

基于低秩分解的联合动态稀疏表示 多观测样本分类算法

胡正平,高红霄,赵淑欢

(燕山大学信息科学与工程学院,河北秦皇岛 066004)

摘 要: 通过互联网易获得同一对象的多个无约束的观测样本,针对如何解决无约束观测样本带来的识别困难及充分利用多观测样本数据信息提高其分类性能问题,提出基于低秩分解的联合动态稀疏表示多观测样本分类算法.该算法首先寻找到一组最佳的图像变换域,使得变换图像可以分解成一个低秩矩阵和一个相关的稀疏误差矩阵;然后对低秩矩阵和稀疏误差矩阵分别进行联合动态稀疏表示,以便充分利用类级的相关性和原子级的差异性,即使多观测样本的稀疏表示向量在类别上分享相同的稀疏模型,而在原子级上采用不同的稀疏模型;最后利用总的稀疏重建误差进行类别判决.在 CMU-PIE 人脸数据库、ETH-80 物体识别数据库、USPS 手写体数字数据库和 UMIST 人脸数据库上进行对比实验,实验结果表明本方法的优越性.

关键词: 模式识别;多观测样本分类;低秩矩阵恢复;联合动态稀疏表示

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 0372-2112 (2015)03-0440-07

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

DOI: 10.3969/j.issn.0372-2112.2015.03.004

Multiple Observation Sets Classification Algorithm Based on Joint Dynamic Sparse Representation of Low-Rank Decomposition

HU Zheng-ping, GAO Hong-xiao, ZHAO Shu-huan

(School of Information Science and Engineering, Yanshan University, Qinhuangdao, Hebei 066004, China)

Abstract: Multiple unconstrained observations of the same object can be easily accessed by the Internet, with regard to overcoming the identification-difficult of the unconstrained samples. Moreover, to exploit the information of multiple observation sets to improve the classification performance, a multiple observation sets classification algorithm based on joint dynamic sparse representation of low-rank decomposition is presented. First of all, we need find the best set of image transform domain, which decomposes the data matrix into a low-rank matrix and an associated sparse error matrix. Secondly, the low-rank matrix and sparse error matrix is represented by joint dynamic sparsity respectively, in order to make full use of the correlation of the class-level and the differences of the atom-level, i.e., the sparse representation vectors for the multiple observations can share the same class-level sparsity pattern while their atom-level sparsity patterns may be distinct. Finally, we compare the classification results with the total sparse reconstruction errors. Three comparative experiments are conducted on CMU-PIE face dataset, ETH-80 object recognition dataset, USPS handwritten digit dataset, and UMIST face dataset, and the results demonstrate the superiority of the proposed algorithm.

Key words: pattern recognition; multiple observation sets classification; low-rank matrix recovery; joint dynamic sparse representation

1 引言

近年来,随着信息技术的发展,图像和视频分享网站日益普及,从而导致在线可用的视觉数据急剧增加,在视觉领域,这大大激发人们对无约束多观测样本分类的兴趣.在实际分类问题中,存在如光照变化,角度变

化,部分遮挡,污损甚至没有校准对齐等问题,这些数据对识别分类具有更高要求.这些要求对现存的视觉分类算法提供机会和挑战,因为域变换难以精确的测量待分类图像的相似度.如何克服这些困难,充分利用无约束多观测样本所提供的大量信息提高实际应用中的辨别精度和鲁棒性,正是国内外学者广泛关注的问题.

针对多观测样本分类问题,国内外学者展开了不少工作,尤其是稀疏表示算法已然成为模式识别和信号处理的前沿热点^[1~3].文献[4]在单观测样本稀疏表示分类^[5~7]的基础上提出多观测样本独立稀疏表示分类(SRC)算法,该方法是将每个观测样本分别通过求解 l_1 范数最小优化问题进行稀疏表示.为了进一步提高该方法的性能,文献[8]在不同类的字典间引入一个非相关项,促使来自不同类的字典尽可能的独立.但是上述改进的字典没有考虑数据的局部特征,因此文献[9]提出一种局部敏感字典学习方法,在字典中保留数据局部结构,然后利用稀疏表示进行分类.由于文献[4]中每个观测样本优化均是可分的,故不能充分利用不同观测样本之间的相关性,因此文献[10]提出了一种稀疏可能构图,利用非负稀疏权重作为聚类指标实现分类目标;而文献[11]则通过将组(类)稀疏和数据局部性结构整合成统一的优化框架,通过对 $l_{1,2}$ 混合范数优化求解进行稀疏表示及识别;不同于上述两种方法,文献[12]提出多任务联合稀疏表示分类(MTJSRC)算法,该方法是将多观测样本当成一个整体,让所有视角分享相同的稀疏模型,利用总的稀疏重建误差进行类别判决,由于多观测样本是通过不同的角度获得的,因此让所有视角分享相同的稀疏模型不能充分利用不同视角的差异性,为了进一步解决这一问题,文献[13,14]提出联合动态稀疏表示分类(JDSRC)算法,该思路利用类级的相关性在该级上分享相同的稀疏模型,同时利用原子级的差异性在该级上采用不同的稀疏模型,再利用总的稀疏重建误差进行类别判决.

