

基于类别相关近邻子空间的 最大似然稀疏表示 鲁棒图像识别算法

胡正平^{1,2} 宋淑芬¹

摘要 为了构建一个快速鲁棒的图像识别算法, 提出基于类别相关近邻子空间的 L_1 最大似然稀疏表示图像识别算法. 考虑到每个测试样本的不同分布特性及训练样本选择的类别代表性原则, 不再将所有训练样本作为稀疏表示的字典, 而是基于距离相近准则选择合适子空间, 从每个类别中选取自适应数量的局部近邻构成新的字典, 在减少训练样本的同时保留了稀疏表示原有的子空间结构. 然后基于最大似然稀疏表示识别方法, 将稀疏表示的保真度表示为余项的最大似然函数, 并将识别问题转化为加权的稀疏优化问题. 在公用人脸与数字识别数据库上的实验证明该算法的合理性, 提高识别速度的同时保证了识别精度和算法的鲁棒性, 特别是对于遮挡与干扰图像具有较好的适应性.

关键词 图像识别, 稀疏表示, 类别相关子空间, 近邻选择, 最大似然估计

引用格式 胡正平, 宋淑芬. 基于类别相关近邻子空间的 L_1 最大似然稀疏表示鲁棒图像识别算法. 自动化学报, 2012, 38(9): 1420–1427

DOI 10.3724/SP.J.1004.2012.01420

Robust Image Recognition Algorithm of Maximum Likelihood Estimation Sparse Representation Based on Class-related Neighbors Subspace

HU Zheng-Ping^{1,2} SONG Shu-Fen¹

Abstract In order to construct a fast and robust image recognition algorithm, an image recognition algorithm of maximum likelihood estimation sparse representation based on class-related neighbors subspace is proposed in this paper. Considering the different distribution characteristics of each test sample and the class-representative principle of training samples' selection, instead of constructing the dictionary of sparse representation by all training samples, suitable subspace is selected and local neighbors of adaptive number that is selected from each class are used to construct the new dictionary based on distance proximity criterion. The training samples are reduced and the original subspace structure of sparse representation is kept at the same time. Then based on the recognition method of maximum likelihood sparse representation, the fidelity of sparse representation is represented by the maximum likelihood function of residuals and the recognition problem is converted to a weighted sparse optimization problem. Experiments results on public available face and handwritten digital databases verify the rationality, recognition speed, and recognition accuracy of the proposed algorithm. The algorithm is robust, especially it can work for in disturbed and occluded images.

Key words Image recognition, sparse representation, class-related subspace, neighbors' selection, maximum likelihood estimation

Citation Hu Zheng-Ping, Song Shu-Fen. Robust image recognition algorithm of maximum likelihood estimation sparse representation based on class-related neighbors subspace. *Acta Automatica Sinica*, 2012, 38(9): 1420–1427

收稿日期 2011-09-01 录用日期 2012-02-22
Manuscript received September 1, 2011; accepted February 22, 2012

国家自然科学基金 (61071199), 河北省自然科学基金 (F2010001297), 中国博士后自然科学基金 (20080440124), 第二批中国博士后基金 (200902356) 资助

Supported by National Natural Science Foundation of China (61071199), Natural Science Foundation of Hebei Province (F2010001297), Postdoctoral Science Foundation of China (20080440124), and the Second Batch of China Postdoctoral Science Foundation (200902356)

本文责任编辑 贾云得

Recommended by Associate Editor JIA Yun-De

1. 燕山大学信息科学与工程学院 秦皇岛 066000 2. 河北省计算机虚拟技术与系统集成重点实验室 秦皇岛 066000

1. School of Information Science and Engineering, Yanshan University, Qinhuangdao 066000 2. Key Laboratory for Com-

puter Virtual Technology and System Integration of Hebei Province, Qinhuangdao 066000

寻求客观事物的稀疏表示方法, 一直是计算机视觉、图像处理、图像识别领域的研究目标. 特别是近来以 L_1 , $L_{1/2}$ 优化为核心的信号稀疏表示思路, 这种思路已经蔓延到图像识别领域. 可以预见, 基于稀疏表示的图像识别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图像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图像之间的相似度.

实用图像识别系统通常应同时具有较好的鲁棒性和实时性: 鲁棒性即能够在许多非理想条件 (光照、遮挡等) 下取得较好的识别效果, 实时性即能够

puter Virtual Technology and System Integration of Hebei Province, Qinhuangdao 066000

在较少运算量的情况下实现快速识别, 构建快速鲁棒的图像识别算法属于关键技术的研究. 目前一些学者将稀疏表示用于图像识别^[1], 该方法将识别问题看作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, 取得了较好的识别效果, 引起了国内外学者的极大兴趣. 为了构建更加鲁棒的稀疏表示识别算法, 针对该方法对图像之间的对齐性要求较高的问题, 文献 [2] 提出具有平移不变性的稀疏表示分类方法, 在求解稀疏表示的同时得到平移参数, 文献 [3] 以稀疏表示作为工具将测试图像训练图像对齐, 并构建光照模型, 实现对于光照变化和图像不对齐等情况的鲁棒性. 文献 [4–5] 则均针对稀疏表示保真度的表达式进行了改进, 保真度由稀疏表示的余项表示, 它代表稀疏表示字典的线性组合对测试图像的逼近程度, 对测试图像的正确分类起着至关重要的作用, 文献 [1–4] 分别将保真度的表达式表示为余项的 L_2 范数和 L_1 范数, 两种方法都限定了余项服从特定的分布, 影响了算法的鲁棒性, 因此, 文献 [5] 将保真度的表达式改写为余项的似然分布函数, 相对于文献 [1–4], 这是一个更加通用的表达式, 可以构建出更鲁棒的稀疏表示模型.

