

核迁移稀疏编码算法在跨域图像分类中的应用

孙登第¹, 孟欠欠^{1,2}, 马云鹏^{1,2}

(1 安徽大学 计算机科学与技术学院, 安徽 合肥 230039;

2 安徽大学 计算智能与信号处理重点实验室, 安徽 合肥 2300601)

摘 要: 针对非线性分布的数据样本在原始特征空间可分性较差的问题, 文中提出一种基于核迁移稀疏编码的跨域图像分类方法, 并应用于图像分类. 首先将图像特征和字典映射到一个高维的再生核希尔伯特空间, 使得线性不可分问题变为线性可分问题. 然后在高维特征空间中对每个样本数据进行表示. 文中算法不仅有效地处理非线性结构数据, 而且考虑了源域和目标域的分布差异以及几何结构信息, 获得更为鲁棒的稀疏表达, 提高跨域图像分类精度.

关键词: 稀疏编码; 核方法; 基学习; 非线性

中图分类号: TP391.4

文献标识码: A

文章编号: 1000-7180(2018)10-0029-07

Kernel Transfer Sparse Coding Algorithms Application for Cross-domain Image Classification

SUN Deng-di¹, MENG Qian-qian^{1,2}, MA Yun-peng^{1,2}

(1 School of Computer Science and Technology, Anhui University, Hefei 230601, China;

2 Key Laboratory of ICSP, Ministry of Education, Anhui University, Hefei 230039, China)

Abstract: To deal with the problem that the data samples with non-linear distribution are poorly separable in the original feature space, in this paper, a kernel transfer sparse coding algorithms is proposed for cross-domain image classification. Firstly, the image features and the dictionary are mapped to a high-dimensional reproducing-kernel hilbert space, which makes the problem of linear inseparable problem become a linear separable problem. Then, the samples are individually represented in high-dimensional feature space. The proposed algorithm not only effectively handles the the nonlinear structure data, but also considers the distribution differences and geometric structure information between source and target domains, which gain more robust sparse representation and improve cross-domain image classification accuracy.

Key words: sparse coding; kernel method; basis learning; non-linear

1 引言

在图像分类中, 稀疏编码方法是一种新型的特征表示方法, 它可以将图像分解成一组超完备图像字典集与表示系数的线性组合, 以获取样本数据的高维稀疏语义特征, 然后利用这些特征进行迁移是实现跨域的关键. 这使得基于特征选择的稀疏表示方法在当前图像研究领域得到了愈多的关注. 为了提高图像处理的效果, 研究者们在其基础上综合考

虑图像的高级语义特征等进行不断的改进^[1], 并取得一定的成果. 谭等人^[2]将回归模型与稀疏编码相结合, 以迭代搜索人脸特征点位置. 迪可新等人^[3]将稀疏编码扩展到同时对图像的行列进行压缩, 可以快速有效地检测空间目标与成像. Zheng 等人^[4]提出图正则化稀疏编码模型(GSC), 在稀疏编码的基础上引入了图像拉普拉斯正则化项, 保留数据的原始几何结构. Long 等人^[5]考虑数据的重合度, 提出了迁移稀疏编码(transfer sparse coding, TSC), 提

高了稀疏表示的质量. 这些改进算法在图像处理与分类领域中均取得了优异的成绩. 然而, 却忽略了数据的特征分布是否线性可分的问题. 大量的研究表明, 当样本数据的特征差异较小时, 数据点之间的分布比较混杂, 呈非线性分布, 利用传统的线性算法很难进行分离, 且效果并不理想.

核方法是一类有效处理非线性问题的方法, 它的基本思想是通过非线性映射(核映射), 将低维数据投影到高维的希尔伯特空间, 使映射后的数据具有能够进行线性分离, 进而利用传统的分类方法对图像数据进行处理. 该方法不仅有效的解决了样本数据线性不可分的问题, 而且在一定程度上降低计算的复杂度, 提高算法效率, 所以将传统线性算法与核方法相结合能够提高非线性分布数据的分类精度^[6].

