\mathbf{a}_1; \mathbf{a}_2; \dots; \mathbf{a}_n]$,其中行向量 \mathbf{a}_i 表示 \mathbf{A} 的第 i 行,则余项表示为 $\mathbf{e} = \mathbf{y} - \mathbf{A}\boldsymbol{\alpha} = [e_1; e_2; \dots; e_n]$,其中 $e_i = y_i - \mathbf{a}_i\boldsymbol{\alpha}$, $i = 1, 2, \dots, n$.假设 e_1, e_2, \dots, e_n 独立同分布,设其概率密度函数为 $f_\theta(e_i)$, θ 表示该分布的参数设置.在不考虑稀疏性表示系数 $\boldsymbol{\alpha}$ 约束的情况下,余项 \mathbf{e} 的最大似然估计函数为:

$$L_\theta(e_1, e_2, \dots, e_n) = \prod_{i=1}^n f_\theta(e_i) \quad (4)$$

最大似然估计理论的目的是最大化式(4)的似然函数,即最小化式(5)的目标函数:

$$-\ln L_\theta = \sum_{i=1}^n \rho_\theta(e_i) \quad (5)$$

式中, $\rho_\theta(e_i) = -\ln f_\theta(e_i)$.

如果加入稀疏性表示系数 $\boldsymbol{\alpha}$ 约束的情况,则高斯混合稀疏表示模型定义为:

$$\min_{\boldsymbol{\alpha}} \sum_{i=1}^n \rho_\theta(y_i - \mathbf{a}_i\boldsymbol{\alpha}) \quad \text{s.t.} \|\boldsymbol{\alpha}\|_1 \leq \sigma \quad (6)$$

该模型本质上是一个稀疏约束最大似然估计问题.令

$$F_\theta(\mathbf{e}) = \sum_{i=1}^n \rho_\theta(e_i), \text{ 可以将式(6)写成更一般的形式:} \\ \min \|\mathbf{F}_\theta(\mathbf{y} - \mathbf{A}\boldsymbol{\alpha})\|_2^2 \quad \text{s.t.} \|\boldsymbol{\alpha}\|_1 \leq \delta \quad (7)$$

3 基于 Gabor 特征和字典学习的高斯混合稀疏表示图像识别算法

由于当测试图像存在部分遮挡、噪声干扰或者局部变形时^[13~15],传统的稀疏表示识别方法通过简单的用 l_1 或 l_2 范数不能够完全准确的来表示残差的分布概率情况,在一定程度上缺少鲁棒性.考虑到 Gabor 特征从图像局部提取,对光照、表情和姿态变化相对于整体特征更具鲁棒性;同时考虑到字典的选择构造尤为

重要,无论是在图像重建和分类识别^[16]的过程中都起着非常关键的作用.综合以上多种考虑,本文提出基于 Gabor 特征和字典学习的高斯混合稀疏表示图像识别算法.图 1 是本文算法的流程框图.

该算法主要思想:首先是提取原始图像的 Gabor 特征,然后将 Gabor 特征集作为初始字典,通过对初始字典的学习得到具有类别标签的新字典,新字典由于引入 Fisher 判别约束编码系数,具有小的类内散度和大的类间散度,使学习得到的字典具有很强的鉴别能力;最后,使用高斯混合稀疏表示识别方法,将稀疏表示的保真度表示为余项的最大似然函数,将识别问题转化为加权范数的优化逼近问题,对不同的余项赋予不同的权值,用迭代重加权稀疏编码算法进行求解.



图1 本文提出算法流程

3.1 基于 Gabor 特征集和 FISHER 准则的字典学习

Gabor 滤波器定义为:

$$\phi_{\mu,v}(z) = \frac{\|k_{\mu,v}\|^2}{\sigma^2} e^{(-\|k_{\mu,v}\|^2 \|z\|^2 / 2\sigma^2)} [e^{ik_{\mu,v}z} - e^{-\sigma^2/2}] \quad (8)$$

式中, $z = (x, y)$ 代表像素, μ 和 v 分别表示方向和尺度; $k_{\mu,v} = k_v e^{i\phi_{\mu,v}}$, 其中 $k_v = k_{\max}/f^v$, $\phi_{\mu} = \pi u/8$, 最大频率 $k_{\max} = \pi/2 \cdot f$ 是频域的核间距因子, σ 表示高斯窗宽度与波长比.

