

基于局部抑制线性编码的图像快速识别方法

陈光喜 龚震霆 温佩芝 任夏荔

(桂林电子科技大学计算机科学与工程学院广西高校图像图形智能处理重点实验室 桂林 541004)

摘要 传统的图像识别方法如 ScSPM、LLC 都是在 SIFT 的基础上提取特征,忽略了人工特征的局限性,且单张图像识别耗时略长。考虑到这些不足,提出了一种基于局部抑制线性编码的图像快速识别方法。该方法首先直接利用局部抑制线性编码提取图像局部特征描述子;然后用线性空间金字塔匹配(LSPM)对特征描述子进行计算;最后将计算结果输入到线性支持向量机(LSVM)中进行训练和测试。在3个常用的图像数据集上的实验结果表明,该方法在类别不多的情况下具有很好的识别准确率,同时大大减少了单张图像识别耗时,从而验证了该方法在图像识别上的有效性。

关键词 局部抑制线性编码,线性空间金字塔匹配,线性支持向量机,单张图像识别耗时

中图分类号 TP391.41 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.5.059

Fast Image Recognition Method Based on Locality-constrained Linear Coding

CHEN Guang-xi GONG Zhen-ting WEN Pei-zhi REN Xia-li

(Guangxi Colleges and Universities Key Laboratory of Intelligent Processing of Computer Images and Graphics, School of Computer Science and Engineering, Guilin University of Electronic Technology, Guilin 541004, China)

Abstract The traditional image recognition methods, such as ScSPM and LLC, are based on the SIFT feature, ignoring the limitations of artificial features, and the single image recognition time-consuming takes slightly longer. Considering these deficiencies, this paper proposed a fast recognition method for image based on locality-constrained linear coding. The method first directly uses locality-constrained linear coding to extract local features' descriptors of image, then uses the linear spatial pyramid matching(LSPM) to calculate feature descriptors, and inputs the results into the linear support vector machine(LSVM) for training and testing. The experimental results for three usual image data sets show that the method has good recognition accuracy, and at the same time greatly reduces the signal image recognition time-consuming, which verifies the effectiveness of this method in the image recognition.

Keywords Locality-constrained linear coding, Linear spatial pyramid match, Linear support vector machine, Single image recognition time-consuming

1 引言

图像识别是依据图像的某些特征将其分类到不同的类别中^[1],是人工智能与模式识别的一个重要领域,如今在生物医学、航天航空、安全监控、大众生活方面都有了广泛的发展与应用。图像识别通常分为预处理、图像分割、特征提取、分类器训练与预测等步骤,而其中任何一步都对识别结果影响重大^[2]。然而由于自然图像类别繁多,且其拓扑结构复杂多变,采用图像分割算法很难精确完整地分割出单个目标,因此计算机无法准确判断事先定义的人工特征,从而导致后面分类器训练与预测的不精确。

利用传统的图像分割算法对图像进行完整的分割,然后提取人工特征,是普遍使用的一种图像识别方法。能否正确地分割出目标决定着特征提取的准确性,如 SIFT(Scale-invariant feature transform)特征的提取及基于 SIFT 的变种方

法^[3-6]。此外,另一种方法主要针对种类繁多、内容拓扑结构复杂多变的图像,该方法从图像中随机地选取大量的 patch,然后采用某种学习模型来提取图像的局部特征,如近几年在图像识别上获得很好效果的稀疏编码(Sparse Coding, SC)算法^[7]。SC 模拟哺乳动物视觉系统主视皮层 V1 区简单细胞感受野,是一个无监督的学习过程,可以对已提取特征进行稀疏表示,也可以直接提取图像的局部特征。

2 相关工作

稀疏编码理论被引入到图像处理,极大地促进了图像分类识别准确率的提高。2004年, Lee 等人^[9]提出了一种基于递归解决凸优化问题的高效稀疏编码算法;特征符号搜索法解决 l_1 范数正则化的最小二乘问题和拉格朗日对偶法解决 l_2 范数约束化最小二乘问题,实验结果表明该算法显著提高了过完备基训练速度,从而可以学习更稀疏的 codes。2005