为了得到对测试图像最稀疏的表示, 通常稀疏表示识别方法计算复杂度、成本大, 运算速度慢, 特别是在训练样本较多的情况下, 求解过程更加耗时, 为改进运算速度, 一种有代表性的方法是将由所有训练样本构成的稀疏表示字典进行压缩. 文献 [6] 通过训练一种用于检测图像中非人脸像素的度量, 将训练集过滤为一个较小的集合, 文献 [7] 则基于同一类的训练图像可建模为一个小的样本集的理论, 通过 K-SVD 算法设计过完备字典, 并将字典进一步压缩为更小的集合, 很大程度上缩短了运算时间, 但该方法造成了识别精度较明显的下降. 文献 [8–9] 则将字典减少为测试图像固定数量的近邻进行稀疏表示, 该方法原理清晰, 并达到了较好的识别效果, 但没有考虑测试样本的不同分布特性及训练样本选择的类别代表性原则, 导致不能合理选择原样本空间的近邻子空间, 进而影响识别精度.

本文提出基于类别相关自适应近邻子空间的快速鲁棒稀疏表示图像识别算法, 充分考虑到每个测试样本的不同分布特性及训练样本选择的类别代表性, 选择合理的近邻子空间, 在减少训练样本的同时保留了稀疏表示原有的子空间结构, 然后基于最大似然估计的稀疏表示图像识别方法, 构建了一个快速鲁棒的稀疏表示识别算法.

1 基于稀疏表示的分类方法

1.1 稀疏表示的分类思想

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^[1], 通过优化求解, 用最稀

疏的训练样本表示测试图像, 从而判断其类别. 定义包含所有训练样本的字典 $A = [A_1, \dots, A_k] \in \mathbf{R}^{m \times n}$, 其中 k 表示类别数目, m 表示样本维数, n 表示样本总数, $A_i = [v_{i,1}, \dots, v_{i,n_i}] \in \mathbf{R}^{m \times n_i}$ 表示第 i 类的训练样本矩阵, n_i 表示第 i 类的训练样本数目.

对某一未知类别的测试样本 $\mathbf{y} \in \mathbf{R}^m$ 的稀疏表示分类可分为以下几步:

- 1) 预处理: 将字典 A 按列归一化, 使各列具有统一的 L_2 范数;
- 2) 将测试样本表示为字典 A 的线性组合, 由 L_1 范数最小化求解得到稀疏系数 $\hat{\mathbf{x}}_1 \in \mathbf{R}^n$

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

式中, ϵ 表示误差容限, 为了保证稀疏表示的保真度, 作为 L_1 范数优化的约束条件要求余项 $\mathbf{e} = \mathbf{y} - A\mathbf{x}$ 满足 $\|\mathbf{e}\|_2 \leq \epsilon$;

- 3) 计算各类别样本对测试样本的逼近残差

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

式中, $\delta_i(\hat{\mathbf{x}}_1)$ 表示 $\hat{\mathbf{x}}_1$ 中对应于第 i 类样本空间的系数向量;

- 4) 将测试样本归为具有最小逼近残差的类别

$$\text{identity}(\mathbf{y}) = \arg \min_{i \in \{1, \dots, k\}} (r_i(\mathbf{y})) \quad (3)$$

1.2 最大似然鲁棒稀疏表示识别模型

稀疏表示的分类方法将稀疏表示的保真度表示为余项 L_2 的范数, 但最大似然估计理论的证明这样的假设要求余项服从高斯分布, 实际中这样的假设可能并不成立, 特别是当测试图像中存在噪声、遮挡等异常像素, 这就导致传统的保真度表达式所构造的稀疏表示模型对上述这些情况缺少足够的鲁棒性. 而文献 [5] 中的最大似然稀疏表示识别模型则基于最大似然估计理论, 将保真度表达式改写为余项的最大似然分布函数, 并将最大似然问题转化为一个加权优化问题, 在稀疏表示的同时引入代表各像素不同权值的矩阵, 使得该算法对于图像中包含异常像素的情况表现出较好的鲁棒性.

最大似然估计稀疏表示分类方法具体思想如下:

- 1) 构建最大似然估计识别模型

首先将字典 A 改写为 $A = [\mathbf{d}_1; \dots; \mathbf{d}_m]$, 其中行向量 \mathbf{d}_i 表示 A 的第 i 行, 余项 $\mathbf{e} = \mathbf{y} - A\mathbf{x} = [e_1; \dots; e_m]$, 其中 $e_i = y_i - \mathbf{d}_i\mathbf{x}, i = 1, \dots, m$. 假设 e_1, \dots, e_m 独立同分布, 设其概率密度函数为 $f_{\theta}(e_i), \theta$ 表示该分布的参数设置, 则余项 \mathbf{e} 的似然函数可表示为

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

最大似然估计理论即最大化式 (4) 的似然函数, 等价于最小化式 (5) 中的目标函数 $F_{\theta}(\mathbf{e})$

$$F_{\theta}(\mathbf{e}) = -\ln L_{\theta} = \sum_{i=1}^m \rho_{\theta}(e_i)$$

其中,

$$\rho_{\theta}(e_i) = -\ln f_{\theta}(e_i) \quad (5)$$

考虑对系数 \mathbf{x} 的稀疏约束, 与传统的稀疏表示识别模型比较, 最大似然估计的稀疏表示识别模型可表示为

$$\hat{\mathbf{x}}_1 = \arg \min F_{\theta}(\mathbf{y} - \mathbf{A}\mathbf{x}) \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq \sigma \quad (6)$$