2 相关技术

假设给定 $\mathbf{X}$ 是待编码的图像矩阵 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n] \in \mathbb{R}^{p \times n}$, 包括源域 $\mathbf{X}^l$ 和目标域 $\mathbf{X}^u$ 两部分, $\mathbf{X} = [\mathbf{X}^l, \mathbf{X}^u]$. 其中, 源域 $\mathbf{X}^l = [\mathbf{x}_1^l, \dots, \mathbf{x}_{n_l}^l] \in \mathbb{R}^{p \times n_l}$, $\mathbf{X}^l = [\mathbf{x}_1^l, \dots, \mathbf{x}_{n_l}^l] \in \mathbb{R}^{p \times n_l}$, 含有 n_l 幅有标签的图像, 而目标域 $\mathbf{X}^u = [\mathbf{x}_1^u, \dots, \mathbf{x}_{n_u}^u] \in \mathbb{R}^{p \times n_u}$, $\mathbf{X}^u = [\mathbf{x}_1^u, \dots, \mathbf{x}_{n_u}^u] \in \mathbb{R}^{p \times n_u}$, 含有 n_u 幅图像. 稀疏编码的目的是需要寻找一个过完备字典 $\mathbf{D} \in \mathbb{R}^{p \times h}$ 和一个系数矩阵 $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_n] \in \mathbb{R}^{h \times n}$ 的乘积近似表示矩阵 $\mathbf{X}$. 为了使编码系数具有稀疏性以及目标函数易求解, 稀疏编码采用系数的 l_1 范数作为正则化项来约束模型. 因此, 稀疏编码的模型可表示为:

$$\min_{\mathbf{U}, \mathbf{A}} \|\mathbf{X} - \mathbf{U}\mathbf{A}\|_F^2 + \lambda \sum_{i=1}^n |\mathbf{a}_i| \quad (1)$$

式中, $\mathbf{X} - \mathbf{U}\mathbf{A}$ 为残差矩阵, 表示重构误差; $|\mathbf{a}_i|$ 是系数矩阵 $\mathbf{A}$ 中每列的 l_1 范数, 用以约束系数向量 $\mathbf{a}_i$ 的稀疏性; λ 为可调节稀疏参数, 用以对系数矩阵惩罚约束.

同时, 为了更好的提取图像的本质特征, 进行鲁棒性的稀疏表示, 图正则化稀疏编码模型(GSC)在 SC 的基础上引入图拉普拉斯矩阵, 将其正则化, 以挖掘了数据集的几何局部特性. 其构造过程如下: 对全部输入数据 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n] \in \mathbb{R}^{p \times n}$, 构造具有 n 个顶点的 k 近邻图 G , 每个顶点表示一幅图像. 那么图 G 的权值矩阵 $\mathbf{W}$, 有如下定义: 如果数据点 x_i 属于 x_j 的 k 近邻, 那么 $\mathbf{W}_{ij} = 1$, 否则 $\mathbf{W}_{ij} = 0$; 它的度矩阵 $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$ 定义为: $d_i = \sum_j \mathbf{W}_{ij}$. 从而计算得到拉普拉斯矩阵: $\mathbf{L} = \mathbf{D} - \mathbf{W}$, 具体推导公

式为:

$$\min \frac{1}{2} \sum_{ij} \|\mathbf{a}_i - \mathbf{a}_j\|^2 \mathbf{W}_{ij} = \text{tr}(\mathbf{A}\mathbf{L}\mathbf{A}^T) \quad (2)$$

数学模型如公式(3)所示.

$$\begin{aligned} \min_{\mathbf{D}, \mathbf{A}} \|\mathbf{X} - \mathbf{D}\mathbf{A}\|_F^2 + \text{tr}(\mathbf{A}(\nu \mathbf{L})\mathbf{A}^T) + \lambda \sum_{i=1}^n |\mathbf{a}_i| \\ \text{s. t } \|\mathbf{d}_i\|^2 \leq t, \forall i = 1, \dots, n \end{aligned} \quad (3)$$

式中, ν 是图拉普拉斯正则化参数, 主要用来约束局部特征的量化误差. 通过将图拉普拉斯矩阵正则化, 保留了样本数据的几何内蕴结构, 同时降低了量化误差, 使具有相似特征数据的稀疏编码保持最大程度的相似性.

由于跨域数据通常来自不同的分布, 其数据本身相对应子空间内的特征分布也可能并不相同, 甚至分布差异较大或完全相反. 此时源域子空间中的特征不一定能促进目标领域的学习, 影响分类效果. 针对这一问题, 迁移稀疏编码模型(TSC)在 GSC 的基础上进一步最小化源域与目标域的数据距离^[5], 同时最大程度保留它们各自的内部特性, 从而使源领域与目标领域的特征分布尽可能相似, 能够准确描述域间的共同特性, 而不仅仅是单个数据域的特性表达. 该算法使用一个经典的公式叫做最大均值差异表示两域的数据距离. 距离的计算公式如下:

$$\begin{aligned} \min \left\| \frac{1}{n_l} \sum_{i=1}^{n_l} \mathbf{a}_i - \frac{1}{n_u} \sum_{j=n_l+1}^{n_u} \mathbf{a}_j \right\|^2 \\ = \sum_{i,j=1}^n \mathbf{a}_i^T \mathbf{a}_j \mathbf{M}_{ij} \\ = \text{Tr}(\mathbf{A}\mathbf{M}\mathbf{A}^T) \end{aligned} \quad (4)$$