图像的 Gabor 特征 $G_{\mu,v}(z)$ 是将图像 $I(z)$ 与 Gabor 滤波器 $\phi_{\mu,v}(z)$ 做卷积实现:

$$G_{\mu,v}(z) = M_{\mu,v}(z) \cdot \exp(j\theta_{\mu,v}(z))$$

其中, $M_{\mu,v}(z)$ 表示幅度信息, 包含图像局部能量; $\theta_{\mu,v}(z)$ 表示相位信息. 为了便于描述图像的局部特征, 通常采用 5 个尺度 $v = \{0, 1, 2, 3, 4\}$ 和 8 个方向 $\mu = \{0, 1, 2, 3, 4, 5, 6, 7\}$ 组成的 40 个 Gabor 滤波器, 因此将 $\chi = [a_{0,0}^T, a_{0,1}^T, \dots, a_{4,7}^T]^T$ 作为图像的 Gabor 描述.

为了得到一个性能更好的字典, 本文将原始图像提取的 Gabor 特征集作为初始字典进行学习. 令 $G = [G_1, G_2, \dots, G_k]$ 表示 k 类训练样本的 Gabor 特征集, G_i 是第 i 类样本的 Gabor 特征子集. 令字典学习后得到的新字典为 D , 即 $G \approx DX$, 其中 $D = [D_1, \dots, D_i, \dots, D_k]$, D_i 是第 i 类的子字典; X 为字典 D 对 Gabor 特征集 G 表示的稀疏系数矩阵, 即 $X = [X_1, \dots, X_i, \dots, X_k]$, X_i 是第 i 个稀疏系数的子矩阵.

假设 G_i 在字典 D 上稀疏系数为 $X_i = [X_i^1; \dots; X_i^j; \dots; X_i^c]$, X_i^j 是 G_i 在子字典 D_i 上的稀疏系数, 将判别项 r 构造为:

$$r = \|G_i - DX_i\|_F^2 + \|G_i - D_i X_i^j\|_F^2 + \sum_{\substack{j=1 \\ j \neq i}}^c \|D_j X_i^j\|_F^2 \quad (9)$$

式中, $\|G_i - DX_i\|_F^2$ 是整个字典 D 对于 G_i 的表示能力; 其次为了强化子字典 D_i 对同类特征 G_i 的鉴别表示能力, 我们添加 $\|G_i - D_i X_i^j\|_F^2$ 使其值足够的小; 同时为了弱化其他子字典 $D_j (i \neq j)$ 对 G_i 的表示能力, 期望 X_i^j 是稀疏矩阵, 使得 $\|D_j X_i^j\|_F^2$ 很小. 因此, 判别项 r 使学习到的字典可以大大的增强对同一类特征的表示能力.

同时, 在字典学习过程中通过对稀疏表示的系数 X 引入 Fisher 判别准则来构造鉴别性约束项 $d(X)$, 对 X 作最小化类内散度 $S_B(X)$ 和最大化类间散度 $S_w(X)$ 来使学习得到的字典 D 具有可鉴别性. 将 $S_B(X)$ 和 $S_w(X)$ 分别定义为:

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

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

其中, m_i 和 m 分别是 X_i 和 X 的均值向量, n_i 是第 i 类样本的个数.

为了使类内特征更紧凑, 类间特征差异变大, 定义鉴别性约束为:

$$d(X) = \text{tr}(S_w(X)) - \text{tr}(S_B(X)) + \eta \|X\|_F^2 \quad (10)$$

其中 η 是一个标量参数. 通过结合式(9)和式(10)得到目标函数:

$$J(D, X) = \arg \min_{D, X} \left\{ \begin{aligned} &\sum_{i=1}^k r(G_i, D, X_i) + \lambda_1 \|X\|_1 \\ &+ \lambda_2 (\text{tr}(S_w(X)) - \text{tr}(S_B(X))) \\ &+ \eta \|X\|_F^2 \end{aligned} \right\} \quad (11)$$

式(11)的目标函数的求解是一个组合迭代优化的过程, 通过分别固定 D 和 X 进行迭代优化获得所需的新字典 D 和鉴别系数 X .

