

基于 Gabor 低秩恢复稀疏表示分类的人脸图像识别方法

杜海顺, 张旭东, 金 勇, 侯彦东

(河南大学图像处理与模式识别研究所, 河南开封 475004)

摘 要: 针对含光照、表情、姿态、遮挡等误差或被噪声污染的人脸图像的识别问题, 本文提出一种基于 Gabor 低秩恢复稀疏表示分类的人脸图像识别方法. 该方法首先用低秩矩阵恢复算法求得训练样本图像对应的误差图像; 然后, 对每一个训练样本图像及其对应的误差图像进行 Gabor 变换, 得到相应的 Gabor 特征向量, 并将这些 Gabor 特征向量组成一个 Gabor 特征字典; 进而, 计算测试样本图像 Gabor 特征向量在该 Gabor 特征字典下的稀疏表示系数, 并用该稀疏表示系数和 Gabor 特征字典, 对测试样本图像的 Gabor 特征向量进行类关联重构, 同时计算相应的类关联重构误差. 最后, 根据测试样本图像 Gabor 特征向量的类关联重构误差, 实现对测试样本图像的分类识别. 在 CMU PIE、Extended Yale B 和 AR 数据库上的实验结果表明, 本文提出的人脸图像识别方法具有较高的识别率和较强的抗干扰能力.

关键词: 人脸图像识别; 稀疏表示; 低秩矩阵恢复; Gabor 变换

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 0372-2112 (2014)12-2386-08

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2014.12.008

Face Image Recognition Method via Gabor Low-Rank Recovery Sparse Representation-Based Classification

DU Hai-shun, ZHANG Xu-dong, JIN Yong, HOU Yan-dong

(Institute of Image Processing and Pattern Recognition, Henan University, Kaifeng, Henan 475004, China)

Abstract: To recognize the face images containing errors of illumination, expression, pose, occlusion, or contaminated by noise, we propose a face image recognition method via Gabor low-rank recovery sparse representation-based classification. In this method, we firstly obtain the error images of the training images using the low-rank matrix recovery algorithm, and then calculate the Gabor feature vectors of the training images and the corresponding error images via the Gabor transform algorithm. With these Gabor feature vectors, we constitute a Gabor feature dictionary. Based on the Gabor feature dictionary, we calculate the sparse representation coefficients of Gabor feature vector of the given test image. For each class, we use the sparse representation coefficients associated with the class and the Gabor feature dictionary to reconstruct the Gabor feature vector of the given test image. And then we calculate the reconstruction error between the Gabor feature vector and its approximation associated with the class. Based on the reconstruction errors associated with different class, we can accurately classify the given test image. Experimental results on CMU PIE, Extend Yale B and AR databases show that the proposed face image recognition method has a higher recognition rate and greater noise immunity.

Key words: face image recognition; sparse representation; low-rank matrix recovery; Gabor transformation

1 引言

近年来, 随着压缩感知理论和 l_1 范数最优化技术的发展, 稀疏表示受到国内外众多学者关注. 研究表明, 稀疏表示模型与人类视觉系统原理非常相似^[1]. 因此, 稀疏表示被广泛用于图像滤波^[2,3]、图像重构^[4,5]、图像

压缩^[6,7]等计算机视觉领域. 2009 年, John Wright 等学者将稀疏表示应用到人脸图像识别领域, 提出一种基于稀疏表示分类 (Sparse Representation-based Classification, SRC) 的人脸图像识别方法^[8]. 该方法在人脸图像受噪声污染或有其它误差干扰情况下, 取得了较好的识别效果. 然而, SRC 仍存在如下不足: (1) SRC 中用单位阵对

图像误差及噪声描述不够准确直观。(2)SRC 中用于描述图像误差和噪声的单位阵维数过高,从而使后续稀疏分解运算计算量很大。

为解决 SRC 中存在的问题, M. Yang 等学者提出一种基于 Gabor 特征稀疏表示分类 (Gabor-feature based SRC, GSRC) 的人脸图像识别方法^[9]。该方法在对 SRC 中描述图像误差和噪声的单位阵 Gabor 变换后, 进行误差字典学习, 得到了一个更加紧凑的 Gabor 误差字典, 从而使后续稀疏分解运算的计算量显著降低。Deng 等学者提出一种扩展 SRC (Extended Sparse Representation-based Classification, ESRC) 方法^[10]。该方法首先用人脸训练样本图像减去其对应的类均值图像得到误差图像, 然后用全部训练样本图像及其对应的误差图像共同组成字典, 并在该字典下得到人脸测试样本图像的稀疏表示, 从而实现测试样本图像的分类识别。由于 ESRC 在字典中使用了误差图像, 因此, 人脸测试样本图像在该字典下的稀疏表示更加准确, 从而提高了人脸图像的识别性能。然而, 研究发现, ESRC 所用的误差图像对人脸图像中的光照、姿态、表情、遮挡等误差描述并不准确。

近年, Z. Lin 等学者提出一种低秩矩阵恢复算法 (Low-rank Matrix Recovery, LR)^[11], 能将训练样本图像向量矩阵 D 分解为一个低秩逼近图像向量矩阵 A 和一个误差图像向量矩阵 E , 且 $D = A + E$ 。本文研究发现, LR 算法求得的误差图像能准确地描述人脸图像中的光照、姿态、表情、遮挡等误差和噪声。此外, 由于 Gabor 小波是对大脑视觉皮层神经元的很好逼近, 因此, 采用 Gabor 特征来处理图像具有良好的视觉特性和生物背景^[12]。基于此, 本文提出一种基于 Gabor 低秩恢复稀疏表示分类 (Gabor Low-rank Recovery Sparse Representation-based Classification, GLR_SRC) 的人脸图像识别方法。一方面, 由于 LR 算法求得的误差图像能准确地描述人脸图像中的误差及噪声, 因此, GLR_SRC 用全部人脸训练样本图像的 Gabor 特征向量和 LR 算法求得的误差图像的 Gabor 特征向量共同组成字典, 能够得到人脸测试样本图像 Gabor 向量的精确稀疏表示; 另一方面, 由于 Gabor 变换能够提取人脸图像不同方向和尺度下的特征, 所以这些特征更适用于图像稀疏表示^[10]。因此, 基于 GLR_SRC 的人脸图像识别方法具有较高的识别率和较强的抗干扰能力。