式中, $\mathbf{M} \in \mathbb{R}^{n \times n}$ 为最大均值差矩阵, 推导公式为:

$$\mathbf{M}_{ij} = \begin{cases} \frac{1}{n_l^2}, & x_i, x_j \in \mathbf{X}^l \\ \frac{1}{n_u^2}, & x_i, x_j \in \mathbf{X}^u \\ -\frac{1}{n_l n_u}, & \text{otherwise} \end{cases} \quad (5)$$

将其融入 GSC 中, 构成迁移稀疏编码模型, 数学模型如公式(6)所示.

$$\begin{aligned} \min_{\mathbf{D}, \mathbf{A}} \|\mathbf{X} - \mathbf{D}\mathbf{A}\|_F^2 + \text{tr}(\mathbf{A}(\mu \mathbf{M} + \nu \mathbf{L})\mathbf{A}^T) + \lambda \sum_{i=1}^n |\mathbf{a}_i| \\ \|\mathbf{d}_i\|^2 \leq t, \forall i = 1, \dots, n \end{aligned} \quad (6)$$

在 TSC 模型中, μ 为均值差参数, 通过控制最大均值差矩阵 $\mathbf{M}$ 以度量不同数据集分布的差异, 准确描述域间共同特性; 同时利用 $\mathbf{L}$ 矩阵保留了源域和目标域的几何结构, 以保证局部特征的一致性. 通过

将这两个正则化项相结合,能够获得更精确的高级语义特征.

3 核迁移稀疏编码模型

当数据呈高度非线性分布时,样本在原始特征空间可分性较差,而迁移稀疏编码无法高效处理非线性数据分布的问题^[7].因此,利用核函数将两个领域的图像数据共同映射到能对数据线性分离的空间中.在映射后的空间中,最小化域间的数据距离,同时保留数据集的几何内蕴结构,然后对样本进行编码和分类.假定对一组服从非线性分布的输入数据 $X = [x_1, \dots, x_n] \in \mathbb{R}^{p \times n}$ 实施以下核映射:

$$\varphi: \mathbb{R}^n \rightarrow F, x \in X \rightarrow \varphi(x) \in F \quad (7)$$

式中, φ 为核映射; F 为高维空间.核方法主要通过核映射 φ 将原始特征空间 X 映射到核空间 F ,解决了样本数据线性不能分离的问题.其本质是使用核函数替换数据点之间复杂的内积变换:

$$(x_i, x_j) \rightarrow k(x_i, x_j) = \langle \varphi(x_i) \varphi(x_j) \rangle \quad (8)$$

式(8)中, $\langle \varphi(x_i), \varphi(x_j) \rangle$ 为数据点 x_i 和 x_j 的内积运算, $k(x_i, x_j)$ 为核函数.核函数的引入巧妙的将高维度中向量的内积计算转化为低维度的核函数计算,避免了直接计算映射 φ 的问题,极大的降低了算法的复杂度.不同的核函数具有各自的优点和缺陷,对算法也会产生不同的影响,目前应用比较广泛的是以下四种核函数,基本公式如式(9)所示.

$$K = \begin{cases} E_{i,j} = x_i \cdot x_j \\ D_{i,j} = [x_i \cdot x_j + a]^d \\ G_{i,j} = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \\ C_{i,j} = \tanh(b(x_i \cdot x_j) - c) \end{cases} \quad (9)$$

式中, $E_{i,j}$ 为线性核函数,主要应用于数据线性分布的情况下,并能取得很好的效果; $D_{i,j}$ 和 $C_{i,j}$ 分别为多项式核和 Sigmoid 核函数,它们均含有两个参数,常导致计算复杂度较高; $G_{i,j}$ 为高斯核函数,只含有一个参数,从计算复杂度方面考虑,高斯核是一个较好的选择^[8].此外,高斯核是一个对应于无限空间的核函数,可以将低维空间投影到无穷维空间中,能够很好的处理数据非线性分布的情况.