3.2 基于高斯混合稀疏表示的分类方法

将经过本节 3.1 字典学习后得到的具有鉴别性的新字典 D 用于高斯混合稀疏表示模型, 求解编码系数和重构残差. 在本文算法中, 式(7)进一步写为下式:

$$\min \|W^{1/2}(y - D\alpha)\|_2^2 \quad \text{s.t.} \quad \|\alpha\|_1 \leq \sigma \quad (12)$$

由式(7)和式(12)比较可得, 本文的分类方法将传统的稀疏表示模型转化为一个加权的稀疏表示模型, 式(12)是一个加权范数逼近求解问题, 此处的 W 定义为对角的权值矩阵, 赋予不同的残差余项不同的权值, 即 $W_{i,i}$ 表示分配给每一个特征点的不同权值. 通常情况下, 由于异常的 Gabor 特征点会导致较大的残差; 因

此,应该赋予其较低的权值,才能获得较好的鲁棒性.基于这样的情况,选择与其有类似性质的 SVM 的 hingeloss 函数作为权值函数,即:

$$W_{i,i} = \omega_{\theta}(e_i) = \frac{\exp(\mu\delta - \mu e_i^2)}{(1 + \exp(\mu\delta - \mu e_i^2))} \quad (13)$$

式中, μ 控制从 1 到 0 的下降速度, δ 控制分界点.

最后,结合式(12)和式(13)得到求解稀疏表示系数的目标函数:

$$\hat{\alpha} = \arg \min \|W^{1/2}(y - D\alpha)\|_2^2 \quad \text{s.t.} \|\alpha\|_1 \leq \sigma \quad (14)$$

对于式(14),通过不断迭代优化更新权值矩阵 W ,直至收敛,求得最优的稀疏表示系数,根据式(2)求出残差余项,最后由最小逼近残差准则完成测试图像的分类.

4 实验结果

为了验证本文算法的有效性,我们在 AR 库^[17]和 FERET 人脸数据库^[18]以及 USPS 手写数据库进行了一系列的实验.实验结果均为在多次实验后得到的平均识别率,以上实验在 Inter(R) core(TM) i5-4130, CPU 主频是 3.4GHz, 4G 内存的 win7 操作系统, matlab2012b 环境下运行.

4.1 AR 人脸库

本实验主要针对光照变化问题,AR 人脸库包含 126 个人的 4000 多幅正面人脸图像,从中选择 50 名男性和 50 名女性,每人 14 幅图像,前 7 幅用于训练,后 7 幅用于测试.实验数据库中图像包含光照变化和表情变化,所有图像的尺寸裁剪为 83×60 ,图像都经过归一化处理,采用 PCA 降维后维数 $d = 300$,本文算法参数 $\lambda_1 = 0.005$, $\lambda_2 = 0.005$.

表 1 分别比较文献[4]中判别 KSVD(DKSVD)、文献[2]中稀疏表示分类算法(SRC)、文献[6]中判别字典学习的稀疏表示(FDDL)算法、基于 Gabor 特征的 FDDL 算法(GFDDL)和本文算法的正确识别率,同时给出了各算

法的在识别阶段的时间运行情况.

表 1 AR 人脸数据库算法识别结果比较

algorithm	recognition rate(%)	time(s)
DKSVD	85.4	526
SRC	88.8	619
FDDL	92.0	635
GFDDL	93.7	628
本文算法	96.2	1016

由于 AR 人脸库中图像存在丰富的光照、表情等变化, SRC 等算法只是简单的把全部训练样本用来构造字典,这样的字典没有充分利用已知的样本类别信息; FDDL 算法虽然在字典学习中添加了判别性约束条件,但是采用整体特征的算法仍不能有效的处理这些问题.本文算法由于使用 Gabor 特征和加入 Fisher 准则构造字典,不但使字典具有更强的表示能力,而且具有较强的鉴别性.同时为了验证高斯混合稀疏模型在分类时比 SRC 模型更具鲁棒性的,本文还做了 GFDDL 算法作为对比实验,实验结果表明本文方法能够更好的处理光照和表情变化的干扰,验证了本文算法的可行性和有效性.