2 稀疏表示分类相关方法介绍

2.1 稀疏表示分类 (SRC) 方法

假设有 C 类训练样本图像, $D_i = [d_{i1}, d_{i2}, \dots, d_{in_i}] \in \mathbb{R}^{m \times n_i}$ 表示第 i 类训练样本图像向量构成的矩阵,

$d_{ij} \in \mathbb{R}^m$ 表示第 i 类第 j 个训练样本图像向量。其中, m 为训练样本图像向量的维数, n_i 为第 i 类训练样本的个数。 $D = [D_1, D_2, \dots, D_C] \in \mathbb{R}^{m \times n}$ 表示训练样本图像向量矩阵, 其中, n 为训练样本图像的总数, C 为训练样本图像的类别数, $n = \sum_{i=1}^C n_i$ 。令矩阵 D 为字典, 则给定测试样本图像向量 y 可由字典 D 表示, 即 $y = Dx$ 。其中, x 为测试样本图像向量 y 在字典 D 下的表示系数向量。SRC 认为, 测试样本图像向量 y 可由与它同类的训练样本图像向量充分表示, 并且在样本图像足够多的情况下, y 在 D 下的表示是稀疏的。同时, SRC 指出, y 的稀疏表示系数向量 x 越稀疏, 越有利于测试样本图像的分类^[8]。

SRC 通过求解如下 l_1 范数最优化问题, 得到 $y = Dx$ 的稀疏解 \hat{x}_1 :

$$\hat{x}_1 = \arg \min \|x\|_1 \quad \text{s.t.} \quad \|y - Dx\|_2 \leq \epsilon \quad (1)$$

迄今为止, 国内外学者针对求解 l_1 范数最优化问题提出了很多有效的算法, 具体可见 A. Yang 等学者在这方面所做的综述^[13]。

令 $\delta_i(\hat{x}_1)$ 为与 \hat{x}_1 同维, 且仅保留 \hat{x}_1 中与第 i 类对应的非零元素而其它元素为零的向量。基于 $\delta_i(\hat{x}_1)$, 重构与第 i 类关联的测试样本图像向量: $\hat{y}_i = D\delta_i(\hat{x}_1)$ 。更进一步, 可得 y 的类关联重构误差:

$$r_i(y) = \|y - D\delta_i(\hat{x}_1)\|_2, \quad i = 1, 2, \dots, C \quad (2)$$

基于 y 的类关联重构误差, SRC 给出如下分类规则:

$$\text{class}(y) = \arg \min_i r_i(y) \quad (3)$$

式(3)中, $\text{class}(y)$ 表示测试样本图像向量 y 的类别。

在实际应用中, 人脸图像一般会包含光照、表情、姿态、遮挡等误差, 另外还可能被噪声污染。为此, SRC 在字典中引入单位阵 $I \in \mathbb{R}^{m \times m}$ 来描述人脸图像中的误差和噪声。给定测试样本图像向量 y , 可由字典 $[D, I] \in \mathbb{R}^{m \times (n+m)}$ 稀疏表示, 即 l_1 范数最优化问题(1)转化为:

$$\begin{bmatrix} \hat{x}_1 \\ \hat{\beta}_1 \end{bmatrix} = \arg \min \left\| \begin{bmatrix} x \\ \beta \end{bmatrix} \right\|_1 \quad \text{s.t.} \quad \left\| y - [D, I] \begin{bmatrix} x \\ \beta \end{bmatrix} \right\|_2 \leq \epsilon \quad (4)$$

相似地, 通过求解 l_1 范数最优化问题(4)得到 \hat{x}_1 和 $\hat{\beta}_1$ 后, 采用式(5)计算 y 的类关联重构误差:

$$r_i(y) = \|y - D\delta_i(\hat{x}_1) - \hat{\beta}_1\|_2, \quad i = 1, 2, \dots, C \quad (5)$$

基于 $r_i(y)$, 采用式(3)的分类规则, 对人脸测试样本图像向量 y 进行分类。

2.2 扩展稀疏表示分类 (ESRC) 方法

SRC 虽然尝试通过用单位阵描述图像中的误差和噪声来解决人脸图像含误差或被噪声污染情况下的识

别问题,然而,单位阵对人脸图像中的误差和噪声描述不够准确直观.此外,由于 SRC 所用单位阵维数过高,因此,将单位阵作为字典的一部分会使后续的稀疏分解运算计算量很大.

为解决这一问题,Deng 等学者^[10]提出了一种扩展 SRC(ESRC)方法.该方法采用 $\mathbf{E} = [\mathbf{D}_1 - \mathbf{m}_1 \mathbf{e}_1, \dots, \mathbf{D}_C - \mathbf{m}_C \mathbf{e}_C] \in \mathbb{R}^{m \times n}$ 描述人脸图像中的误差和噪声.其中, \mathbf{D}_i 为第 i 类训练样本图像向量构成的矩阵; \mathbf{m}_i 为第 i 类训练样本图像的均值向量; \mathbf{e}_i 为 $[1, \dots, 1] \in \mathbb{R}^{1 \times n_i}$. ESRC 用训练样本图像向量矩阵 \mathbf{D} 和 \mathbf{E} 共同组成字典.给定测试样本图像向量 \mathbf{y} ,可由字典 $[\mathbf{D}, \mathbf{E}] \in \mathbb{R}^{m \times 2n}$ 稀疏表示.求解如下 l_1 范数最优化问题,可得 \mathbf{y} 在该字典下稀疏表示系数向量 $[\hat{\mathbf{x}}_1^T, \hat{\boldsymbol{\beta}}_1^T]^T$:

$$\begin{bmatrix} \hat{\mathbf{x}}_1 \\ \hat{\boldsymbol{\beta}}_1 \end{bmatrix} = \arg \min \left\| \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\beta} \end{bmatrix} \right\|_1 \quad \text{s.t.} \quad \left\| \mathbf{y} - [\mathbf{D}, \mathbf{E}] \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\beta} \end{bmatrix} \right\|_2 \leq \varepsilon \quad (6)$$

更进一步,ESRC 采用下式计算 \mathbf{y} 的类关联重构误差:

$$r_i(\mathbf{y}) = \left\| \mathbf{y} - \mathbf{D}\delta_i(\hat{\mathbf{x}}_1) - \mathbf{E}\hat{\boldsymbol{\beta}}_1 \right\|_2, \quad i = 1, 2, \dots, C \quad (7)$$

基于 $r_i(\mathbf{y})$,ESRC 同样采用式(3)的分类规则,对测试样本图像向量 \mathbf{y} 进行分类.