对于输入数据集 $\{x_1, \dots, x_n\}$, 高斯核函数计算的是每一个数据点 x_i 和所有数据点之间的距离,若 $x_i x_i$ 和 x_j 距离很近,那么 $\|x_i - x_j\|^2 \approx 0$, 核函数的值为 1; 反之若 $x_i x_i$ 和 x_j 距离很远,核函数的值为 0, 在同样的距离下,核函数变化的敏感度可以通过

δ 灵活控制.因此本文选择高斯核函数作为映射函数,在本文中定义为:

$$k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (10)$$

将其引入到迁移稀疏编码过程中,更容易得到鲁棒稀疏编码,提高分类精度.这也是本文所提出的核迁移稀疏编码模型,该模型的优化函数如下:

$$\min_{\varphi(D), A} \|\varphi(X) - \varphi(D)A\|_F^2 + \text{Tr}(A(\mu M + \nu L)A^T) + \lambda \sum_{i=1}^n |a_i| \quad (11)$$

式中, $\varphi(X)$ 表示核映射后的数据集; $\varphi(D)$ 表示核映射后的字典,也称为核字典.该模型利用核方法建立点到点的映射,有效解决了数据在原始特征空间不能线性分离的问题.

4 模型求解

由于目标函数(11)存在两个自变量 $\varphi(D)$ 与 A , 其求解过程可以分为字典学习与稀疏表示学习两个过程.因此,我们将目标函数分解为两个子函数,并分别交替迭代寻优,最终使目标函数收敛.

4.1 稀疏表示学习

首先初始化字典,学习稀疏编码 A , 进入第一轮迭代,此时目标函数(11)可简化为如下子函数:

$$\min_A \|\varphi(X) - \varphi(D)A\|_F^2 + \text{Tr}(A(\mu M + \nu L)A^T) + \lambda \sum_{i=1}^n |a_i| \quad (12)$$

为叙述方便,我们将式(11)写成向量形式并求解,过程如下:

$$\begin{aligned} J(a_i) &= \|\varphi(x_i) - \varphi(D)a_i\|_F^2 + \text{Tr}(a_i(\mu M + \nu L)a_i^T) + \lambda \sum_{i=1}^n |a_i| \\ &= (\varphi(x_i) - \varphi(D)a_i)^T (\varphi(x_i) - \varphi(D)a_i) \\ &\quad + \text{Tr}(a_i(\mu M + \nu L)a_i^T) + \lambda \sum_{i=1}^n |a_i| \\ &= k(x_i, x_i) + a_i^T \mathbf{K}_{PP} a_i - 2a_i^T \mathbf{K}_P(x_i) \\ &\quad + \text{Tr}(a_i(\mu M + \nu L)a_i^T) + \lambda \sum_{i=1}^n |a_i| \quad (13) \end{aligned}$$

根据式中 $\mathbf{K}_{PP}$ 表示 $P \times P$ 大小的核矩阵; $\mathbf{K}_P(x)$ 表示 $p \times 1$ 的矢量,根据核函数的定义,具有如下关系:

$$\begin{cases} \{\mathbf{K}_{PP}\}_{ij} = \langle \varphi(p_i) \varphi(p_j) \rangle = k(p_i, p_j) \\ \mathbf{K}(x, x) = 1 \\ \{\mathbf{K}_P(x)\}_i = \langle \varphi(p_i) \varphi(x) \rangle = k(p_i, x) \end{cases} \quad (14)$$

$$\text{令 } L(a_i) = 1 + a_i^T \mathbf{K}_{PP} a_i - 2a_i^T \mathbf{K}_P(x_i) +$$

$\text{Tr}(a_i(\mu M + \nu L)a_i^T)$, 优化函数为:

$$J(a_i) = L(a_i) + \lambda \sum_{i=1}^n |a_i| \quad (15)$$

本文采用文献[5]提出的特征符号搜索方法(feature-sign search algorithm)对式(15)求解, 获取编码矩阵 $\mathbf{A}$.