4.2 FERET 人脸数据库

本实验主要针对姿态变化问题,选择 FERET 库的一个姿态库作为实验数据,该姿态库由 200 个人的 1400 幅图像组成,每人 7 幅图像,其中 3 幅为正面图像,4 幅为姿态偏转图像,分别标记为 ba、bd、be、bf、bg、bj 和 bk. 在实验中,所有图像的尺寸归一化为 80×80 . 我们总共做了 5 组不同姿态角度的实验,在实验 1(姿态角度为 0 度)中,每一类图像中标记为 ba、bj 的图像作为训练集,标记为 bk 的图像作为测试集; 剩余的四组实验中,用标记为 ba、bj 和 bk 的图像作为训练集,然后分别用 bg(-25°)、bf(-15°)、be(+15°)、bd(+25°)作为测试集,得到不同姿态角度下的识别结果. 在实验中,用 PCA 降到 350 维,参数设置为 $\lambda_1 = 0.005$, $\lambda_2 = 0.005$. 来自同一人的 人脸图像如图 2 所示.



图2 FERET库同一人的不同姿态角度的人脸图像

图 3 画出了几种不同算法分别在不同姿态角度下的正确识别率,本文提出的算法明显高于 SRC 以及其他的算法.表 2 给出了各算法在实验 1 中不同姿态角度下的运行时间情况.从图中可以看出,当姿态角度偏转适中的情况(0° 和 $\pm 15^\circ$)下, SRC 对姿态变化很敏感,本文算法的相比于 SRC 算法,识别率至少提高了 12%. 随着偏转角度的增加,当姿态变化较大($\pm 25^\circ$)时,各算法

的识别率都在下降,但本文的算法仍然高于其他算法. 由于 Gabor 特征描述图像局部信息的优越性,因此与单纯的采用判别式字典学习(FDDL)算法相比, GFDDL 算法表明 Gabor 局部特征对姿态变化有较好的鲁棒性,显著地提高了识别率.同时由于高斯混合稀疏表示对局部变形的不敏感性,本文算法相比其他算法更好的克服了姿态变化的影响.实验结果表明本文算法在姿态

变化适中的情况下显著的提高了识别效果,在实际中也具有一定的应用价值.

表 2 FERET 库在实验 1(姿态角度为 0 度)算法时间对比

algorithm	time(s)
SRC	203
FDDL	195
GFDDL	217
本文算法	490

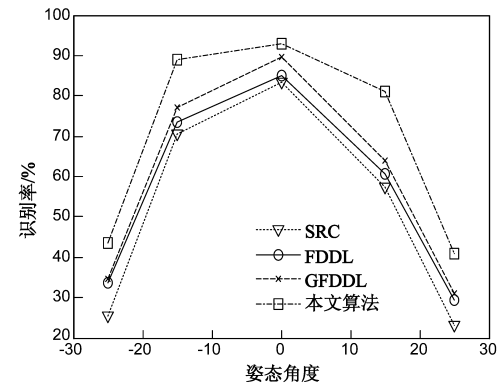


图3 FERET人脸数据库算法识别结果比较

4.3 USPS 手写数字库

美国国家邮政局手写数字数据库(USPS)包含 0~9 共 10 个数字的手写体图像,其中训练图像 7291 幅,测试图像 2007 幅.图 4 是该库 3 组不同字符的图像.实验从每一类数字分别随机选取 100 幅作为训练集,100 幅作为测试集.图像尺寸大小为 16×16,图像均经过归一化处理,用 PCA 变换降维后维数 $d=120$,参数设置为 $\lambda_1=0.005$, $\lambda_2=0.05$.