观察可知, 式 (6) 的模型本质是一个稀疏约束的最大似然估计问题. 最小收缩-选择算子 (Least absolute shrinkage and selection operator, LASSO)^[10], 主要用于解决一类有约束的优化问题, 可表示为

$$\hat{\mathbf{x}}_1 = \arg \min \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq \sigma \quad (7)$$

其中, σ 是一个大于零的常数. 比较可知, 式 (7) 与式 (1) 中的稀疏优化问题是对偶问题, 两者是等价的, 因此, 可用 LASSO 问题等价代替式 (1) 的 L_1 范数优化问题. 故式 (6) 与式 (7) 比较可知, 式 (7) 中 $F_{\theta}(\mathbf{e}) = \|\mathbf{e}\|_2$, 则 $f_{\theta}(e_i) = \exp(-e_i)$, 即要求余项服从高斯分布, 因此式 (7) 仅是式 (6) 的一个特例, 也就是说, 相对于式 (1) 的稀疏表示识别方法, 式 (6) 的稀疏约束最大似然估计模型是一个更加通用的稀疏表示识别模型.

2) 加权优化求解稀疏表示

通常假设概率密度函数 $f_{\theta}(e_i)$ 是对称的, 并满足当 $|e_i| > |e_j|$ 时 $f_{\theta}(e_i) < f_{\theta}(e_j)$. 因此, $\rho_{\theta}(e_i)$ 有以下特性: a) $\rho_{\theta}(0)$ 是 $\rho_{\theta}(e_i)$ 的全局最小值, 即 $\rho_{\theta}'(0) = 0$; b) $\rho_{\theta}(e_i) = \rho_{\theta}(-e_i)$; c) 当 $|e_i| > |e_j|$ 时, $\rho_{\theta}(e_i) > \rho_{\theta}(e_j)$.

$F_{\theta}(\mathbf{e})$ 在定义域内某点 \mathbf{e}_0 处的一阶泰勒展式可表示为

$$F_{\theta}'(\mathbf{e}) = F_{\theta}'(\mathbf{e}_0) + (\mathbf{e} - \mathbf{e}_0)^T F_{\theta}''(\mathbf{e}_0) + \mathbf{R}_1(\mathbf{e}) \quad (8)$$

其中, $F_{\theta}'(\mathbf{e}_0) = [\rho_{\theta}'(e_{0,1}); \dots; \rho_{\theta}'(e_{0,m})]$, $\mathbf{R}_1(\mathbf{e})$ 表示高阶余项, 设 $\mathbf{R}_1(\mathbf{e}) = \frac{1}{2}(\mathbf{e} - \mathbf{e}_0)^T \mathbf{W}(\mathbf{e} - \mathbf{e}_0)$, 其中, \mathbf{W} 是一个对角阵. 由式 (8) 对余项 \mathbf{e} 求导可得:

$$F_{\theta}''(\mathbf{e}) = F_{\theta}''(\mathbf{e}_0) + \mathbf{R}_1'(\mathbf{e}) \quad (9)$$

其中, $\mathbf{R}_1'(\mathbf{e}) = [W_{1,1}(e_{i,1} - e_{0,1}); \dots; W_{n,n}(e_{i,m} - e_{0,m})]$, 由 $\rho_{\theta}'(0) = 0$ 可得:

$$W_{i,i} = \frac{\rho_{\theta}''(e_{0,i})}{e_{0,i}} \quad (10)$$

$W_{i,i}$ 是一个非负的标量, 因此, $F_{\theta}(\mathbf{e})$ 可改写为 $F_{\theta}(\mathbf{e}) = \frac{1}{2}\|\mathbf{W}^{1/2}\mathbf{e}\|_2^2 + b$, b 是一个由 \mathbf{e}_0 决定的常数, 因此可将式 (6) 改写为

$$\hat{\mathbf{x}}_1 = \arg \min \|\mathbf{W}^{1/2}(\mathbf{y} - \mathbf{A}\mathbf{x})\|_2^2 \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq \sigma \quad (11)$$

由式 (11) 与式 (7) 比较可知, 最大似然稀疏表示分类方法将原稀疏表示模型转化为一个加权的稀疏表示模型, 权值矩阵 \mathbf{W} 有明确的物理意义, $W_{i,i}$ 表示分配给每个像素的不同权值, 很明显对于异常像素应赋予较低的权值. 因此, 选择与其具有相似性质 SVM 的 hinge loss 函数^[11] 作为权值函数, 通过不断迭代优化更新 \mathbf{W} , 直至收敛, 求得最终的稀疏表示系数, 最后同样根据最小逼近残差准则实现测试图像的分类.

2 基于类别相关近邻子空间的快速鲁棒稀疏表示识别算法

为了得到测试图像的稀疏表示, 通常稀疏优化求解算法计算量较大, 而最大似然稀疏表示识别算法, 通过不断迭代优化求得权值矩阵, 在提高算法鲁棒性的同时, 很大程度增加了计算量. 因此, 为了在保证识别精度的同时提高识别速度, 本文提出一种快速鲁棒稀疏表示图像识别算法, 不再是将所有训练样本作为稀疏表示的字典, 而是合理选择原样本空间的近邻样本子空间, 构成新的字典, 然后基于最大似然稀疏表示模型, 实现对测试图像的分类. 图 1 为本文提出的快速稀疏表示识别算法的框图.