联合动态稀疏表示是直接训练图像和测试图像上进行的,由于无约束数据本身所存在的问题,直接对其进行联合动态稀疏表示可能会对分类造成一定的影响.如何对数据进行预处理,使其对噪声和奇异值有更强的鲁棒性?本文提出基于低秩分解^[15,16]的联合动态稀疏表示分类(LR-JDSRC)算法.该方法的主体思路是将训练图像和测试图像分解成一个低秩矩阵和相对应稀疏误差矩阵;然后充分利用类级的相关性和原子级的差异性对低秩矩阵和稀疏误差矩阵分别进行联合动态稀疏表示,最后利用加权后的总的稀疏重建误差进行类别判决.

2 基于低秩分解的联合动态稀疏表示多观测样本分类

2.1 低秩矩阵恢复

在实际分类问题中,不能期望训练样本总是在可控制的环境下进行采集,不可避免的会存在光照变化,视角变化,部分遮挡,污损等.使用这样的图像作为训练样本,识别算法可能会拟合这些不可控制因素所引

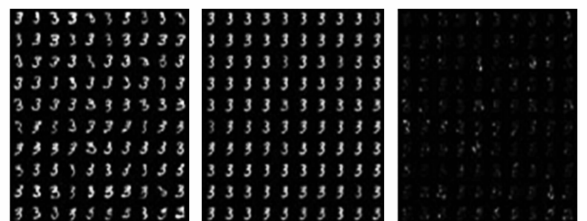
起的噪声而不是对待识别的特征进行建模.而低秩矩阵恢复则可以通过将采集来的数据矩阵分解成一个低秩矩阵和一个相对应的稀疏误差矩阵来缓解上述问题,其中所提取的低秩矩阵要保存数据结构,因此相对应的误差矩阵是稀疏的.将低秩矩阵恢复应用到 N 个对象的物体识别,采集训练样本 $D = [D_1, D_2, \dots, D_N]$ 如图1(a)所示,其中 $D_i \in \mathbf{R}^{m \times n_i}$ 是第 i 个对象的训练数据矩阵, m 是图像的维数,而 n_i 是该对象的样本数.低秩矩阵恢复的目的是将数据矩阵 $D = [D_1, D_2, \dots, D_N]$ 分解成一个低秩矩阵 $A = [A_1, A_2, \dots, A_N]$ (如图1(b)所示)和一个对应的稀疏误差矩阵 $E = [E_1, E_2, \dots, E_N]$ (如图1(c)所示);更精确的描述是解决下列优化问题:

$$\min_{A, E} \text{rank}(A) + \lambda \|E\|_0 \quad \text{s.t. } D = A + E \quad (1)$$

其中, $\|\cdot\|_0$ 指的是 l_0 范数,即矩阵中非零元素的个数, $\lambda > 0$ 是一个正规化参数,由于上述优化问题难以解决,故进行下列优化使最初的低秩矩阵恢复易处理:

$$\min_{A, E} \|A\|_* + \lambda \|E\|_1 \quad \text{s.t. } D = A + E \quad (2)$$

其中,核范数 $\|A\|_*$ (即奇异值的总和)接近 A 的秩;用 l_1 范数 $\|E\|_1$ 代替 l_0 范数 $\|E\|_0$,它是 E 中所有元素绝对值的总和; $\lambda = C/\sqrt{\max(m, n_i)}$, C 是一个常数,通常为1.只要被恢复的矩阵 A 的秩不是太大且 E 中错误数量是稀疏的,那么解决这个凸相关问题相当于解决原始低秩矩阵逼近问题.考虑到计算效率,在此应用不精确的增广拉格朗日乘子(ALM)解决式(2)的优化问题.只要被恢复的矩阵 A 的秩不是太大且 E 中误差数量是稀疏的,那么解决这个凸相关问题相当于解决原始低秩矩阵逼近问题.