图4 USPS手写数字库样本

表 3 给出了本文方法与其他几种方法识别效果的比较以及各算法的在识别阶段的时间消耗情况.由于书写风格的不同造成了各种字符变形,使字符的模式特征变得不稳定和极其复杂,从图 3 可以看出该字库中各字符笔画的形态、粗细和灰度等差别都很显著.与 SRC 算法相比,本文结合 Gabor 特征和 Fisher 准则学习得到的字典能够很好的提取每一类的判别性信息,具有更好的鉴别稀疏表示能力;通过对比 GFDDL 算法实验证明了高斯混合稀疏模型对各种字体的形态变化相比 SRC 模型具有更好的鲁棒性.可以看出,相比其他算法,本文方法取得了很好的识别效果,证明了算法的合理性和有效性.

表 3 USPS 手写数字库算法识别结果比较

algorithm	recognition rate(%)	time(s)
SRSC	91.8	241
SRC	92.1	278
FDDL	90.1	243
GFDDL	93.1	265
本文算法	95.9	485

为了充分评价算法的性能,在实验中记录了各算法的在识别阶段所消耗的时间,下面对各算法的复杂度进行分析.由于 SRC 算法只是简单的将全部训练样本用于构造字典,而 FDDL,GFDDL 以及本文算法需要进行 Gabor 特征的提取以及字典学习过程,所以 FDDL,GFDDL 和本文算法计算时间消耗相对要大;但显然这种情况出现在训练阶段,因此我们在这里比较各算法在测试识别阶段的时间复杂度情况.在识别阶段,上面算法耗时主要是集中在稀疏编码的过程,即求解式(2)或者式(14).假设训练字典原子的个数为 m ,式(2)的求解(即传统的稀疏表示分类)通常采用 l_1 正则化稀疏编码方式,它的时间复杂度为 $O(m^\epsilon)$,一般 $\epsilon \approx 1.5$,SRC、FDDL 和 GFDDL 算法在识别阶段采用传统的稀疏分类方法,只需要执行一次稀疏编码过程,然后利用残差来实现分类;而高斯混合稀疏分类算法需要几次迭代过程才完成编码过程,即时间复杂度为 $O(k(m^\epsilon))$, k 为迭代次数(本文实验 k 设置为 2).在这种情况下,本文算法的时间复杂度要高于前面的算法,大约是它们的 k 倍.由于本文算法存在 2 次迭代,所以识别的时间是它们的 2 倍左右,从各组实验结果也得到了验证.

5 总结

针对图像识别过程中存在的光照,表情和姿态变化等问题,考虑到 Gabor 特征相比于整体特征更具鲁棒性,本文算法对提取出的 Gabor 特征进行字典学习,在字典学习过程中引入 Fisher 准则施加在稀疏表示系数上,最后采用高斯混合稀疏表示算法进行分类识别.在国际公开的图像数据库上的实验结果表明,相对于目前比较流行的稀疏表示算法,本文算法在处理光照、姿态变化等方面具有更好的识别效果.

参考文献

[1] Huang K, Aviyente S. Sparse representation for signal classification[A]. Proceedings of Advances in Neural Information Processing Systems [C]. Vancouver, Canada: The MIT Press, 2006. 609 – 616.

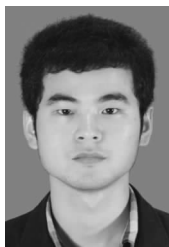
[2] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2009, 31(2): 210 – 227.