3 基于 Gabor 低秩恢复稀疏表示分类 (GLR_SRC) 的人脸图像识别方法

虽然 ESRC 所用误差图像比 SRC 采用单位阵对人脸图像中的误差和噪声描述更加直观,但 ESRC 所用误差图像对人脸图像中的光照、姿态、表情、遮挡等误差和噪声的描述并不准确.众所周知,低秩矩阵恢复(LR)算法^[11]能将训练样本图像向量矩阵 \mathbf{D} 分解为一个低秩逼近图像向量矩阵 \mathbf{A} 和一个误差图像向量矩阵 \mathbf{E} ,且 $\mathbf{D} = \mathbf{A} + \mathbf{E}$.本文研究发现,LR 算法求得的误差图像能准确地描述人脸图像中的误差和噪声.LR 采用增广拉格朗日乘子法求解如下最优化问题得到 \mathbf{A} 和 \mathbf{E} :

$$\min_{\mathbf{A}, \mathbf{E}} \|\mathbf{A}\|_* + \lambda \|\mathbf{E}\|_1 \quad \text{s.t.} \quad \mathbf{D} = \mathbf{A} + \mathbf{E} \quad (8)$$

具体求解步骤如算法 1^[11]所示.

算法 1 低秩矩阵恢复(LR)算法

输入:数据矩阵 \mathbf{D} ,参数 $\lambda > 0$

(1) $\mathbf{Y}_0 = \mathbf{D} / \max(\|\mathbf{D}\|_2, \lambda^{-1} \|\mathbf{D}\|_\infty)$; $\mu_0 > 0$; $\rho > 1$; $k = 0$.

重复

(2) $(\mathbf{U}, \mathbf{S}, \mathbf{V}) = \text{svd}(\mathbf{D} - \mathbf{E}_k + \mu_k^{-1} \mathbf{Y}_k)$;

(3) $\mathbf{A}_{k+1} = \mathbf{U} \mathbf{T}_{\mu_k^{-1}}[\mathbf{S}] \mathbf{V}^T$;

(4) $\mathbf{E}_{k+1} = \mathbf{T}_{\lambda \mu_k^{-1}}[\mathbf{D} - \mathbf{A}_{k+1} + \mu_k^{-1} \mathbf{Y}_k]$;

(5) $\mathbf{Y}_{k+1} = \mathbf{Y}_k + \mu_k(\mathbf{D} - \mathbf{A}_{k+1} - \mathbf{E}_{k+1})$;

$\mu_{k+1} = \rho \mu_k$;

(6) $k \leftarrow k + 1$;

直到收敛

输出:低秩逼近矩阵 \mathbf{A} 和误差矩阵 \mathbf{E}

其中,算法 1 步骤 3 和步骤 4 中的 $T_\varepsilon[\mathbf{X}]$ 为矩阵软阈值操作^[14],即:

$$T_\varepsilon[\mathbf{X}] = \begin{cases} \mathbf{X}_{ij} - \varepsilon, & \text{if } \mathbf{X}_{ij} > \varepsilon \\ \mathbf{X}_{ij} + \varepsilon, & \text{if } \mathbf{X}_{ij} < -\varepsilon \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

其中 $\varepsilon > 0$.

本文以 CMU PIE 数据库部分人脸图像为例,对比了 LR 求得的误差图像和 ESRC 所用误差图像.由图 1 可以看出,LR 求得的误差图像对人脸图像中的误差和噪声描述更加准确.

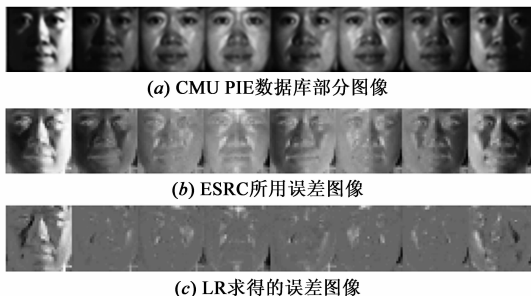


图1 人脸图像误差描述对比图

基于上述发现,本文首先对第 i 类训练样本图像向量矩阵 $\mathbf{D}_i = [\mathbf{d}_{i1}, \mathbf{d}_{i2}, \dots, \mathbf{d}_{in_i}] \in \mathbb{R}^{m \times n_i}$,采用算法 1 求得其低秩逼近图像向量矩阵 \mathbf{A}_i 和误差图像向量矩阵 \mathbf{E}_i ;然后,对每一类训练样本都做如此处理,可得训练样本图像向量矩阵 \mathbf{D} 的低秩逼近图像向量矩阵 $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_C] \in \mathbb{R}^{m \times n}$ 和误差图像向量矩阵 $\mathbf{E} = [\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_C] \in \mathbb{R}^{m \times n}$.其中 $\mathbf{A}_i = [\mathbf{a}_{i1}, \mathbf{a}_{i2}, \dots, \mathbf{a}_{in_i}] \in \mathbb{R}^{m \times n_i}$, \mathbf{a}_{ij} 为第 i 类第 j 个训练样本图像向量 \mathbf{d}_{ij} 的低秩逼近图像向量; $\mathbf{E}_i = [\mathbf{e}_{i1}, \mathbf{e}_{i2}, \dots, \mathbf{e}_{in_i}] \in \mathbb{R}^{m \times n_i}$, \mathbf{e}_{ij} 为第 i 类第 j 个训练样本图像向量 \mathbf{d}_{ij} 对应的误差图像向量.