(a) 原始图像 (b) 低秩图像 (c) 稀疏误差图像

图1 低秩矩阵恢复结果实例

2.2 联合动态稀疏表示多观测样本分类

2.2.1 稀疏表示分类思想

假设第 i 类的测试样本为 $y \in \mathbf{R}^{d \times 1}$,训练样本矩阵为 $A_i = [a_{i1}, a_{i2}, \dots, a_{in_i}] \in \mathbf{R}^{d \times n_i}$,其中 d 为其特征维数, n_i 为其样本数目.由线性子空间理论可知, y 可以由 $a_{i1}, a_{i2}, \dots, a_{in_i}$ 线性表示,即

$$y = x_{i1}a_{i1} + x_{i2}a_{i2} + \dots + x_{in_i}a_{in_i} = A_i X_i \quad (3)$$

其中, $X_i = [x_{i1}, x_{i2}, \dots, x_{in_i}]^T$ 是包含第 i 类原子权重的系

数向量. 因为在实际分类中, \mathbf{y} 的类别并不是已知的, 所以定义了涵盖所有类别的训练样本矩阵(即过完备字典) $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_{C_i}] \in \mathbf{R}^{d \times N}$, 其中, C_i 是类别数目,

$N = \sum_{i=1}^{C_i} n_i$ 为样本总数. 故 \mathbf{y} 可以表示为下列线性组合:

$$\mathbf{y} = \mathbf{A}\mathbf{X} \in \mathbf{R}^{d \times 1} \quad (4)$$

上式中, $\mathbf{X} = [\mathbf{0}^T, \mathbf{0}^T, \dots, \mathbf{X}_i^T, \dots, \mathbf{0}^T]^T \in \mathbf{R}^N$ 是系数向量, \mathbf{X} 中除与 \mathbf{y} 同类别样本的系数不为零外, 其余均为零. 当类别数目 C_i 足够大时, 式(4)的解是稀疏的, 当此解越稀疏时, 对测试样本待定类别的求取越有利. 因此将式(4)的求解问题转化为基于最小化 l_0 范数的最稀疏求解问题, 如下式所示:

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \|\mathbf{X}\|_0 \quad \text{s.t. } \mathbf{y} = \mathbf{A}\mathbf{X} \quad (5)$$

由于实际图像中包含噪声, 所以测试样本几乎不可能精确地由训练样本线性表示, 故上式改写成下列形式:

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \|\mathbf{X}\|_0 \quad \text{s.t. } \|\mathbf{y} - \mathbf{A}\mathbf{X}\|_2 \leq \epsilon \quad (6)$$

或者

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \|\mathbf{y} - \mathbf{A}\mathbf{X}\|_2^2 \quad \text{s.t. } \|\mathbf{X}\|_0 \leq K \quad (7)$$

其中 ϵ 为噪声容限.

根据各类训练样本对测试样本的稀疏逼近残差进行分类, 将待定类别定为最小稀疏逼近残差对应的类别, 如下式所示

$$\hat{i} = \arg \min_i \|\mathbf{y} - \mathbf{A}\delta_i(\hat{\mathbf{X}})\|_2^2 \quad (8)$$

其中, $\delta_i(\cdot)$ 是一个矢量算子, 保持第 i 类对应元素不变, 其他类元素均为零.

2.2.2 联合稀疏表示模型

在实际情况中, 容易得到同一物体的多个视角, 因此很自然地考虑利用同一物体的多个不同视角的信息制定一个相对独立的联合决策. 给出同一物体的 M 个视角, 则将 M 个稀疏表示问题式(7)结合在一起为:

$$\begin{cases} \hat{\mathbf{X}}_i \}_{i=1}^M = \arg \min_{\{\mathbf{X}_i\}} \sum_{i=1}^M \|\mathbf{y}_i - \mathbf{A}\mathbf{X}_i\|_2^2 \\ \text{s.t. } \|\mathbf{X}_i\|_0 \leq K, \forall 1 \leq i \leq M \end{cases} \quad (9)$$

然而, 因为每个视角优化是可分的, 所以这种规划不能充分利用不同视角之间的相关性, 如图 2(a)所示. 为了利用多视角之间相关信息进行识别, 在此引入联合稀疏假设模型, 多个稀疏表示向量共享相同的稀疏模型, 即选择相同的一组原子来表示每一视角, 然而相同的选择原子对应的系数值可能不同, 如图 2(b)所示. 根据这一假设, 多视角稀疏表示转化为下面的优化问题:

$$\begin{cases} \hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2 \\ \text{s.t. } \|\mathbf{X}\|_{l_0 \setminus l_2} \leq K \end{cases} \quad (10)$$

其中, $\|\cdot\|_F$ 表示酉不变范数, $\|\mathbf{X}\|_{l_0 \setminus l_2}$ 是混合范数, 在每一行进行 l_2 范数, 在结果向量中进行 l_0 范数. 这是联合稀疏表示的基础分类方法, 通过式(10)得到的稀疏表示矩阵具有行稀疏的特征. 该方法将多观测样本当成一个整体(多个观测向量合在一起), 利用总的稀疏重建误差进行类别判决, 由于多观测样本是通过不同的角度获得的, 因此让所有视角分享相同的稀疏模型也具有一定的局限性.