4.2 核字典学习

固定获得的编码 $\mathbf{A}$, 目标函数(11)可以简化为如下形式:

$$\min_{\varphi(D)} \|\varphi(X) - \varphi(D)\mathbf{A}\|_F^2 \quad (16)$$

为了优化字典, 本文采用文献[8]提出的方法, 假设随机选择 N 个特征进行字典训练, 则公式(16)变为:

$$\begin{aligned} f(\varphi(D)) &= \frac{1}{N} \sum_{i=1}^n \|\varphi(x_i) - \varphi(D)a_i\|^2 \\ &= \frac{1}{N} \sum_{i=1}^n \left[1 + \sum_{s=1}^k \sum_{t=1}^k a_{is}a_{it}k(\varphi(D)_s, \varphi(D)_t) \right] \\ &\quad - \frac{1}{N} \sum_{i=1}^n \left[2 \sum_{s=1}^k a_{is}k(\varphi(D)_s, x_i) \right] \end{aligned} \quad (17)$$

式中, $\varphi(D)_s, \varphi(D)_t$ 分别表示字典中第 s, t 个基向量; h 为字典长度. 为了获得字典的每一列, 设当前要优化的字典第 m 列为 $\varphi(D)_m, m \in (1, 2, \dots, h)$. 为了求解方便, 对目标函数 $f(\varphi(D))$ 求导为:

$$\begin{aligned} \frac{\partial f}{\partial \varphi(D)_m} &= \\ &= \frac{4\gamma}{N} \sum_{i=1}^n \left[\sum_{t=1}^h a_{im}a_{it}k(\varphi(D)_m, \varphi(D)_t) (\varphi(D)_m - \varphi(D)_t) \right] \\ &\quad - \frac{4\gamma}{N} \sum_{i=1}^n \left[\sum_{t=1}^h a_{im}k(\varphi(D)_m, x_i) (\varphi(D)_m - x_i) \right] \end{aligned} \quad (18)$$

为了方便求解, 本文使用第 $(n-1)$ 次结果计算核函数项 $k(\varphi(D), x)$. 目标函数变为:

$$\begin{aligned} \frac{\partial f}{\partial (\varphi(D)_m)_n} &\cong \\ &= \frac{4\gamma}{N} \sum_{i=1}^n \left[\sum_{t=1}^h a_{im}a_{it}k((\varphi(D)_m)_{n-1}, \varphi(D)_t) \right. \\ &\quad \left. ((\varphi(D)_m)_n - \varphi(D)_t) \right] \\ &\quad + \frac{4\gamma}{N} \sum_{i=1}^n [a_{im}k((\varphi(D)_m)_{n-1}, x_i) ((\varphi(D)_m)_n - x_i)] \end{aligned} \quad (19)$$

式中, $(\varphi(D)_m)_{n-1}$ 为第 $(n-1)$ 次迭代的基向量, $(\varphi(D)_m)_n$ 为优化后的核字典. 令 $\partial f / \partial (\varphi(D)_m)_n = 0$, 求得核字典为:

$$(\varphi(D)_m)_n = \sum_{i=1}^n \sum_{t=1}^h$$

$$\frac{a_{it}\varphi(D)_t - k((\varphi(D)_m)_{n-1}, x_i)x_i}{a_{it}k((\varphi(D)_m)_{n-1}, \varphi(D)_t) - k((\varphi(D)_m)_{n-1}, x_i)}$$

4.3 KTSC 算法实现

本文所提出的核迁移稀疏编码(KTSC)算法是一个不断交替迭代寻优的过程, 简单介绍算法实现步骤.

step1. 利用高斯核 K 对全部输入数据 X 进行核处理, 将它们映射到高维特征空间;

step2. 初始化 $\varphi(D)$. 利用核化后的训练样本和测试样本对核字典 $\varphi(D)$ 进行初始化;

step3. 更新稀疏编码 $\mathbf{A}$. 固定 $\varphi(D)$, 根据 feature-sign search 算法优化式(15)求出稀疏编码 $\mathbf{A}$;

step4. 更新核字典 $\varphi(D)$. 固定编码 $\mathbf{A}$, 根据式(19)求出优化后的核字典 $\varphi(D)$;

step5. 输出 $\varphi(D)$ 和 $\mathbf{A}$. 重复步骤 3 和步骤 4, 直至达到收敛标准 ϵ , 此时得到优化后的核字典 $\varphi(D)$ 和稀疏编码 $\mathbf{A}$;

step6. 分类. 根据得到的字典 $\varphi(D)$ 和编码 $\mathbf{A}$ 对测试数据进行处理和分类.

5 实验与讨论

为充分验证本文所提算法的合理性、有效性和优越性, 本文分别在手写数字数据集(USPS、MNIST)和自然图像数据集(Office、Caltech-256)上进行分类实验. 同时本节还将 KTSC 与逻辑回归(LR)、主成份分析(PCA)、稀疏编码(SC)、基于图正则化的稀疏编码(GSC)、基于分布差异正则化的稀疏编码(MSC)和迁移稀疏编码(TSC)算法进行对比分析.