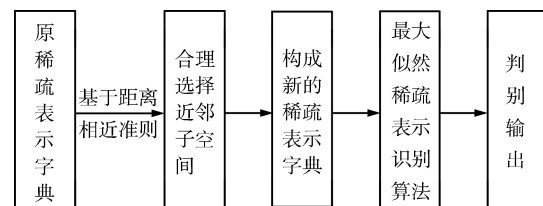


图 1 快速稀疏表示识别算法框图

Fig. 1 Diagram of fast recognition algorithm based on sparse representation

传统的近邻选择方法依据距离相近准则, 选择测试样本的固定数目近邻作为新的字典^[8]. 但该方法有明显的缺点: 1) 不同近邻数目对识别结果有很大影响, 缺乏对不同数据的适应性; 2) 对所有的测试样本选择同一个近邻数目, 没有考虑每个测试样本的不同分布特性, 一些比较“简单”的测试样本, 仅需要少数的近邻即可达到较高的识别效果, 而一些比较“困难”的样本则需要较大数量的近邻. 因此, 提出对每个测试样本自适应选择不同数量近邻的方法, 首先求得测试样本与所有训练样本的平均距离, 然后仅选择小于或等于平均距离的训练样本, 实现

每个测试样本自适应数量的近邻选择.

在图像识别中, 样本选择通常应满足代表性和典型性的原则, 即应尽量选择各类别中有代表性的样本, 并尽量避免同一类的很多重复选择, 使得所选样本更具有“价值”. 而上述近邻选择方法, 都是仅依靠距离相近准则在全体训练样本中选择近邻, 使得某些类别的样本被大量重复选择, 而某些类别的样本没有被选择, 破坏了原样本空间的结构, 一定程度影响识别精度. 因此, 本文提出基于类别相关的近邻子空间选择方法, 类别相关就是无论与测试样本距离较近或较远的类别, 都参与样本选择, 充分体现样本选择的类别代表性, 并同样采用自适应原则, 该方法构成的样本子空间基本保留了原样本空间的结构.

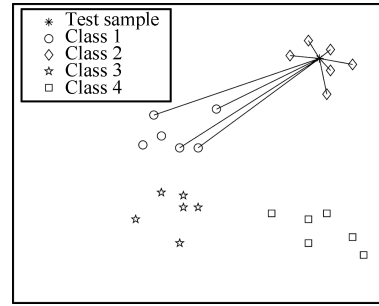
选择类别相关近邻子空间方法的具体过程如下: 设类别数目为 k , 则分别对每个类别自适应的选择近邻, 如某类别 i , 共有 n_i 个训练样本, 首先求得测试样本与该类训练样本的平均距离 $\bar{D}_i = \text{mean}(\text{dist}(\mathbf{y}, A_{i,h})), h = 1, \dots, n_i$, $\text{dist}(\cdot)$ 表示距离度量, 本文采用欧氏距离作为度量, $\text{mean}(\cdot)$ 表示求平均值, 然后在该类别中选择小于或等于平均距离 \bar{D}_i 的训练样本. 对每类都进行同样的样本选择操作, 得到新的样本子空间, 构成新的字典, 即

$$A = \{A_j | j = p_1, \dots, p_N\} \quad (12)$$

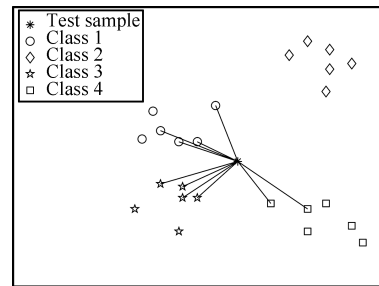
其中, $[p_1, \dots, p_N] = \arg_{i=1, \dots, k, h=1, \dots, n_i} \text{dist}(\mathbf{y}, A_{i,h}) \leq \bar{D}_i$, N 表示测试样本选择的近邻数量.

为了进一步阐释上述近邻选择方法, 将传统和自适应的近邻选择及基于类别相关的近邻子空间选择方法分别用图 2~图 4 进行表示, 设共有 4 个类别构成原样本空间, 且理论上来说同类样本的数据点在高维空间中也应是相关的, 则同类样本的数据应该位于该类数据所构成的子空间中, 从而形成了图中包含类别信息的原稀疏表示的样本子空间结构. 图中不同类别分别用不同标志表示, 每类有 6 个训练样本, “*” 代表测试样本, 每种方法以 (a)、(b) 两图中两个不同的测试样本分别表示.

由图 2 和图 3 比较可知, 子图 (a) 中的测试样本仅与少数类别的训练样本距离较近, 属于较“简单”的测试样本, 而子图 (b) 中的测试样本同时与很多类别的训练样本距离接近, 属于较“困难”的测试样本, 因此图 3 中分别选择 7 个和 14 个训练样本作为其近邻, 而图 2 则统一将近邻数目固定为 10. 可见, 自适应近邻选择方法实现对不同测试样本的区别处理. 由图 4 与图 2、图 3 比较可知, 基于类别相关的近邻子空间选择方法, 不再仅依赖样本个体之间的相似性, 而是自适应地选择各个类别中有代表性的样本, 相对于图 2 和图 3, 图 4 形成了与原样本空间的分布最接近的近邻子空间.