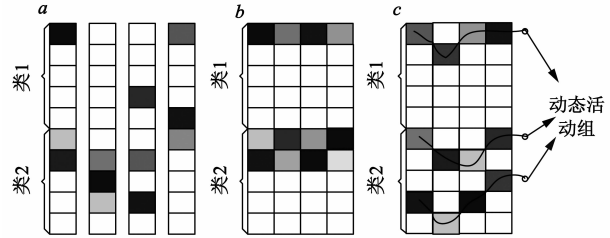


图2 不同稀疏模型的系数矩阵 \mathbf{X} 的说明^[14]

2.2.3 联合动态稀疏表示模型

为了达到充分利用类级的相关性和原子级的差异性的目的, 多观测样本理想的稀疏表示向量应该是在类级上分享相同的稀疏模型, 并在原子级上采用不同的稀疏模型, 称作联合动态稀疏表模型, 如图 2(c)所示. 联合动态稀疏模型进行联合动态稀疏的关键之一是引入了动态活动组, 每个动态活动组 $g_s \in \mathbf{R}^M$, $s = 1, 2, \dots$, 指的是系数矩阵 \mathbf{X} 中同一类的一组系数的索引(行), 在多观测样本稀疏表示时, 大量的动态活动组共同地活动. 每个动态活动组 g_s 只包含 \mathbf{X} 每一列中的一个索引, 即 $g_s(m)$ 表示的是第 s 个动态活动组中系数矩阵 \mathbf{X} 第 m 列中选择原子的行坐标. 因此在该方法中, 每个视角的稀疏表示是不同的, 但他们分享相同的类级结构.

该方法的特性是:(1) 在联合稀疏表示过程中, 由于来自同一物体的多个视角的信息被合并, 因此加强了联合稀疏恢复的鲁棒性;(2) 通过动态活动组促进稀疏, 从而在恢复多稀疏表示向量时产生联合动态稀疏模型. 在一个动态活动组中为了加强并结合所有原子的信息, 故在每一个动态活动组上应用 l_2 范数. 在联合稀疏表示中, 为了促进稀疏允许小数量的动态活动组参与, 并使动态活动组通过 l_2 范数后再进行 l_0 范数, 因此得到下面的联合动态稀疏优化条件:

$$\|\mathbf{X}\|_G = \|\left[\|\mathbf{X}_{g_1}\|_2, \|\mathbf{X}_{g_2}\|_2, \dots\right]\|_0 \quad (11)$$

其中, \mathbf{X}_{g_s} 表示的是由与第 s 个动态活动组相关的系数集合组成的向量:

$$\begin{aligned} \mathbf{X}_{g_s} &= \mathbf{X}(g_s) \\ &= [\mathbf{X}(g_s(1), 1), \mathbf{X}(g_s(2), 2), \\ &\quad \dots, \mathbf{X}(g_s(M), M)]^T \in \mathbf{R}^M \end{aligned} \quad (12)$$

为了得到多观测样本 $\{y_m\}_{m=1}^M$ 在联合动态稀疏约束下的稀疏表示系数矩阵 X , 提出了下列联合动态稀疏表示模型:

$$\begin{aligned} \hat{X} = \arg \min_x & \|Y - AX\|_F^2 \\ \text{s.t. } & \|X\|_C \leq K \end{aligned} \quad (13)$$

其中 K 为稀疏度, $\|X\|_C$ 是 l_2 与 l_0 的组合, 其主要优点为: l_2 作用于每一个动态活动集, 便于在类内多观测样本中灵活进行原子选择, 进而在类内多观测样本中选择出最有代表性的原子; l_0 作用于整个动态活动集, 以便于选择最有代表性的动态活动集, 即在类间选择出最有代表性的类模式。

由于 l_0 范数和联合动态稀疏约束共存, 对式(13)的解决带来挑战, 为了解决式(13), 在此引入贪婪动态稀疏表示算法, 该算法结构与 CoSOMP^[17] 相类似, 包括下列步骤: (1) 在当前残差下选择新的候选者; (2) 将新的候选集与之前选择的原子集合并; (3) 通过合并的原子集估算表示系数; (4) 在新估计的表示系数的基础上将合并的原子集修剪到特定的稀疏水平上; (5) 更新残差, 直到满足文献[17]的收敛条件. 联合动态稀疏表示 (JDSR) 具体步骤如算法 1.

算法 1 联合动态稀疏表示

输入: 来自 C_l 个类别的训练样本 A , 训练样本类别标记 u , 多观测样本 Y , 稀疏度 K , 多观测样本数目 M

输出: 稀疏系数矩阵 X

1. 给定初始值: $R \leftarrow Y, I \leftarrow \emptyset$
2. 重复步骤 2 直到满足收敛条件
 - 2.1 计算 $E = A^T R$
 - 2.2 联合动态稀疏映射 (JDSM) 选择原子: $I_{\text{new}} \leftarrow P_{\text{JDSM}}(E, 2K)$
 - 2.3 索引矩阵更新: $I \leftarrow [I^T, I_{\text{new}}^T]^T$
 - 2.4 更新表示稀疏进行 for 循环: for $m = 1 : M$
 - (a) 计算 $i \leftarrow I(:, m)$
 - (b) 计算 $C(i, m) \leftarrow (A(:, i)^T A(:, i))^{-1} A(:, i)^T Y(:, m)$
- 2.5 通过联合动态稀疏映射 (JDSM) 进行裁剪: $I \leftarrow P_{\text{TDSM}(C, K)}$
- 2.6 $X \leftarrow \mathbf{0}$
- 2.7 进行 for 循环: for $m = 1 : M$
 - (a) 计算 $i \leftarrow I(:, m)$
 - (b) 计算 $X(i, m) \leftarrow C(i, m)$
- 2.8 更新残差: $R = AX - Y$