5.1 评价标准

准确率(accuracy, AC)是针对分类结果的一种求解方法, 是分类算法常采用的评价标准, 表示预测的标签中正确的标签数所占的比例. 因此实验采用测试集上的分类准确率作为算法性能的评价标准, 其计算如公式(20)所示.

$$\text{Accuracy} = \frac{|x: x \in D_s \wedge \hat{y}(x) = y(x)|}{|x: x \in D_s|} \quad (20)$$

式中, D_s 为测试数据; $y(X)$ 是图像 x 的真正标签; $\hat{y}(x)$ 是根据模型预测到的标签.

5.2 数据集

手写数字数据集: USPS 是美国手写数字数据集, 包含 7 291 幅训练数据和 2 007 幅测试数据, 本文从中随机抽取 1 800 幅图像做为实验数据;

MNIST 包含 60 000 个训练数据和 10 000 个测试数据,本文从中随机抽取 2 000 幅图像作为实验数据. 本文将所有实验数据尺寸缩放成 10×10 大小,每幅图像表示成 100 维的向量. 实验分别对两个数据进行了不同程度的随机遮挡,主要包括无遮挡的、随机 5% 遮挡后的以及随机 10% 遮挡后的实验数据集. 共组成 U-M, M-U, 5U-5M, 5M-5U, 10U-10M, 10M-10U 等六个分类任务,分别用 U, M 表示 USPS 和 MNIST 数据集, U 和 M 前面数字表示数据随机遮挡的程度. 每个分类任务包含两部分,训练集和测试集,比如 5M-5U 表示随机遮挡 5% 的 MNIST 作为训练集,随机遮挡 5% USPS 数据集作为测试集^[9]. 实验结果如表 1 所示.

表 1 字符数据集的分类准确率 %						
数据集	M-U	5M-5U	10M-10U	U-M	5U-5M	10U-10M
LR	27.44	10.11	6.78	27.8	28.35	25.5
PCA	37.78	17.89	10.44	31.75	32.25	28.45
SC	43.66	37.15	31.91	38.69	37.23	29.45
GSC	59.29	56.15	49.08	44.56	41.83	35.43
MSC	49.52	45.29	41.63	44.47	41.72	35.28
TSC	54.14	53.19	48.74	52.32	50.77	41.73
KTSC	57.50	56.61	50.05	59.45	55.70	45.05

自然图像数据集:Office 是用于领域自适应的标准数据集,含有三个子数据集 Amazon、Dslr 和 Webcam,共包含 31 类 4652 幅图像. Caltech-256 是计算机视觉标准数据集,共包含 256 类 30607 幅自然图像. 本文采用 long 等人提出的 Office+Caltech 数据集^[10],共包括 Amazon、Caltech-256、Dslr、Webcam 等四个实验数据集,分别用 A、C、D、W 表示,任选其中两个试验数据集分别作为训练集和测试集,共组成 12 个分类任务. 实验结果如表 2 和表 3 所示.

表 2 自然图像数据集的分类准确率 a %						
数据集	A-C	A-D	A-W	C-A	C-D	C-W
LR	15.85	24.84	22.71	15.03	10.19	9.15
PCA	17.28	24.71	26.75	15.87	13.38	7.80
SC	15.91	28.22	22.92	16.98	13.89	8.07
GSC	14.79	22.93	24.71	18.05	17.52	13.56
MSC	18.43	27.64	15.83	18.68	23.06	18.98
TSC	18.99	32.10	22.14	19.34	23.31	13.86
KTSC	24.98	36.49	29.32	20.11	28.55	21.69

表 3 自然图像数据集的分类准确率 b %						
数据集	D-A	D-C	D-W	W-A	W-C	W-D
LR	17.11	11.04	55.25	13.88	8.99	59.23
PCA	15.45	12.11	52.20	15.76	9.53	60.51
SC	18.42	12.00	62.27	14.44	10.40	60.19
GSC	14.24	15.33	56.20	14.71	11.72	62.29
MSC	16.45	12.12	41.12	16.33	11.62	54.01
TSC	14.39	12.00	49.93	18.58	13.07	57.13
KTSC	23.48	15.85	52.61	21.85	13.17	61.24

5.3 实验分析

从表 1 可以看出,所有方法的分类精度都随着遮挡程度加大而下降. 其中 MSC 和 GSC 是 TSC 算法的退化情况,因为 MSC 仅保留了其中的一个分布差异的正则化项,GSC 仅保留了图像拉普拉斯正则化项. 实验结果显示,GSC 分类精度高于 MSC 分类精度,表明在字符数据集中,数据集几何特性的保留对于分类效果有更为重要的作用. 同时看到 KTSC 算法在六个分类任务上的分类精度均高于其它六种方法,与 TSC 方法相比依次提高了 3.36%、3.42%和 1.31%,以及 7.13%、4.93%和 3.32%. 表明了融入核方法的迁移稀疏编码模型能够解决数据非线性分布问题,体现了 KTSC 算法的鲁棒性和有效性.