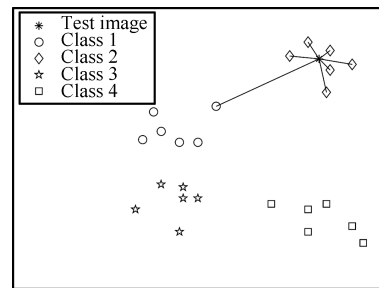
(a) “简单”测试样本的传统近邻选择
(a) Traditional neighbors' selection of easy test image



(b) “困难”测试样本的传统近邻选择
(b) Traditional neighbors' selection of difficult test image

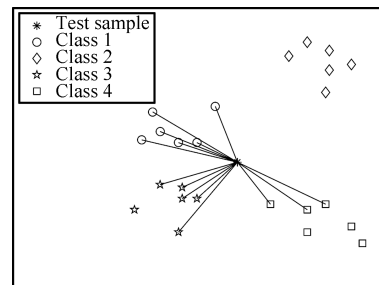
图 2 传统的近邻选择方法示意图 ($N = 10$)

Fig. 2 Sketch maps of traditional method for neighbors' selection ($N = 10$)



(a) “简单”测试样本的自适应近邻选择

(a) Adaptive neighbors' selection of easy test image

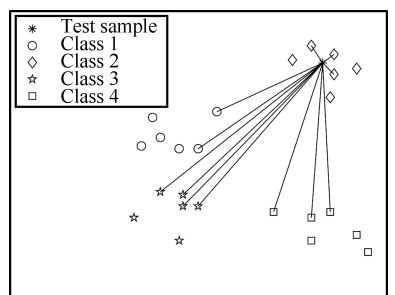


(b) “困难”测试样本的自适应近邻选择

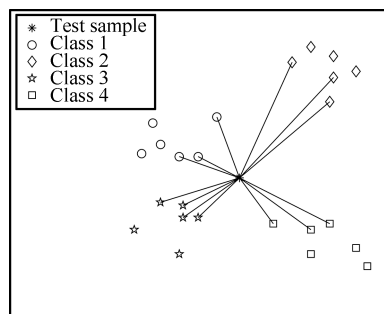
(b) Adaptive neighbors' selection of difficult test image

图 3 自适应近邻选择方法示意图

Fig. 3 Sketch maps of adaptive method for neighbors' selection



(a) “简单”测试样本的类别相关近邻选择
(a) Class-related neighbors' selection of easy test image



(b) “困难”测试样本的类别相关近邻选择
(b) Class-related neighbors' selection of difficult test image

图 4 基于类别相关的近邻子空间选择方法示意图

Fig. 4 Sketch maps of the method for neighbors subspace selection based on the class-related

稀疏表示的识别算法本质是在由所有类别的训练样本构成的高维字典空间中用最稀疏的原子最好的重构测试样本, 该识别过程依赖于整个字典的子空间结构, 即由于所有训练样本类别的存在, 使得稀疏优化算法可以在测试样本与各个类别样本的差异性和相似性中做出比较, 从而选择出最相似的样本来表示测试样本。

而传统的样本选择方法, 在整个字典中寻找测试样本的近邻, 会造成仅选择字典中的几类甚至是一类作为新的字典, 这样的样本选择会导致字典内的类别信息减少, 所选样本缺乏类别代表性。本文提出的基于类别相关的近邻子空间选择方法在每个样本子空间中均保留部分样本, 保证了原字典的样本子空间不被破坏, 仅自适应地减少各子空间中的一些冗余样本 (如图 4(a) 去掉了类别 2 的一些重复样本), 并引入与测试样本差异较大的样本类别 (如图 4(a) 引入了类别 3 和类别 4 的一些样本), 差异性较大的样本可以更加突出某一类别与测试样本的相似性, 从而更加容易获得测试样本的稀疏表示, 相对于传统的近邻选择和自适应近邻选择方法, 更有利于测试样本的正确识别。

3 实验仿真研究

为了检验上述算法的有效性, 本文采用 AR 人

脸库和 USPS 手写数字库进行实验, 为了验证算法对于包含异常像素图像的鲁棒性, 采用 AR 数据库中经过伪装 (戴墨镜、戴围巾) 的人脸子集作为测试样本进行实验。实验所用平台为 AMD Athlon (TM)2 处理器, 2.9 G 主频, 2 G 内存。

3.1 AR 人脸库

选用 AR 人脸库中 100 个目标, 每个目标有 14 张图片, 前 7 张用于训练, 后 7 张用于测试, 数据包括人脸的多种表情和光照等细节变化, 将所有图像剪裁为 60 像素 \times 43 像素, 并对图像进行下采样实现降维, 实验中采用 1/2 的下采样率, 并将数据经过规范化处理。

表 1 列出了 AR 数据库上几种算法识别率的比较, 可以看出基于稀疏表示识别方法和最大似然稀疏表示识别方法优于其他分类器的识别效果, 证明了稀疏表示分类思想的合理性, 且由于 AR 人脸库中存在光照、表情变化等, 最大似然稀疏表示识别方法表现出了更好的鲁棒性, 达到了 96.9% 的最高识别率, 相对于文献 [1] 的稀疏表示识别算法提高了近 4%。

表 1 AR 库上几种方法识别率的比较

Table 1 Comparison of recognition rate of several methods on AR database

实验方法	识别率 (%)
NN (Nearest neighbor)	73.7
NS (Nearest subspace)	82.1
基于稀疏表示识别方法	93.1
基于最大似然估计的稀疏表示识别方法	96.9