由于 Gabor 小波在不同尺度和方向上对图像的代表特性与人类视觉系统的相关特性非常相似.因此, Gabor 小波提取的图像特征更适用于图像表示.二维 Gabor 小波变换函数定义如下^[15]:

$$\varphi_{\mu, \nu}(\mathbf{z}) = \frac{\|k_{\mu, \nu}\|^2}{\sigma^2} \exp\left(-\frac{\|k_{\mu, \nu}\|^2 \|\mathbf{z}\|^2}{2\sigma^2}\right) \times \left[\exp(i k_{\mu, \nu} \mathbf{z}) - \exp\left(-\frac{\sigma^2}{2}\right) \right] \quad (10)$$

其中, μ 和 ν 分别表示 Gabor 函数的方向和尺度.本文取 8 个方向和 5 个尺度,即 $\mu \in \{0, \dots, 7\}$, $\nu \in \{0, \dots, 4\}$.另取 $\sigma = \pi$.式中 $\mathbf{z} = (x, y)$ 为坐标, $\|\cdot\|$ 表示模运算, $\|\mathbf{z}\| = \sqrt{x^2 + y^2}$; $k_{\mu, \nu} \triangleq k_\nu \exp(i\phi_\mu)$, 其中 $k_\nu = 2^{-\frac{\nu+2}{2}} \pi$, ϕ_μ

$= \pi\mu/8$.

一幅图像的 Gabor 表示是由该图像与一组 Gabor 变换核卷积求得的. 假设 $I(z)$ 为一幅图像, 则它在 μ 方向和 v 尺度上的 Gabor 表示为:

$$L_{\mu,v} = I(z) * \varphi_{\mu,v}(z) \quad (11)$$

其中, $*$ 表示卷积运算. 那么, $S = \{L_{\mu,v}; \mu \in \{0, \dots, 7\}, v \in \{0, \dots, 4\}\}$ 为图像 $I(z)$ 的 Gabor 表示.

为了初步降低图像 Gabor 特征向量的维数, 本文首先对 $L_{\mu,v}$ 以采样因子 $\rho = 8 \times 8$ 为间隔进行下采样, 得到 $L_{\mu,v}^{\rho}$. 然后, 将 $L_{\mu,v}^{\rho}$ 按列连接起来得到向量 $l_{\mu,v}^{\rho}$, 并进一步对 $l_{\mu,v}^{\rho}$ 归一化处理, 即:

$$\bar{l}_{\mu,v}^{\rho} = \frac{l_{\mu,v}^{\rho} - \eta}{\delta} \quad (12)$$

式中 η 和 δ 分别为向量 $l_{\mu,v}^{\rho}$ 的均值和标准差. 从而得到图像 $I(z)$ 的最终 Gabor 特征向量为 $G = [\bar{l}_{0,0}^{\rho}; \dots; \bar{l}_{7,4}^{\rho}]$.

对任一人脸训练样本图像向量 d_{ij} 及其对应的误差图像向量 e_{ij} , 将它们转换成矩阵形式后, 进行上述 Gabor 变换, 可求得相应的 Gabor 特征向量 d_{ij}^g 和 e_{ij}^g , 进而可得全部训练样本图像及其对应误差图像的 Gabor 特征向量矩阵 $D^g = [d_1^g, d_2^g, \dots, d_n^g]$ 和 $E^g = [e_1^g, e_2^g, \dots, e_n^g]$. 同样地, 对给定人脸测试样本图像 y , 进行 Gabor 变换可求得其 Gabor 特征向量 y^g . 更进一步, 将 D^g 和 E^g 中的 Gabor 特征向量以及 y^g 进行 PCA 降维并按 l_2 范数归一化. 为方便叙述, 本文仍用 D^g 、 E^g 和 y^g 表示 PCD 降维归一化后的矩阵和向量.

由 D^g 和 E^g 共同组成字典 $[D^g, E^g]$, 则 y^g 可由字典 $[D^g, E^g]$ 稀疏表示. 求解如下 l_1 范数最优化问题, 可得到 y^g 在字典 $[D^g, E^g]$ 下的稀疏表示系数向量 $[\hat{x}_1^T, \hat{\beta}_1^T]^T$:

$$\begin{bmatrix} \hat{x}_1 \\ \hat{\beta}_1 \end{bmatrix} = \arg \min \left\| \begin{bmatrix} x \\ \beta \end{bmatrix} \right\|_1 \quad \text{s.t.} \quad \left\| y^g - [D^g, E^g] \begin{bmatrix} x \\ \beta \end{bmatrix} \right\|_2 \leq \epsilon \quad (13)$$

然后, 本文采用与 SRC 相同的方式定义 $\delta_i(\hat{x}_1)$. 并基于此, 对 y^g 进行类关联重构: $\hat{y}_i^g = D^g \delta_i(\hat{x}_1) + E^g \hat{\beta}_1$, $i = 1, 2, \dots, C$. 更进一步, 可得 y^g 的类关联重构误差:

$$r_i(y^g) = \|y^g - \hat{y}_i^g\|_2, \quad i = 1, 2, \dots, C \quad (14)$$

基于 $r_i(y^g)$, GLR_SRC 采用如下分类规则, 对人脸测试样本图像 y 进行分类:

$$\text{class}(y) = \arg \min_i r_i(y^g) \quad (15)$$

GLR_SRC 方法具体步骤如算法 2 所示.

算法 2 GLR_SRC 算法

输入: 训练样本图像向量矩阵 $D = [D_1, D_2, \dots, D_C] \in \mathbf{R}^{m \times n}$ 和测试样本图像向量 y , 参数 $\epsilon > 0$.

(1) 采用算法 1 计算 D_i 对应的 A_i 和 E_i , 进而得到 D 的低秩逼近图像

向量矩阵 A 和误差图像向量矩阵 E .

(2) 将 D 中图像向量 d_{ij} 和 E 中的误差图像向量 e_{ij} 转换成矩阵形式后进行 Gabor 变换, 得到相应的 Gabor 特征向量 d_{ij}^g 和 e_{ij}^g , 进而得到全部训练样本图像及其对应误差图像的 Gabor 特征向量矩阵 D^g 和 E^g .

(3) 将测试样本图像向量 y 转换成矩阵形式后, 进行 Gabor 变换得到其 Gabor 特征向量 y^g .

(4) 对 D^g 和 E^g 中的 Gabor 特征向量以及 y^g 进行 PCA 降维并按 l_2 范数归一化.