- [3] Zhuolin Jiang, Zhe Lin, Larry S Davis. Label consistent K-SVD: Learning a discriminative dictionary for recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(11): 2651 – 2664.
- [4] 朱杰, 杨万扣, 唐振民. 基于字典学习的核稀疏表示人脸识别方法[J]. 模式识别与人工智能, 2012, 25(5): 859 – 864.
ZHU Jie, YANG Wan-kou, TANG Zhen-Min. A dictionary learning based kernel sparse representation method for face recognition[J]. Pattern Recognition and Artificial Intelligence, 2012, 25(5): 859 – 864. (in Chinese)
- [5] Elad M, Aharon M. Image denoising via sparse and redundant representations over learned dictionaries[J]. Image Processing, IEEE Transactions on, 2006, 15(12): 3736 – 3745.
- [6] Zhang Q, Li B. Discriminative K-SVD for dictionary learning in face recognition[A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. San Francisco, CA: IEEE, 2010. 2691 – 2698.
- [7] Yang M, Zhang L, Feng X, et al. Fisher discrimination dictionary learning for sparse representation[A]. Proceedings of IEEE International Conference on Computer Vision (ICCV) [C]. Barcelona, Spain: IEEE, 2011. 543 – 550.
- [8] Yang M, Zhang L. Gabor feature based sparse representation for face recognition with gabor occlusion dictionary[A]. Proceedings of Computer Vision-ECCV 2010 [C]. Crete, Greece: Springer Berlin Heidelberg, 2010. 448 – 461.
- [9] 王科俊, 邹国锋. 基于子模式的 Gabor 特征融合的单样本人脸识别[J]. 模式识别与人工智能, 2013, V26(1): 50 – 56.
WANG Ke-Jun, ZOU Guo-Feng. A sub-pattern Gabor features fusion method for single sample face recognition[J]. Pattern Recognition and Artificial Intelligence, 2013, V26(1): 50 – 56. (in Chinese)
- [10] 周家锐, 纪震, 等. 基于 Gabor 小波与 Memetic 算法的人脸识别方法[J]. 电子学报, 2012, 40(4): 642 – 646.
ZHOU Jia-rui, JI Zhen, et al. Face recognition using Gabor wavelets and Memetic algorithm[J]. Acta Electronic Sinica, 2012, 40(4): 642 – 646. (in Chinese)
- [11] 詹曙, 张启祥, 等. 基于 Gabor 特征核协作表达的三维人脸识别[J]. 光子学报, 2013, 42(12): 1448 – 1453.
ZHAN Shu, ZHANG Qi-xiang. 3D face recognition by kernel collaborative representation based on gabor feature[J]. Acta Photonica Sinica, 2013, 42(12): 1448 – 1453. (in Chinese)
- [12] Yang M, Zhang L, Yang J, et al. Robust sparse coding for face recognition[A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Colorado Springs: IEEE, 2011. 625 – 632.
- [13] Ou W, You X, Tao D, et al. Robust face recognition via occlusion dictionary learning[J]. Pattern Recognition, 2014, 47(4): 1559 – 1572.
- [14] Deng W, Hu J, Guo J. Extended SRC: Undersampled face recognition via intraclass variant dictionary[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2012, 34(9): 1864 – 1870.
- [15] 平强, 庄连生, 俞能海. 姿态和光照可变条件下的仿射最小线性重构误差人脸识别算法[J]. 电子学报, 2013, 40(10): 1965 – 1970.
PING Qiang, ZHUANG Liansheng. Affine minimum linear reconstruction error face recognition under varying pose and illumination[J]. Acta Electronic Sinica, 2013, 40(10): 1965 – 1970. (in Chinese)
- [16] 胡正平, 李静. 基于低秩子空间恢复的联合稀疏表示人脸识别算法[J]. 电子学报, 2013, 41(5): 987 – 991.
HU Zheng-ping, LI Jing. Face recognition of joint sparse representation based on low-rank subspace recovery [J]. Acta Electronic Sinica, 2013, 41(5): 987 – 991. (in Chinese)
- [17] Martinez A M. The AR Face Database[R]. Barcelona: Computer Vision Center, Universitat Autònoma de Barcelona, 1998.
- [18] Phillips P J, Wechsler H, Huang J, et al. The FERET database and evaluation procedure for face-recognition algorithms[J]. Image and Vision Computing, 1998, 16(5): 295 – 306.

作者简介



詹曙 男, 1968 年生于安徽合肥. 副教授、硕士生导师, 合肥工业大学计算机与信息学院. 研究方向为三维人脸识别和医学图像处理.
E-mail: shu_zhan@hfut.edu.cn



王俊 男, 1989 年生于安徽合肥. 硕士研究生, 合肥工业大学计算机与信息学院. 研究方向为数字图像处理和模式识别.
E-mail: wangjunhfut@163.com



杨福猛 男, 1968 年生于安徽和县. 副教授, 三江学院电子信息工程学院. 研究方向: 电子与通讯, 图像及信号处理等.