图 5 列出了各种近邻选择稀疏表示识别方法在 AR 数据库上的比较, 分别列出了固定近邻数目的传统样本选择 (50-N、100-N、200-N、300-N、400-N 分别代表选择的近邻数目为 50、100、200、300、400)、自适应近邻选择 (A-N)、不经过样本选择的最大似然估计稀疏表示识别方法 (MLE) 和基于类别相关的近邻子空间选择 (C-N-S) 四种方法的识别精度和完成所有测试样本的识别所需的运算时间。

由图 5 可以看出, 最大似然估计稀疏表示识别算法识别精度最高, 但耗时也最长, 运算时间是其他方法的几倍或十几倍; 固定数目的近邻选择方法, 随着所选近邻数目的增加, 识别精度得到提高, 运算时间也随之增加, 且近邻数目的设定没有很合理的参考值; 自适应近邻选择方法达到了较高的识别精度, 但与类别相关的近邻子空间选择方法在识别精度上仍有差距; 而类别相关的近邻子空间选择方法则在运算时间缩短大于 2 倍的情况下, 达到了与最大似

然估计识别算法相同的识别精度, 从而证明基于类别相关的近邻子空间选择在识别精度和运算时间上达到了较合理的平衡.

近邻选择方法的有效性.

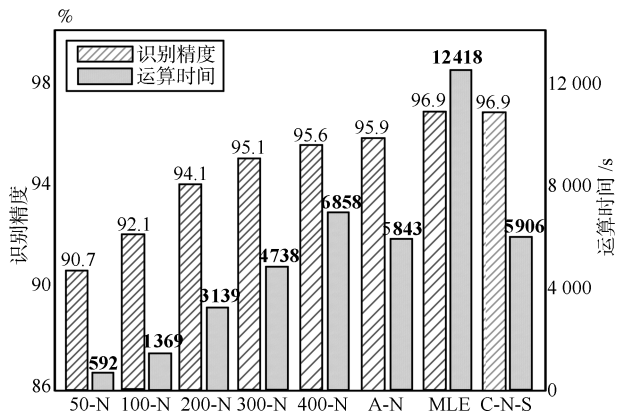


图 5 AR 库上各种近邻选择方法识别率和运行时间的比较
Fig. 5 Comparison of recognition rate and run time of methods for neighbors' selection on AR database

3.2 USPS 手写数字库

本组实验选择 USPS 手写数字库作为实验对象, 该数据库包含 0~9 共 10 个数字的手写体图像, 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, 图像大小为 16 像素 × 16 像素, 将数据经过规范化处理. 表 2 列出了 USPS 手写数字库上几种算法识别率的比较, 可以看出相对于其他分类器, 文献 [1] 的稀疏表示识别算法与基于最大似然估计的识别方法达到了较好的识别效果.

表 2 USPS 手写数字库上几种方法识别率的比较
Table 2 Comparison of recognition rate of several methods on USPS database

实验方法	识别率 (%)
NN (Nearest neighbor)	91.3
NS (Nearest subspace)	91.8
基于稀疏表示识别方法	92.1
基于最大似然估计的稀疏表示识别方法	93.1

图 6 列出了各种近邻选择方法在 USPS 数据库上的比较, 由图中可以看出, 最大似然估计稀疏表示识别算法识别精度最高, 但耗时却是其他方法的几倍甚至几十倍; 固定数目的近邻选择方法, 在近邻达到一定数量的时候, 识别精度的提高有限, 可以看出传统的近邻选择方法由于没有选择合适的子空间, 识别精度受到了限制; 本文提出基于类别相关的近邻子空间选择方法则在运算时间缩短接近 3 倍的同时, 达到了较高的识别精度, 从而证明基于类别相关

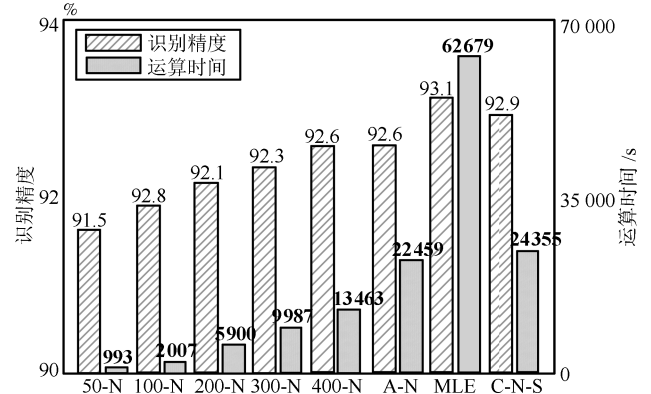


图 6 USPS 库上各种近邻选择方法识别率和运行时间的比较

Fig. 6 Comparison of recognition rate and run time of methods for neighbors' selection on USPS database

3.3 AR 经过伪装的人脸库

采用 AR 数据库中经过伪装 (戴墨镜、戴围巾) 的人脸子集进行实验, 分别选取 100 个目标的 200 张戴墨镜和 200 张戴围巾的图像作为测试图像, 将所有图像剪裁为 42 像素 × 30 像素, 并将数据经过规范化处理. 其中墨镜大约遮挡了 20% 的人脸, 围巾则大约遮挡了 40% 的人脸图像, 如图 9 和图 10 的子图 (a) 所示, 对识别造成很大的难度. 表 3 和表 4 分别列出了戴墨镜和戴围巾 AR 人脸图像库上几种算法识别率的比较, 可以看出基于最大似然估计的稀疏表示识别方法明显优于其他分类器的识别效果. 因此, 对于图像中包含异常像素的情况, 该方法表现出了更好的鲁棒性.