表 2 中,与字符数据集不同,在自然图像数据集中,图像的分类准确率总体不高. 其均分类精度分别为 21.94%、23.32%、24.18%、24.32%、21.91%、24.12%和 29.11%. 其中 LR 使用最原始的图像像素特征,计算代价不大,但精度不高,PCA 通过数据降维提取有用的特征,SC 提取了高级语义特征,与 LR 相比,均分类精度得到了提高,表明通过提取有用特征对图像跨域分类起着重要作用. 此外,KTSC 算法的均准确率相比其它方法分别提高了 7.17%、5.79%、4.93%、4.79%、7.2%和 4.99%,也体现了核迁移稀疏编码算法的有效性.

5.4 参数分析

为了检验 KTSC 中各参数对分类精度影响,本文分别对核带宽参数 γ ,最大均值差参数 μ 、图正则化参数 ν ,局部稀疏性参数 λ 进行实验分析. 实验主要在 M-U、A-C 以及 D-A 三个分类任务进行,实验过程中分别固定其它参数值,调节另一个参数值,其实验结果如图 1~5 所示.

首先对于核带宽参数 γ ,实验结果如图 1 所示. 可以看出, γ 取 $[0.1, 1]$ 图像分类精度比较大,这是

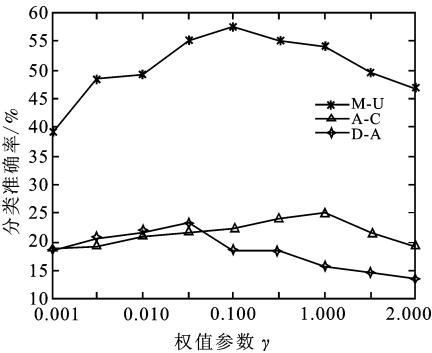


图 1 核带宽参数 γ 对分类精度的影响

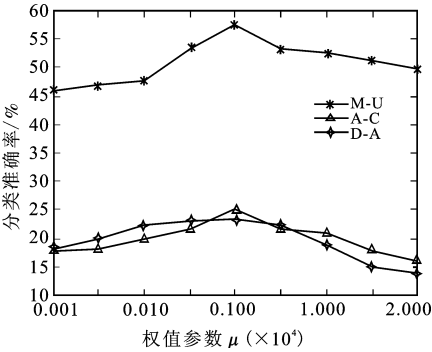


图 2 均值差参数 μ 对分类精度的影响

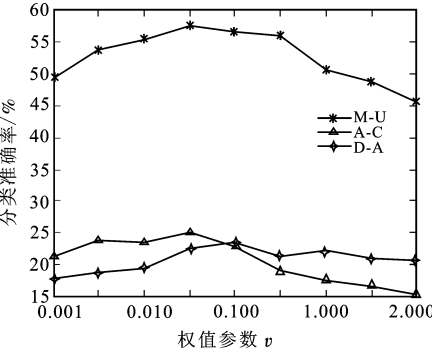


图 3 图正则化参数 ν 对分类精度的影响

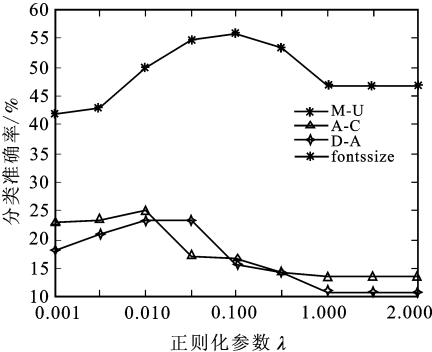


图 4 稀疏正则化参数 λ 对分类精度的影响

由于将数据集和字典映射到高维的希尔伯特空间,更好的进行了图像的特征表示. 而当 γ 过大或过小

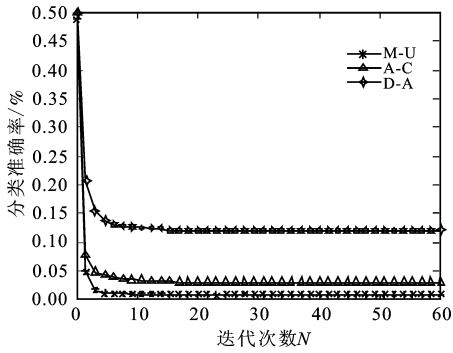


图 5 迭代次数 N 对目标函数收敛的影响
时,会忽略数据集的其它几何特性,导致分类精度下降.