(5) 求解如下 l_1 范数最优化问题, 得到 y^g 在字典 $[D^g, E^g]$ 下的稀疏表示系数向量 $[\hat{x}_1^T, \hat{\beta}_1^T]^T$:

$$\begin{bmatrix} \hat{x}_1 \\ \hat{\beta}_1 \end{bmatrix} = \arg \min \left\| \begin{bmatrix} x \\ \beta \end{bmatrix} \right\|_1 \quad \text{s.t.} \quad \left\| y^g - [D^g, E^g] \begin{bmatrix} x \\ \beta \end{bmatrix} \right\|_2 \leq \epsilon$$

(6) 计算 y^g 的类关联重构误差:

$$r_i(y^g) = \|y^g - D^g \delta_i(\hat{x}_1) - E^g \hat{\beta}_1\|_2, \quad i = 1, 2, \dots, C$$

输出: 测试样本图像 y 的类别标签:

$$\text{class}(y) = \arg \min_i r_i(y^g)$$

在 GLR_SRC 方法中, 虽然对 LR 分解后的图像(包括原始人脸图像和误差图像)进行了 8 个方向、5 个尺度上的 Gabor 小波分解, 但由于本文以采样 $\rho = 8 \times 8$ 为间隔对每个尺度和方向上的 Gabor 特征进行了下采样, 从而使图像的 Gabor 特征向量维数低于其原始特征向量维数. 例如, 一幅分辨率为 $r \times c$ 的人脸图像, 在 8 个方向 5 个尺度上进行 Gabor 小波分解, 以采样因子 $\rho = 8 \times 8$ 为间隔进行下采样后的 Gabor 特征向量维数为 $40/64(r \times c)$, 仅为该人脸图像原始特征向量维数 $r \times c$ 的 $5/8$. 从这一方面来说, 本文采用人脸图像的 Gabor 特征向量进行运算, 其运算复杂度低于 SRC 采用图像原始特征向量进行运算的运算复杂度. 另一方面, 假设人脸图像原始特征向量维数为 m , 训练样本数为 n , 通常 $m \gg n$. SRC 所用误差字典(单位阵)的维数为 $m \times m$, 而 GLR_SRC 所用误差字典的维数为 $5/8m \times n$. 这也使 GLR_SRC 的运算复杂度低于 SRC 方法.

4 实验

为了评估 GLR_SRC 方法的性能, 本文在 CMU PIE、Extended Yale B 和 AR 三个人脸图像数据库上, 分别用 Gabor + 最近邻分类器(NN)、Gabor + LRC^[16]、Gabor + SRC^[8]、Gabor + ESRC^[10] 和 GLR_SRC 方法进行人脸图像识别实验. 在 CMU PIE 和 Extended Yale B 数据库上, 本文随机从每人全部人脸图像中抽取 k 幅图像构成训练样本集, 剩余图像构成测试样本集. 为了减少随机抽样对实验结果的影响, 对任一训练样本数 k , 独立进行 5 次实验, 最后取 5 次实验的平均识别率为最终识别率; 在 AR 数据上, 本文设计了四种实验方案来评估人脸图像有遮挡误差时 GLR_SRC 方法的性能. 另外需要说明的是, 本文采用 PCA 方法将人脸图像的 Gabor 特征向量降到不同维度, 以评估 GLR_SRC 方法在不同特征维数

下的性能. 对 l_1 范数最优化问题, 本文采用 l_1 - l_s 方法^[17]求解.

4.1 CMU PIE 数据库

CMU PIE 数据库^[18]共有 68 个人的 41368 张图像, 其中每个人的图像共有 13 种姿态、43 种光照和 4 种表情. 本文选择 CMU PIE C07 子集作为实验用数据库. CMP PIE C07 子集中共有 68 人的 1629 张人脸图像, 每人大约 24 张人脸图像. 本文对 CMP PIE C07 子集中的人脸图像, 以眼睛为中心进行了剪切和对齐处理, 并将其压缩为 32×32 像素.

为了全面评估 GLR_SRC 的性能, 在 CMU PIE 数据库上, 本文针对人脸图像未添加噪声和添加了 5% 椒盐噪声两种情况独立进行实验. 实验时, 分别从每人的全部人脸图像中随机选取 6 张和 10 张图像构成训练样本集, 剩余图像构成测试样本集. 图 2 给出了 CMU PIE 数据库上人脸图像未添加噪声的情况下, GLR_SRC 与其它 4 种方法随特征维数变化的识别率曲线; 图 3 给出了 CMU PIE 数据库上人脸图像添加了 5% 椒盐噪声的情况下, GLR_SRC 与其它 4 种方法随特征维数变化的识别率曲线. 表 1 给出了 CMU PIE 数据库上, GLR_SRC 与其它 4 种方法的最高人脸图像识别率对比.

表 1 CMU PIE 数据库上, GLR_SRC 与其它 4 种方法的最高人脸图像识别率 (%) 对比

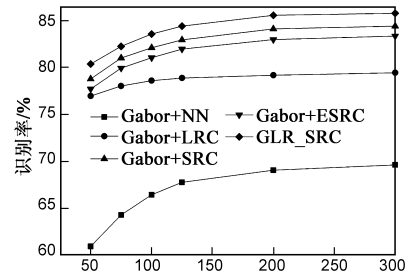
算法	图像未添加噪声		图像添加了 5% 椒盐噪声	
	6	10	6	10
Gabor + NN	69.62	83.52	38.72	48.98
Gabor + LRC	79.44	89.48	53.15	66.51
Gabor + SRC	84.42	91.93	65.29	74.50
Gabor + ESRC	83.36	91.57	64.34	73.76
GLR_SRC	85.78	92.90	68.35	77.41

由图 2 和表 1 可以看出, 在 CMU PIE 数据库中人脸图像未添加噪声情况下, GLR_SRC 方法的识别率明显优于 Gabor + SRC、Gabor + ESRC 方法. 这说明 GLR_SRC 所用的误差图像能更准确地描述人脸图像中的光照、表情和姿态误差; 由图 3 和表 1 可以看出, 在 CMU PIE 数据库中图像添加了 5% 椒盐噪声情况下, GLR_SRC 方法同样明显优于其它 4 种方法. 这进一步说明 GLR_SRC 所用的误差图像不但能准确地描述人脸图像中的光照、表情、姿态等误差, 而且对噪声的描述也比较准确. 另外还可以看出, 在两种情况下, GLR_SRC、Gabor + SRC、Gabor + ESRC 的识别率都明显优于 Gabor + NN 和 Gabor + LRC 方法. 这说明稀疏表示相关分类器的人脸图像识别性能优于最近邻分类器和线性回归分类器.