图 7 和图 8 列出了各种近邻选择方法在戴墨镜、戴围巾 AR 人脸库上的比较, 由图中可以看出, 最大似然估计稀疏表示识别算法识别精度最高, 分别为 100% 和 96.5%, 但是由于该算法不断迭代优化得到权值矩阵, 因此耗时极长; 由于戴墨镜的人脸库识别难度相对较小, 因此传统的近邻选择方法也达到了较高的识别精度, 但对于戴围巾的人脸库, 识别难度较大, 该方法没有得到较好的识别精度, 同样, 在此数据库上, 自适应的近邻选择方法也没有达

表 3 戴墨镜 AR 人脸图像库上几种方法识别率的比较
Table 3 Comparison of recognition rate of several methods on AR database with sunglasses

实验方法	识别率 (%)
NN (Nearest neighbor)	70.0
NS (Nearest subspace)	64.5
基于稀疏表示识别方法	87.0
基于最大似然估计的稀疏表示识别方法	100.0

表 4 戴围巾 AR 人脸图像库上几种方法识别率的比较
Table 4 Comparison of recognition rate of several methods on AR database with scarf

实验方法	识别率 (%)
NN (Nearest neighbor)	52.0
NS (Nearest subspace)	52.5
基于稀疏表示识别方法	69.5
基于最大似然估计的稀疏表示识别方法	96.5

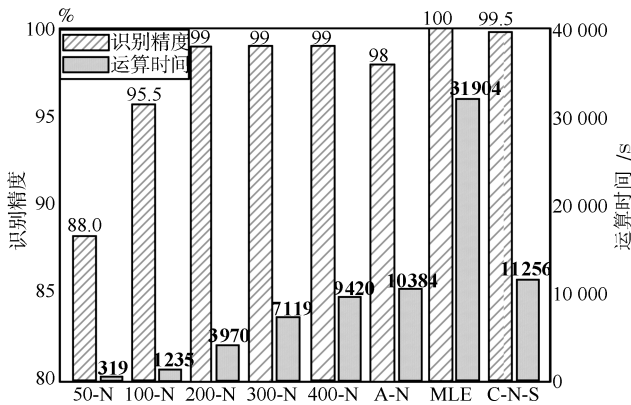


图 7 戴墨镜 AR 库上各种近邻选择方法识别率和运行时间的比较

Fig. 7 Comparison of recognition rate and run time of methods for neighbors' selection on AR database with sunglasses

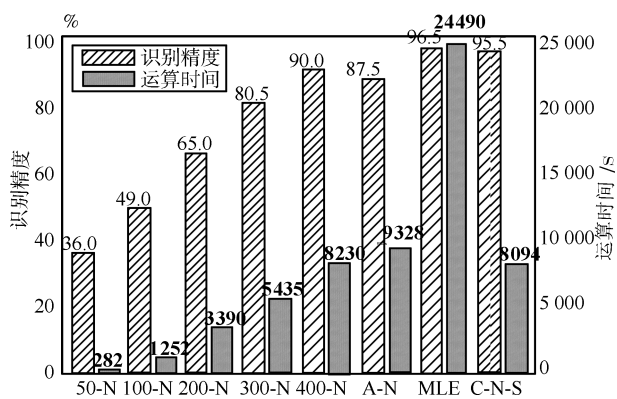


图 8 戴围巾 AR 库上各种近邻选择方法识别率和运行时间的比较

Fig. 8 Comparison of recognition rate and run time of methods for neighbors' selection on AR database with scarf

到较好的识别效果; 而本文提出的基于类别相关的近邻子空间选择方法则在运算时间大大缩短的情况下, 达到了与最大似然稀疏表示识别算法基本相同的识别精度, 分别为 99.5% 和 95.5%. 由此证明基于类别相关的近邻选择方法由于选择了合适的近邻

子空间, 对原样本空间合理压缩, 无论对于图像中包含异常像素或存在光照、表情等变化的情况, 都在大大缩短运算时间的同时基本保持了原算法的识别效果.

图 9 和图 10 列出了两幅戴墨镜和戴围巾的人脸图像, 在经过类别相关近邻子空间选择前后, 所得到的权值 W 的二值图, 可见对于包含异常像素的区域赋予了较低的权值, 比较子图 (b) 和子图 (c) 可知经过类别相关近邻选择前后基本无差别, 由此证明本文提出的算法保持了较好的鲁棒性.

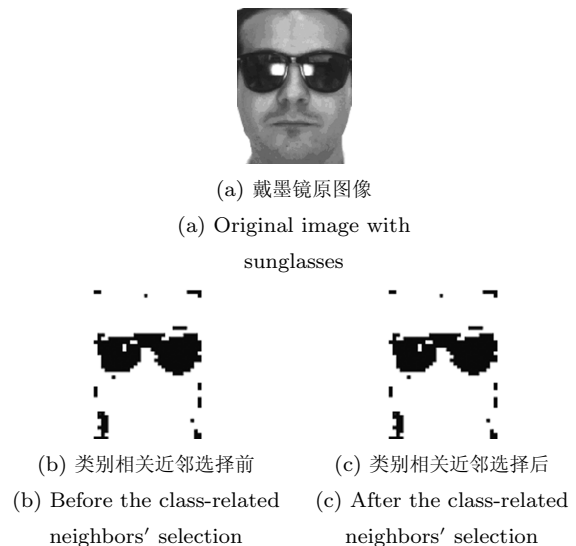


图 9 戴墨镜图像经过类别相关近邻选择前后的权值二值图
Fig. 9 The binary images of weight before and after the class-related neighbors' selection with sunglasses