3. $X \leftarrow \mathbf{0}$

4. 进行 for 循环: for $m = 1 : M$

(a) 计算 $i \leftarrow I(:, m)$

(b) 计算

$$X(i, m) \leftarrow (A(:, i)^T A(:, i))^{-1} A(:, i)^T Y(:, m)$$

end

在联合稀疏表示算法每一次迭代过程中, (a) 和 (d) 均给出一个系数矩阵 $Z \in \mathbf{R}^{N \times M}$, 为了在 Z 中寻找 K 个表示动态活动组, 即利用 K 个动态活动组构建 Z 的最佳逼近 \hat{Z} (即 $\|\hat{Z}_K\|_C = K$), 需解决下列优化问题:

$$\begin{aligned} \hat{Z}_K = \arg \min_{Z \in \mathbf{R}^{N \times M}} & \|Z - Z_K\|_F \\ \text{s.t. } & \|\hat{Z}_K\|_C \leq K \end{aligned} \quad (14)$$

通过下列联合动态稀疏映射 (JDSM) 解决式 (14) 的问题:

$$I_K = P_{\text{JDSM}}(Z, K) \quad (15)$$

$I_K \in \mathbf{R}^{K \times M}$ 是 M 个观测样本的 K 个动态活动组的索引, 联合动态稀疏映射 P_{JDSM} 具体的步骤如下:

Step 1 输入: 系数矩阵 Z , 理想动态活动组数目 K , 字典原子标记 u , 类别数 C_l , 多观测样本数目 M ;

Step 2 索引矩阵 I_K 为空;

Step 3 对动态活动组数目进行 for 循环: for $i = 1 : K$

Step 4 对类别数进行 for 循环: for $j = 1 : C_l$

Step 5 查找标记 u 中第 j 类的索引向量 c ;

Step 6 对观测样本数目进行 for 循环: for $m = 1 : M$

Step 7 查找 $|Z(c, m)|$ 中的最大值 v 保存到

$V(c, m)$ 中且其索引为 t , 将 $c(t)$ 保存到 $\hat{I}(c, m)$;

Step 8 结束步骤 Step 6 的循环: end;

Step 9 合并每一类的最大系数:

$$s(c) = \sqrt{\sum_{m=1}^M V(c, m)^2}$$

Step 10 结束 Step 4 的循环: end

Step 11 贯穿所有类别, 查找同一组中最佳的原子组:

$$[\hat{v}, \hat{t}] = \max(s); I_K(i, :) = \hat{I}(\hat{t}, :),$$

$$Z(I_K(i, :)) = \mathbf{0}^T;$$

Step 12 end

Step 13 输出: 索引矩阵 I_K

2.3 基于低秩的联合动态稀疏表示多观测样本分类算法的描述

LR-JDSR 多观测样本分类算法具体步骤如算法 2.

算法 2 LR-JDSR 多观测样本分类算法

1. 输入:

(1) D : 来自 C_l 个类别的训练样本

(2) u : 训练样本类别标记

(3) Y : 多观测样本

(4) K : 稀疏度

(5) M : 多观测样本数目

2. 处理:

(1) 利用公式(2)对 D 和 Y 进行联合低秩分解, 得到低秩字典 A , 稀

疏误差字典 E , 低秩观测样本 Y^1 及稀疏误差样本 Y^2 ;

(2) 分别对 A 和 Y^1 及 E 和 Y^2 进行联合动态稀疏表示得到稀疏系数矩阵 X^1 和 X^2 ;

(3) 优化:

$$\hat{i} = \arg \min_i (\alpha \| Y^1 - A\delta^i(X^1) \|_F^2 + (1 - \alpha) \| Y^2 - E\delta^i(X^2) \|_F^2)$$

3. 输出:

\hat{i} : 多观测样本类别

3 实验结果与分析

3.1 CMU-PIE 人脸识别

本组实验数据来自于美国卡耐基梅隆大学所创建的 CMU-PIE 人脸数据库, 该数据库包含 68 个人在 4 个不同表情、13 个姿势、43 种不同光照下获得的面部照片共计 41368 张图像. 其中姿态和光照变化图像的采集都是在严格控制条件下进行的. 本实验采用接近正脸姿势的五组图像, 分别是 $c05$ 、 $c07$ 、 $c09$ 、 $c27$ 和 $c29$. 图 3 显示一多观测样本的变化图像.