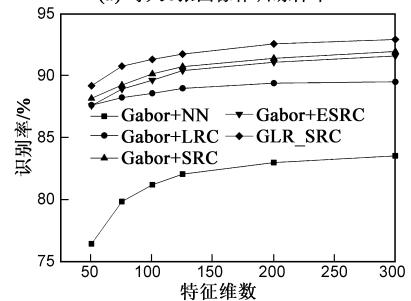
4.2 Extended Yale B 数据库

Extended Yale B 数据库^[19]包含 38 个人的 2414 张人

脸图像. 每张人脸图像为 256 个灰度级, 分辨率为 243×320 像素. 该数据库中人脸图像光照、表情以及拍摄角度都有较大程度的变化. 本文对 Extended Yale B 数据库中每张人脸图像都以眼睛为中心进行了剪切和对齐处理, 并将其压缩为 32×32 像素.

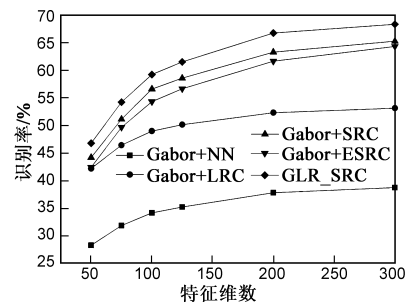


(a) 每人 6 张图像作训练样本

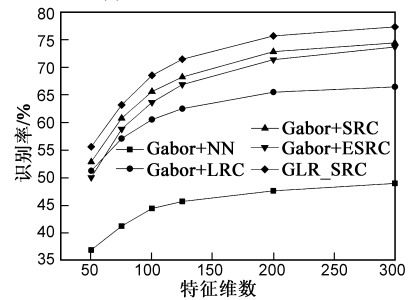


(b) 每人 10 张图像作训练样本

图 2 CMU PIE 数据库上, 图像未添加噪声情况下, 5 种方法随特征维数变化的识别率曲线



(a) 每人 6 张图像作训练样本



(b) 每人 10 张图像作训练样本

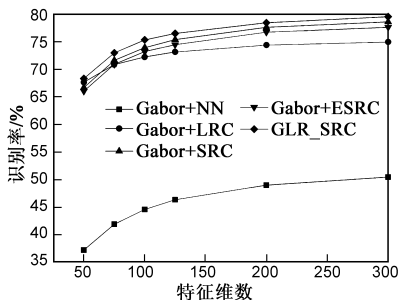
图 3 CMU PIE 数据库上, 图像添加了 5% 椒盐噪声情况下, 5 种方法随特征维数变化的识别率曲线

同样地, 在 Extended Yale B 数据库上, 本文也针对人脸图像未添加噪声和添加了 5% 椒盐噪声两种情况独立

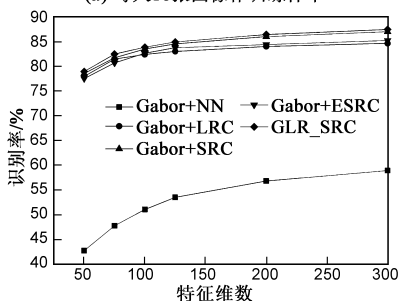
进行实验.实验时,分别从每人的全部人脸图像中随机抽取 16 张和 32 张图像构成训练样本集,其余图像构成测试样本集.图 4 和图 5 分别给出了 Extended Yale B 数据库上,图像未添加噪声和添加了 5% 椒盐噪声两种情况下,GLR_SRC 与其它 4 种方法随特征维数变化的识别率曲线.表 2 给出了 Extended Yale B 数据库上,GLR_SRC 与其它 4 种方法的最高人脸图像识别率对比.

表 2 Extend Yale B 数据库上,GLR_SRC 与其它 4 种方法的最高人脸图像识别率 (%) 对比

算法	图像未添加噪声		图像添加了 5% 椒盐噪声	
	16	32	16	32
Gabor + NN	50.45	58.91	39.29	46.93
Gabor + LRC	75.01	84.59	59.99	76.21
Gabor + SRC	78.69	86.94	68.28	77.81
Gabor + ESRC	77.71	85.18	65.98	75.69
GLR_SRC	79.63	87.41	70.19	79.76

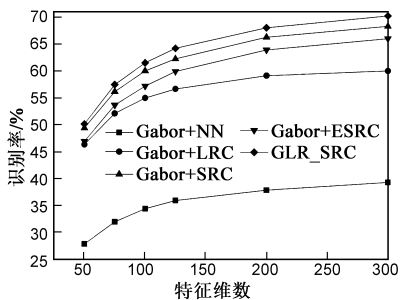


(a) 每人 16 张图像作训练样本

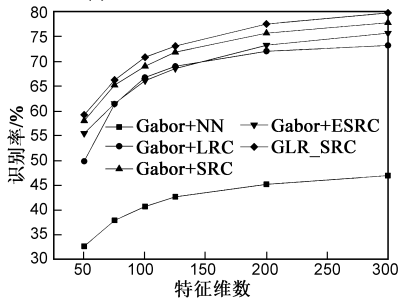


(b) 每人 32 张图像作训练样本

图 4 Extended Yale B 数据库上,图像未添加噪声情况下,5 种方法随特征维数变化的识别率曲线



(a) 每人 16 张图像作训练样本



(b) 每人 32 张图像作训练样本

图 5 Extended Yale B 数据库上,图像添加了 5% 椒盐噪声情况下,5 种方法随特征维数变化的识别率曲线

由图 4 和表 2 可以看出,在 Extended Yale B 数据库中图像未添加噪声情况下,GLR_SRC 方法的识别率优于 Gabor + SRC 和 Gabor + ESRC 方法.这说明了 GLR_SRC 所用的误差图像能准确地描述人脸图像中的光照、表情以及拍摄角度等误差;由图 5 和表 2 可以看出,在 Extended Yale B 数据库中图像添加了 5% 椒盐噪声情况下,GLR_SRC 方法明显优于其它 4 种方法.这充分说明 GLR_SRC 采用 LR 算法求得的误差图像来描述人脸图像中的误差和噪声,比其他 4 种方法抗干扰能力更强.