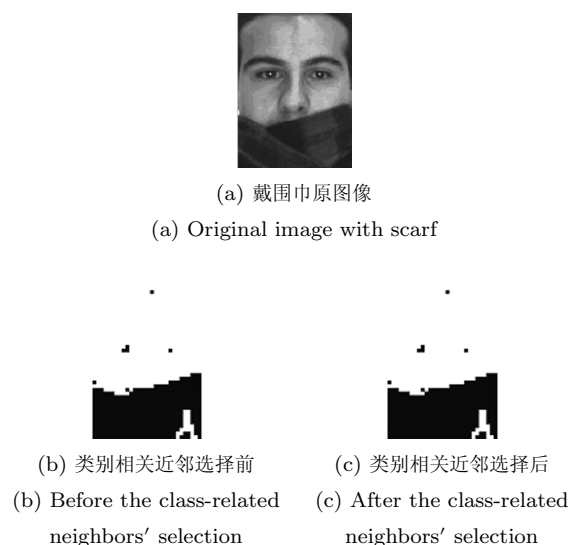


图 10 戴围巾图像经过类别相关近邻选择前后的权值二值图
Fig. 10 The binary images of weight before and after the class-related neighbors' selection with scarf

4 结论

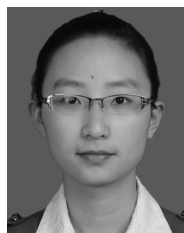
考虑到每个测试样本的不同特性及训练样本选择的类别代表性原则, 提出基于类别相关近邻子空间的快速鲁棒稀疏表示图像识别算法. 基于距离相近准则, 从每个类别中选取自适应数量的局部近邻作为原样本空间的近邻子空间, 构成新的字典, 旨在通过合理减少训练样本的数量, 实现运算时间和识别精度的平衡. 然后基于极大似然估计的稀疏表示识别方法, 将保真度表示为余项的极大似然函数, 并将识别问题转化为一个加权的稀疏优化问题. 在 AR 人脸库、USPS 手写数字库以及 AR 中经过伪装 (戴墨镜、戴围巾) 的人脸子集上的实验证明了该算法的合理性. 构建快速鲁棒的识别系统是图像识别技术研究的关键, 因此如何更加合理的压缩训练样本字典或更合理的构建样本子空间等, 是值得进一步研究的问题.

References

- 1 Wright J, Yang A Y, Ganesh A, Sastry S S, Ma Y. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, **31**(2): 210–227
- 2 Huang J Z, Huang X L, Metaxas D. Simultaneous image transformation and sparse representation recovery. In: Proceedings of the 26th IEEE Conference on Computer Vision and Image Recognition. Anchorage, United States: IEEE, 2008. 1–8
- 3 Wagner A, Wright J, Ganesh A, Zhou Z H, Ma Y. Towards a practical face recognition system: robust registration and illumination by sparse representation. In: Proceedings of the 2009 IEEE Computer Society Conference on Computer Vision and Image Recognition Workshops. Miami, United States: IEEE, 2009. 597–604
- 4 Wright J, Ma Y. Dense error correction via l_1 minimization. *IEEE Transactions on Information Theory*, 2010, **56**(7): 3540–3560
- 5 Yang M, Zhang L, Yang J, Zhang D. Robust sparse coding for face recognition. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Image Recognition. Springs, United States: IEEE, 2011. 625–632
- 6 He R, Hu B G, Zheng W S, Guo Y Q. Two-stage sparse representation for robust recognition on large-scale database. In: Proceedings of the 24th AAAI Conference on Artificial Intelligence and the 22nd Innovative Applications of Artificial Intelligence Conference. Atlanta, United States: AAAI, 2010. 475–480
- 7 Huang J B, Yang M H. Fast sparse representation with prototypes. In: Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Image Recognition. San Francisco, United States: IEEE, 2010. 3618–3625
- 8 Li C G, Guo J, Zhang H G. Local sparse representation based classification. In: Proceedings of the 2010 International Conference on Pattern Recognition. Istanbul, Turkey: ICPR, 2010. 649–652
- 9 Zhang N, Yang J. K nearest neighbor based local sparse representation classifier. In: Proceedings of the 2010 Chinese Conference on Pattern Recognition. Chongqing, China: CCPR, 2010. 400–404
- 10 Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B*, 2011, **73**(3): 273–282
- 11 Zhang J, Jin R, Yang Y M, Hauptmann A G. Modified logistic regression: an approximation to SVM and its applications in large-scale text categorization. In: Proceedings of the 20th International Conference on Machine Learning. Washington, United states: ICML, 2003. 888–895



胡正平 燕山大学信息科学与工程学院电子与通信工程系教授. 主要研究方向为统计学习理论与模式识别, 图像处理. 本文通信作者. E-mail: hzp@ysu.edu.cn (**HU Zheng-Ping** Professor in the Department of Electronics and Communication Engineering, Institute of Information Science and Engineering, Yanshan University. His research interest covers statistical learning theory and pattern recognition, and image processing. Corresponding author of this paper.)



宋淑芬 燕山大学信息科学与工程学院硕士研究生. 2009 年获燕山大学信息科学与工程学院学士学位. 主要研究方向为稀疏表示模式识别. E-mail: songshufen520@163.com (**SONG Shu-Fen** Master student at the Institute of Information Science and Engineering, Yanshan University. She received her bachelor degree from Yanshan University in 2009. Her main research interest is pattern recognition based on sparse representation.)