图3 CMU-PIE数据库中外观变化例子

实验中, 所有图像都裁成中心脸的部分(如无背景、密切切割)调整为 32×32 大小. 在 68 个人中随机选择 55 个人进行实验, 从五组图像中随机选择三组, 每个人每组随机选出 6 幅图像共 990 幅图像作为训练样本, 同时五组图像均进行测试, 每个人每组随机选出 6 幅图像共 1650 幅图像作为测试样本. 经多次实验选出最优加权值 α 为 0.5, 稀疏度 K 通过搜索一系列可能值来确定. 为了进行公平的比较, LR-JDSRC 选取 K 的方法与 JDSRC 中选取 K 的方法一致. 针对不同稀疏度, 每次均进行十次实验求平均识别率选出最优的稀疏度, 此实验中 LR-JDSRC 算法 K 取值 13, JDSRC 算法 K 取值 15. 随机选择 10 次观测样本, 其平均识别率和方差如表 1 所示.

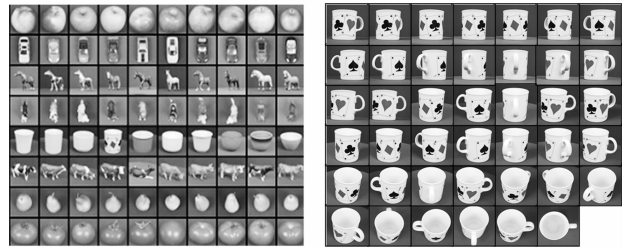
表 1 在 PIE 人脸数据库上的平均识别率

算法	ISM-Graph-GEDA	JDSRC	LR-JDSRC
识别率(%)	77.56	92.36	93.64
标准差	2.05	2.68	2.61

实验结果表明, 本文提出的 LR-JDSRC 算法相比较于 JDSRC 算法和 ISM-Graph-GEDA 算法有更高的识别率. LR-JDSRC 算法通过对图像进行预处理完善字典, 使用较少的字典原子进行稀疏表示得到较高的识别率, 这说明基于低秩的联合动态稀疏表示的多观测样本分类算法的可行性.

3.2 ETH-80 物体识别

本组实验数据来自于 ETH-80 物体数据库. 该数据库包括 8 个种类, 分别为 apple, car, horse, dog, cup, cow, pear, tomato, 每个种类包含 10 个不同的物体类, 如图 4(a) 所示; 每个物体类又由 41 个样本组成, 这 41 个样本是由该物体在上半视角球面上经过均匀旋转而形成的, 图 4(b) 显示一种杯子所有样本. 本组实验采用 cropped-close128 数据, 大小是 128×128 .



(a) 8 个种类的 10 个物体类

(b) 一种杯子的 41 个样本

图4 ETH-80物体数据库示意图

实验中, 所有图像都归一化成 32×32 大小. 将该数据库的 8 个种类均作为已知类, 从每个种类的 10 个不同物体类中分别随机抽取出 20 幅图像作为该物体类的训练样本, 再从剩下的每个种类中每个物体类的 20 幅图像中随机选择 10 幅图像作为其相应物体类的测试图像. 并按照上述实验选取参数的方法选择合适的参数, 进行 10 次随机实验的平均结果如表 2. 本实验还分析了不同多观测样本数目的识别率, 结果如图 5 所示.

表 2 在 ETH-80 数据库上的平均正确识别率

算法	ISM-Graph-GEDA	JDSRC	LR-JDSRC
识别率	82.37	92.75	95.37
标准差	7.2	3.81	2.43

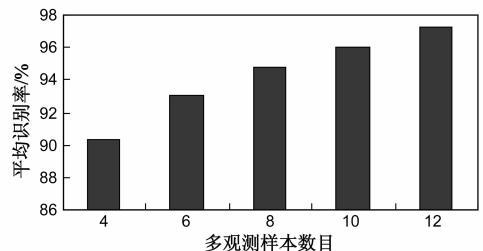


图5 不同数目的多观测样本的平均识别率

由表 2 结果分析可知, 本方法不仅平均识别率高于以前的经典方法, 并且标准差相对较低, 即该方法不仅辨别能力强且具有更强的鲁棒性; 由图 5 可知, 随着多观测样本数目的增多, 本文研究的 LR-JDSRC 算法的平均识别率升高, 由此可见多观测样本数目越多, 提供的关于某一个类别的信息量就越多, 从而判断出来的类别就更准确.

3.3 USPS 手写体数字识别

本组的实验数据采用 USPS 手写体数字数据库. 该数据库包含 0~9 共 10 组手写体数字, 每组数据包括相应数字的 1100 个灰度样本, 所有样本都归一化到 16×16 大小, 图 6 显示了该数据库的部分样本.