到稿日期:2015-05-23 返修日期:2015-08-10 本文受广西高校科研资助项目(2013YB086),广西自然科学基金(2013GXNSFAA019330),国家自然科学基金(61363030),广西高校图像图形智能处理重点实验室开放基金资助。

陈光喜(1971-),男,博士,教授,主要研究方向为可信计算、图像处理;龚震霆(1991-),男,硕士生,主要研究方向为模式识别、图像处理;温佩芝(1963-),女,博士,教授,主要研究方向为模式识别、三维重建。

年 Lazebnik^[10] 等人提出超越“词袋”(Bags of Features, BOF)^[11] 的基于近似全局几何相应的空间金字塔匹配(Spatial Pyramid Matching, SPM)。BOF 把图像表示为局部特征的一个无序的收集,忽略了特征的空间布局信息,无法对图像进行精确的识别;而 SPM 是 BOF 的改进,是把图像分割成越来越细小的子区域,然后计算每一个子区域内局部特征的直方图。目前,稀疏编码在很多方面获得了较好的应用,如人脸识别^[20]、音频分类^[21]、医学图像识别^[22] 等。

近年来,SVMs 使用 SPM 核在图像识别分类上已经取得了很大的成功,但是这些非线性 SVMs 在训练上的复杂度是 $O(n^2 \sim n^3)$,在测试上复杂度是 $O(n)$,其中 n 是训练规模。由此可知,当扩展此方法去处理超过几千张的图像时的复杂度是不容小觑的。2009 年, Yang 等人^[12] 提出了线性空间金字塔匹配(Linear Spatial Pyramid Matching, LSPM), LSPM 是 SPM 方法的一种扩展。Yang 等人用稀疏编码替代矢量量化(Vector Quantization, VQ),接着用 max pooling 方法进行多尺度空间池化,并提出基于 SIFT 稀疏编码的线性 SPM 核,实验表明此方法显著地把 SVMs 在训练上的复杂度降低到 $O(n)$,在测试上的复杂度降低到某一常数;大量实验表明,基于 SIFT 稀疏编码的 LSPM 的性能明显优于基于直方图的线性 SPM 核,甚至优于非线性的 SPM 核。

传统的 SPM 方法需要结合非线性的分类器才能获得很好的识别分类性能。2010 年 Wang 等人^[13] 提出一种简单但有效的局部抑制线性编码方案(Locality-constrained Linear Coding, LLC),取代了传统 SPM 中的 VQ 编码。LLC 使用局部抑制把每一个描述子映射到它的局部坐标系中,然后使用 max pooling 方法集合映射后的坐标生成最后的特征表示,最后将其代入线性 SVM 分类器进行训练和测试。实验结果表明该方法的性能明显优于传统的非线性 SPM 方法。

3 本文方法

上述方法都是在 SIFT 的基础上提取特征,忽略了人工特征的局限性^[14]。而自然图像本身具有冗余性,直接利用局部抑制线性编码(LLC)的无监督学习同样可以提取特征,结合线性 SPM 和线性 SVM 进行识别分类的准确率也很高,并且速度更快。本文提出一种基于局部抑制线性编码的快速识别自然图像的 LLCL 方法,如图 1 所示。

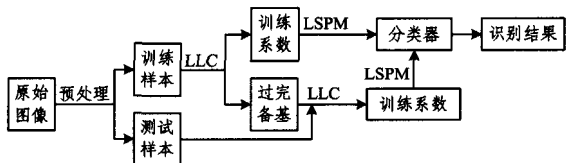


图 1 本文方法

方法总体上分为训练过程和测试过程。训练过程为:经过预处理的训练样本通过 LLC 学习得到一组过完备基 B 和投影到这组基上的训练系数 s ,接着用 LSPM 计算这些系数,并把计算结果代入线性 SVM 中进行训练得到分类器;测试过程的步骤和训练过程基本相同,只是不需要再次训练过完备基,而是直接利用训练过程中得到的过完备基 B 来求测试系数。