4.3 AR 数据库

AR 数据库^[20]有 126 个人的超过 4000 张正面人脸图像.其中,每人的 26 张人脸图像分别采集于两个不同的阶段,其中第 1~13 张图像采集于第一阶段,第 14~26 张图像采集于第二阶段.该数据库中人脸图像光照、脸部表情变化都比较大,并且部分人脸图像有眼镜和围巾遮挡.其中,被眼镜遮挡的人脸图像的遮挡面积约占整个脸部的 20%;被围巾遮挡的人脸图像的遮挡面积约占整个脸部的 40%.在 AR 数据库上,本文选择 65 个男人和 55 个女人,每人 26 张人脸图像,共 3120 张人脸图像做实验.另外需要说明的是,实验时本文将 AR 数据库中图像都转换为灰度图像,并以眼睛为中心进行了剪切和对齐处理,且将其大小压缩为 50×40 像素.

为了评估人脸图像有遮挡误差时 GLR_SRC 方法的性能,本文设计了如下四种实验方案:

方案 1 训练样本集中不含遮挡图像、测试样本集中同时包含眼镜和围巾遮挡图像实验方案

本方案从每人第一阶段图像中选取 7 张无遮挡图像构成训练样本集,用第二阶段图像中的 7 张无遮挡图像以及 1 张眼镜遮挡和 1 张围巾遮挡图像构成测试样本集.图 6(a)给出了该实验方案下,GLR_SRC 与其它 4 种方法随特征维数变化的识别率曲线.

方案 2 训练样本集和测试样本集中都含眼镜遮挡图像实验方案

本方案从每人第一阶段图像中选取 7 张无遮挡图像以及 1 张眼镜遮挡图像构成训练样本集,用第二阶段图像中的 7 张无遮挡图像,以及两个阶段图像中剩余的 5 张眼镜遮挡图像构成测试样本集,即每人 8 张图像作

训练样本,12张图像作测试样本.图6(b)给出了该实验方案下,GLR_SRC与其它4种方法随特征维数变化的识别率曲线.

方案3 训练样本集和测试样本集中都含围巾遮挡图像实验方案

本方案选取每人8张图像(第一阶段图像中7张无遮挡图像和1张围巾遮挡图像)构成训练样本集和12张图像(第二阶段图像中7张无遮挡图像和两个阶段中剩余的5张围巾遮挡图像)构成测试样本集.图6(c)给出了该实验方案下,GLR_SRC与其它4种方法的随特征维数变化的识别率曲线.

方案4 训练样本集和测试样本集中都同时包含眼镜和围巾遮挡图像实验方案

本方案选取每人第一阶段图像中7张无遮挡图像,再加上1张眼镜遮挡图像和1张围巾遮挡图像,共9张人脸图像构成训练样本集,其余17张图像构成测试样

本集.图6(d)给出了该实验方案下,GLR_SRC与其它4种方法随特征维数变化的识别率曲线.

表3给出AR数据库上,四种实验方案下GLR_SRC与其它4种方法的最高人脸识别率对比.由图6和表3可以看出,在人脸图像含遮挡误差时,GLR_SRC方法的识别率优于其它4种方法.这说明GLR_SRC方法所用的误差图像能准确地描述人脸图像中的遮挡误差,同时也说明GLR_SRC方法在处理人脸图像“遮挡”问题时,比其它4种方法更有效.

表3 AR数据库上,GLR_SRC与其它4种方法的最高人脸识别率(%)对比

算法	方案1	方案2	方案3	方案4
Gabor + NN	65.46	76.11	70.56	73.63
Gabor + LRC	67.59	78.19	73.81	76.67
Gabor + SRC	75.26	83.33	81.04	82.4
Gabor + ESRC	75.19	83.06	80.49	82.21
GLR_SRC	77.13	83.96	82.08	83.28

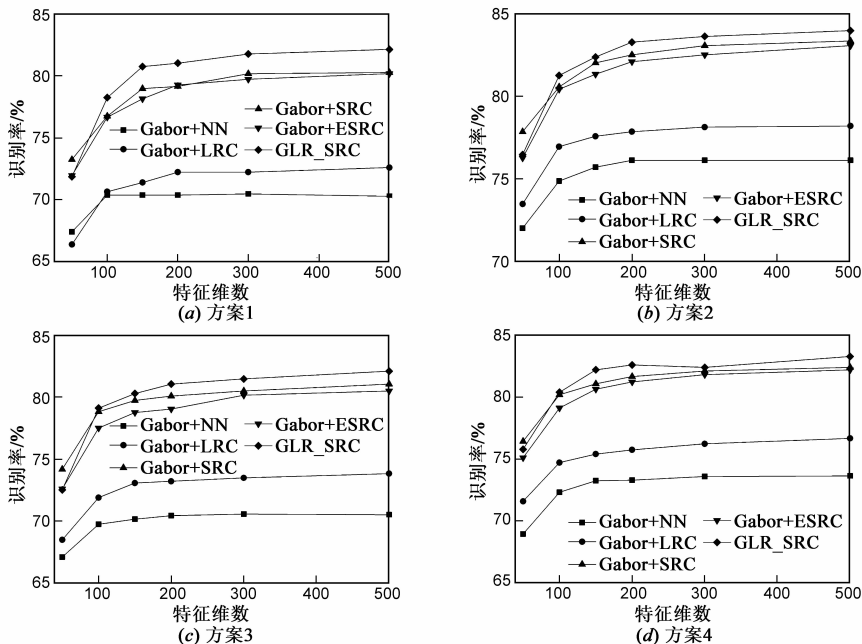


图6 AR数据库上5种方法随特征维数变化的识别率曲线

5 结束语

本文研究发现,低秩矩阵恢复(LR)算法求得的误差图像可以准确地描述人脸图像中的光照、表情、姿态、遮挡等误差和噪声.基于此,提出了一种基于Gabor低秩恢复稀疏表示分类(GLR_SRC)的人脸图像识别方法.GLR_SRC用全部人脸训练样本图像的Gabor特征向量和LR算法求得的误差图像的Gabor特征向量共同组成字典,并在该字典下计算人脸测试样本图像Gabor特征向量的稀疏表示,以实现测试样本图像的识别.在CMU PIE、Extended Yale B和AR数据库上的实验结果

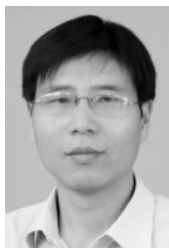
表明,基于GLR_SRC的人脸图像识别方法具有较高的识别率和较强的抗干扰能力.