图6 手写体数字的部分样本

将该数据库的 10 个类别均看作已知类, 从每组数据中随机抽取 5 个数据组成训练样本, 而从每组剩下的 1095 个数据中随机抽取 $m = \{5, 10, 15, 20, 25\}$ 个数据组成多观测样本. 针对不同的 m 值, 按照上述两种实验选取参数的方法选择合适的值, 然后从每组数字中随机抽取 10 次训练样本和测试样本分别进行实验, 即每个类别均做 100 次随机实验, 实验结果中的每一个值都是 10 个类别进行 100 次随机实验的平均值, 图 7 显示了不同多观测样本数目不同方法对应的平均识别率.

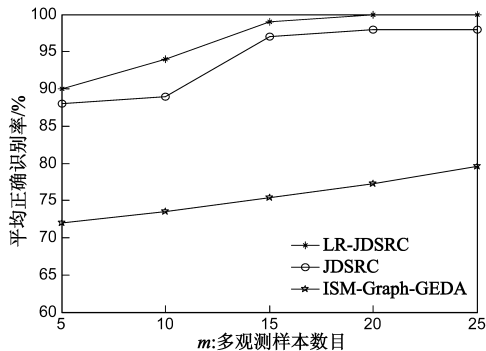


图7 在USPS手写体数字数据库的识别

通过对实验结果分析可知, 随着多观测样本数目的增加, 所有方法的识别率都在有所提高, 这说明在一定程度上多观测样本的数目会影响多观测样本所提供的信息量. 同时从图中也可以看出, 无论多观测样本数目如何变化, LR-JDSRC 算法的平均识别率始终高于其他两种算法. 这说明, 基于低秩的联合动态稀疏表示多观测样本分类算法能够更有效地利用多观测样本所提供的信息量, 得到更好的分类结果.

3.4 UMIST 人脸识别

本组的实验数据来自 UMIST 人脸数据库. 该数据库包含 20 个人的 564 幅图像, 其中每个人均有一系列从侧面到正面的图像, 如图 8 所示.

从每个人中随机选出 6 幅图片作为训练样本, 并从剩下的图像中随机抽取 $m = \{2, 4, 6, 8\}$ 个作为测试样

本, 在稀疏度为 6 时, 10 次随机实验的平均结果如表 3; 在稀疏度为 4 时, 10 次随机实验的平均结果如表 4; 在稀疏度为 3 时, 10 次随机实验的平均结果如表 5.



图8 UMIST人脸数据库中角度变化例子

表3 疏度 $K=6$ 时的平均识别率

观测数	ISM-Graph-GEDA		JDSRC		LR_JDSRC	
	识别率	标准差	识别率	标准差	识别率	标准差
2	75.5	9.85	81	14.3	92.5	8.25
4	74.75	8.37	94.5	11.41	96.5	4.12
6	71.67	7.83	84.5	17.07	93.5	7.47
8	72.5	6.56	92.5	14.58	96.5	4.74

表4 疏度 $K=4$ 时的平均识别率

观测数	ISM-Graph-GEDA		JDSRC		LR_JDSRC	
	识别率	标准差	识别率	标准差	识别率	标准差
2	75.5	9.85	90.5	5.99	91.5	5.3
4	74.75	8.37	97	6.32	96	4.59
6	71.67	7.83	90.5	9.85	92.5	7.17
8	72.5	6.56	99.5	2.11	99	1.58

表5 稀疏度 $K=3$ 时的平均识别率

观测数	ISM-Graph-GEDA		JDSRC		LR_JDSRC	
	识别率	标准差	识别率	标准差	识别率	标准差
2	75.5	9.85	92	8.23	92.5	6.35
4	74.75	8.37	95.5	6.43	95.5	2.84
6	71.67	7.83	94.5	7.98	94.5	4.97
8	72.5	6.56	97.5	5.4	98.5	2.42

由表 3 可知, 本文方法的识别率高于其他两种方法; 由表 4 可知, 在不同角度下, LR_JDSRC 的识别率或高于 JDSRC, 或低于 JDSRC; 由表 4 可知, LR_JDSRC 的识别率比 JDSRC 稍高或相等. 综上三表可知, LR_JDSRC 的识别率和 JDSRC 的识别率与稀疏度相关, 但 LR_JDSRC 相较于 JDSRC 总是有较高的稳定性.

4 结论

本文通过对无约束数据信息进行研究, 提出基于低秩的联合动态稀疏表示多观测样本分类算法 (LR-JDSRC). 该方法对训练样本和测试样本同时进行预处理, 分别将其分解成低秩矩阵和稀疏误差矩阵, 然后利用改进的字典对修正后的多观测样本分别进行联合动态

稀疏表示,充分利用类级的相关性和原子级的差异性,在类别上分享相同的稀疏模型,在原子级上采用不同的稀疏模型;最后利用加权后的总的稀疏重建误差进行分类判决.在 ETH-80 物体识别数据库、CMU-PIE 人脸数据库、USPS 手写体数字数据库和 UMIST 人脸数据库上进行对比实验,实验结果验证了该方法的良好性能,但是,相对于传统稀疏表示算法,该方法进行低秩分解时复杂度相对较高,时间成本较大.