对于均值差参数 μ ,它的大小决定了域间分布的差异性. 具体来说,在图 2 中,可以看到,分布差异的正则化参数 μ 取值为 10^4 左右时 10^4 时,分类性能在三组跨域图像集上都达到最优. 它的大小主要控制着域间的数据距离,然而当 $\mu \rightarrow 0$ 时,即几乎忽视数据集间的分布差异,当 $\mu \rightarrow \infty$ 时,即几乎忽略数据集的几何信息和稀疏特性,都会导致分类精度下降. 在图 3 中,图正则化参数 ν 用来表示数据集的几何内蕴结构在所提取特征中所占的比重. 从图中可以看出,当 $\nu \rightarrow 0$ 时,即几乎忽略数据集的几何特性,和 $\nu \rightarrow \infty$ 即几乎忽略域间差异和稀疏结构,均对分类精度产生较大的影响.

在图 4 中,参数 λ 控制编码的局部稀疏性,当 $\lambda > 1$ 时,分类精度降低很快,主要是因为当稀疏正则化参数超过一定程度时,随着参数的增大,对编码稀疏度惩罚越大,导致非零分量减少,降低分类精度.

为说明本文所提方法的收敛性,图 5 展示了迭代次数对目标函数的影响. 可以看到,KTSC 在 N 接近 10 时开始收敛,表明目标函数的收敛速度较快,同时也验证了 KTSC 算法的合理性与可行性.

6 结束语

针对原始特征空间线性不可分的问题,提出核迁移稀疏编码模型,与迁移稀疏编码不同的是,将样本数据和字典映射到一个高维的线性可分离空间中,使数据能够线性可分;同时考虑了数据集的分布差异与几何特征,提取图像的高级语义信息,获得了更鲁棒性的稀疏编码. 在问题的求解上,编码的学习仍采用特征符号搜索方法,对于核字典的学习,本文通过迭代算法进行求解. KTSC 模型将迁移稀疏编码与核方法相结合,不仅解决了数据线性不可分的问题,而且降低了计算的复杂度. 实验结果也证实了

KTSC方法可以比其它相关方法取得更好的分类精度.

参考文献:

[1] 尚丽. 稀疏编码算法及其应用研究[D]. 合肥: 中国科学技术大学, 2006.

[2] 谭暑秋, 陈东义, 郭承刚, 等. 一种基于稀疏编码的人脸特征点检测方法[J]. 微电子学与计算机, 2017, 35(9):7-10.

[3] 迪可新, 魏平, 柯钧, 等. 基于压缩感知的空间稀疏目标成像方法研究[J]. 微电子学与计算机, 2015, 33(10):98-104.

[4] Zheng M. Graph regularized sparse coding for image representation[J]. IEEE Transactions on Image Processing, 2011, 20(5): 1327-1336.

[5] Long M. Transfer sparse coding for robust image representation[J]. Computer Vision & Pattern Recognition, 2013, 9(4):407-414.

[6] 王铁建, 吴飞, 荆晓远. 核字典学习在软件缺陷预测中的应用[J]. 小型微型计算机系统, 2017, 38(7):

1501-1505.

[7] 姜海燕, 刘昊天, 舒欣, 等. 基于最大均值差异的多标记迁移学习算法[J]. 信息与控制, 2016, 45(4): 463-470.

[8] 刘涛. 基于核拉普拉斯稀疏编码的图像分类[D]. 大连: 大连理工大学, 2013.

[9] 赵鹏, 王维. 基于迁移鲁棒稀疏编码的图像表示方法[J]. 计算机学报, 2017, 39(10):2421-2432.

[10] Long M, Wang J, Ding G, et al. Transfer Feature Learning with Joint Distribution Adaptation [C]// IEEE International Conference on Computer Vision. IEEE, 2014:2200-2207

作者简介:

孙登第 男,(1983-),博士,副教授. 研究方向为机器学习与模式识别. E-mail:mq_ahu@126.com.

孟欠欠 女,(1994-),硕士研究生. 研究方向为图像处理与分类.

马云鹏 男,(1988-),硕士研究生. 研究方向为机器学习与模式识别.