参考文献

- [1] VINJE W E, GALLANT J L. Sparse coding and decorrelation in primary visual cortex during natural vision [J]. Science, 2000, 287(5456): 1273 - 1276.
 - [2] 梁栋, 梁昭, 鲍文霞, 等. 基于非局部正则化稀疏表示的图像去噪算法[J]. 系统工程与电子技术, 2013, 35(5): 1104 - 1109.
- LIANG Dong, LIANG Zhao, BAO Wen-xia, et al. Image denoising algorithm based on non-local regularized sparse representation

- sensation[J]. *Systems Engineering and Electronics*, 2013, 35(5):1104–1109. (in Chinese)
- [3] 张晓阳, 柴毅, 李华峰. 基于 K-SVD 和残差比的低信噪比图像稀疏表示去噪算法[J]. *光学技术*, 2012, 38(1):23–29. ZHANG Xiao-yang, CHAI Yi, LI Hua-feng. Low SNR image denoising via sparse and redundant representations over K-SVD algorithm and residual ratio iteration termination[J]. *Optical Technique*, 2012, 38(1):23–29. (in Chinese)
- [4] 孔繁镛, 井庆丰, 计振兴. 图像压缩感知的自适应方向提升稀疏表示及重构算法[J]. *宇航学报*, 2013, 34(1):121–127. KONG Fan-qiang, JING Qing-feng, JI Zhen-xing. Imagesparse representation and reconstruction algorithm for compressed sensing by adaptive direction lifting[J]. *Journal of Astronautics*, 2013, 34(1):121–127. (in Chinese)
- [5] 练秋生, 张伟. 基于图像块分类稀疏表示的超分辨率重构算法[J]. *电子学报*, 2012, 40(5):920–925. LIAN Qiu-sheng, ZHANG Wei. Image super-resolution algorithms based on sparse representation of classified image patches[J]. *Acta Electronica Sinica*, 2012, 40(5):920–925. (in Chinese)
- [6] 宋琳, 程咏梅, 赵永强. 基于稀疏表示模型和自回归模型的高光谱分类[J]. *光学学报*, 2012, 32(3):1–7. SONG Lin, CHENG Yong-mei, ZHAO Yong-qiang. Hyperspectrum classification based on sparse representation model and auto-regressive model[J]. *Acta Optica Sinica*, 2012, 32(3):1–7. (in Chinese)
- [7] 胡正平, 宋淑芬. 基于类别相关邻域空间的极大似然稀疏表示鲁邦图像识别算法[J]. *自动化学报*, 2012, 38(9):1420–1427. HU Zheng-ping, SONG Shu-fen. Robust image recognition algorithm of maximum likelihood estimation sparse representation based on class-related neighbors subspace[J]. *Acta Automatica Sinica*, 2012, 38(9):1420–1427. (in Chinese)
- [8] WRIGHT J, YANG A Y, A. Ganesh, et al. Robust face recognition via sparse representation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(2):210–227.
- [9] YANG M, ZHANG L. Gabor feature based sparse representation for face recognition with Gabor occlusion dictionary[A]. *Proceedings of 11th European Conference on Computer Vision [C]*. Crete, Greece: Springer, 2010. 448–461.
- [10] DENG W H, HU J, GUO J. Extended SRC: Undersampled face recognition via intra-class variant dictionary[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(9):1864–1870.
- [11] LIN Z, LIU R S, SU Z X. Linearized alternating direction method with adaptive penalty for low-rank representation [A]. *Proceedings of 25th Annual Conference on Neural Information Processing Systems [C]*. Granada, Spain: Springer, 2011. 1–9.
- [12] 侯毅, 周石琳, 雷琳, 等. 基于 Gabor 滤波器组的多尺度不变特征提取方法[J]. *电子学报*, 2013, 41(6):1146–1152. HOU Yi, ZHOU Shi-lin, LEI Lin, et al. Invariant feature with multi-characteristic scales using Gabor filter bank[J]. *Acta Electronica Sinica*, 2013, 41(6):1146–1152. (in Chinese)
- [13] YANG A Y, ZHOU Z H, BALASUBRAMANIAN A G, et al. Fast ℓ_1 -minimization algorithms for robust face recognition[J]. *IEEE Transactions on Image Processing*, 2013, 22(8):3234–3246.
- [14] CAI J, CANDES E, SHEN Z. A singular value thresholding algorithm for matrix completion[J]. *SIAM Journal of Optimization*, 2010, 20(4):1956–1982.
- [15] LEE T S. Image representation using 2D Gabor wavelet[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1996, 18(10):959–971.
- [16] NASEEM I, TOGNERI R, BENNAMOUN M. Linear regression for face recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(11):2106–2112.
- [17] KIM S J, KOH K, LUSTIG M, et al. A method for large-scale ℓ_1 -regularized least squares[J]. *IEEE Journal on Selected Topics in Signal Processing*, 2007, 1(4):606–617.
- [18] SIM T, BAKER S, BSAT M. The CMU pose, illumination, and expression database[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(12):1615–1618.
- [19] GEORGHIADES S, BELHUMEUR P N, KRIEGMAN D J, et al. From few to many: illumination cone models for face recognition under variable lighting and pose[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(6):643–660.
- [20] MARTINEZ A M, BENAVENTE R. *The AR Face Database [R]*. Barcelona: Centre de Visio per Computador Univeresit, 1998.

作者简介



杜海顺 男, 1977年2月生于河南延津. 2007年毕业于东南大学, 获工学博士学位. 现为河南大学计算机与信息工程学院副教授, 硕士生导师, 主要从事模式识别、图像处理等领域的研究工作.

E-mail: jddhs@henu.edu.cn

张旭东 男, 1985年5月生于河南上蔡. 现为河南大学计算机与信息工程学院控制理论与控制工程专业硕士研究生, 主要研究方向为生物特征识别、稀疏表示等.

E-mail: xudong2366@163.com