参考文献

- [1] Dong Xu, Yi Huang, Zinan Zeng, Xinxing Xu. Human gait recognition using patch distribution feature and locality-constrained group sparse representation[J]. IEEE Transactions on Image Processing, 2012, 21(1): 316 – 326.
- [2] Zhao Nan, Xu Xin, Yang Yi. Sparse representations for speech enhancement[J]. Chinese Journal of Electronics, 2011, 20(2): 268 – 272.
- [3] Zhu Hao, Giannakis Georgios B. Sparseovercomplete representations for efficient identification of power line outages[J]. IEEE Transactions on Power Systems, 2012, 27(4): 2215 – 2224.
- [4] Wright John, Yang Allen Y, Ganesh Arvind, Ma Yi. Robust face recognition via sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210 – 227.
- [5] Zhang Lihe, Zhang Kunyu. Weighted discriminative sparse coding for image classification[J]. Chinese Journal of Electronics, 2014, 23(1): 104 – 108.
- [6] Huang Ke, Aviyente Selin. Sparse representations for signal classification[A]. 20th Annual Conference on Neural Information Processing Systems[C]. Vancouver BC: NIPS, 2007. 609 – 616.
- [7] Elad Michael, Aharon Michal. Image denoising via sparse and redundant representations over learned dictionaries[J]. IEEE Transactions on Image Processing, 2006, 15(12): 3736 – 3745.
- [8] Ramirez Ignacio, Sprechmann Pablo, Sapiro Guillermo. Classification and clustering via dictionary learning with structured incoherence and shared features[A]. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition[C]. San Francisco: IEEE, 2010. 3501 – 3508.
- [9] Chia-Po Wei, Yu-Wei Chao, Yi-Ren Yeh, Yu-Chiang Frank Wang. Locality-sensitive dictionary learning for sparse representation based classification[J]. Pattern Recognition, 2013, 46(5): 1277 – 1287.
- [10] He Ran, Zheng Wei-shi, Hu Bao-Gang, Kong Xiang-Wei. Nonnegative sparse coding for discriminative semi-supervised learning[A]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition[C].

Colorado Springs: IEEE, 2011. 2849 – 2856.

- [11] Chao Yu-Wei, Yeh Yi-Ren, Chen Yu-Wen, Lee Yuh-Jye, Wang Yu-Chiang Frank. Locality-constrained group sparse representation for robust face recognition[A]. IEEE International Conference on Image Processing[C]. Brussels: IEEE, 2011. 761 – 764.
- [12] Yuan Xiaotong, Yan Shuicheng. Visual classification with multi-task joint sparse representation[A]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition[C]. San Francisco: IEEE, 2010. 3493-3500.
- [13] Zhang Haichao, Nasrabadi Nasser M, Zhang Yanning, Huang Thomas S. Multi-observation visual recognition via joint dynamic sparse representation[A]. IEEE International Conference on Computer Vision[C]. Barcelona: IEEE, 2011. 595 – 602.
- [14] Zhang Haichao, Nasrabadi Nasser M, Zhang Yanning, Huang Thomas S. Joint dynamic sparse representation for multi-view face recognition[J]. Pattern Recognition, 2012, 45(4): 1290 – 1298.
- [15] Chen Chih-Fan, Wei Chia-Po, Wang Yu-Chiang Frank. Low-rank matrix recovery with structural incoherence for robust face recognition[A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. Providence: IEEE, 2012. 2618 – 2625.
- [16] Peng Yigang, Ganesh Arvind, Wright John, Xu Wenli, Ma Yi. RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated Images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2233 – 2246.
- [17] Duarte Marco F, Cevher Volkan, Baraniuk Richard G. Model-based compressive sensing for signal ensembles[A]. Annual Allerton Conference on Communication, Control, and Computing[C]. Monticello: IEEE, 2009. 244 – 250.

作者简介



胡正平 男, 1970 年生于四川仪陇县, 教授, 博士生导师, 燕山大学通信电子工程系主任, 1996 年于燕山大学无线电专业获得学士学位, 并获得推荐研究生资格, 1999 年获得电路与系统硕士学位, 2007 年于哈尔滨工业大学获得信息与通信工程专业博士学位, 目前为中国电子学会高级会员, 中国图像图形学会高级会员, 目前研究方向为稀疏模式识别.

E-mail: hzp@ysu.edu.cn

高红霄 女, 1987 年生于河北石家庄, 燕山大学信息科学与工程学院通信与信息系统专业, 主要研究方向为多观测样本分类.

E-mail: 970353066@qq.com