3.1 稀疏编码

与生物视觉系统有相似的自适应统计结构的图像称为自

然图像,它涉及到人类感知系统进化发展对自然环境的作用。从统计学上讲,自然图像有显著的冗余性,当一幅自然图像传给感知系统时,只有少部分神经元(主视皮层 V1 区简单细胞感受野)对该图像有较强的响应,而大多数神经元的响应很弱甚至没有响应,神经元细胞感受野的位置选择性、方向选择性以及频率选择性这 3 大反应特性与这种现象的产生有着直接关系。

1996 年 Olshausen 和 Field 提出稀疏编码基础模型^[7],该模型用稀疏惩罚项作为稀疏性约束,这样可以模拟哺乳动物视觉系统主视皮层 V1 区简单细胞感受野的 3 大响应特性,是一种人工神经网络方法。

然而 Yu 等人^[15]发现稀疏编码往往不具有局部性,因为其为了拥有稀疏性,经常会使并不相似的两组基进行加权线性组合来表示相似的两个输入向量,显然这是不合理的。为此, Yu 等人修改 SC,提出局部坐标编码(Local Coordinate Coding, LCC),并指出在某些假设下局部性比稀疏性更加重要。但是,与稀疏编码相似,LCC 仍需要解决 ℓ_1 范数优化问题,因此还是需要高昂的计算代价。

3.2 LLC

正如 LCC 中提到,局部性比稀疏性更重要,局部性能够保证编码的稀疏性,但稀疏性并不能保证编码具有局部性。2010 年, Wang 等人^[13]提出局部抑制线性编码 LLC,用局部抑制项替代稀疏惩罚项。LLC 相当于 LCC 的快速实现版本,它使用局部抑制把每一个描述子映射到它的局部坐标系中。具体地, LLC 编码公式如下:

$$\min \sum_{i=1}^N \|y_i - B s_i\|^2 + \beta \|d_i \odot s_i\|^2 \quad (1)$$

$$\text{s. t. } 1^T s_i = 1, \forall i$$

式中, \odot 代表 *element-wise* 乘法, $d_i \in R^M$ 是局部适配器,具体地:

$$d_i = \exp\left(\frac{\text{dist}(y_i, B)}{\sigma}\right) \quad (2)$$

式中, $\text{dist}(y_i, B) = [\text{dist}(y_i, b_1), \dots, \text{dist}(y_i, b_M)]^T$ 是 y_i 到 b_j 的欧几里德距离,这样可以通过减少距离较大项对应的系数 s_i 来最小化 $\|d_i \odot s_i\|^2$,从而使得剩下的系数具有局部性, σ 是用来调整局部适配器的权重衰减速度。通常情况下,从 $\text{dist}(y_i, B)$ 中去掉 $\max(\text{dist}(y_i, B))$ 来标准化 d_i 到 $(0, 1]$ 之间,约束项 $1^T s_i = 1$ 保证了 LLC 编码的平移不变性。注意,式(1)采用的不是 ℓ_1 范数,然而由于其解中只有少数有效的值,因此从某种意义上来说是稀疏的,实际上是通过简单地设置阈值使较小系数为零。

LLC 学习到的系数具有更高的区分度,因为其根据相似的输入向量,可以学习到相似的系数,而对于不相似的输入向量,学习到的则是不相似的系数,因此 LLC 在识别分类上的性能将会更好。相比于 VQ 和 SC, LLC 中的局部正则项 $\|d_i \odot s_i\|^2$ 表现出几个很吸引人的性能^[13]。

(1)更好地重构^[8]:在 VQ 中,相似的两个输入向量分别由不同的两个基向量重构表示,这存在着很大的量化误差,此外, VQ 忽略了不同基向量之间的相关性;而在 LLC 中每一个输入向量由多个基向量加权线性组合更精确地表示,如图 2(a)、图 2(c)所示,而且通过共享基向量 LLC 编码捕获了相似输入向量之间的相关性,这显著地减小了重构误差,从而使重构性能更好。

(2)局部更加平滑:SC 因使用多个基向量表示某一输入向量而有着少量的重构误差,但是 SC 或许选择非常不同的两组基向量来表示两个相似的输入向量以支持稀疏性,如图 2(b)所示,这丢失了基向量之间的相关性;而 LLC 中的局部适配器能确保学习得到相似的系数,因此使局部更平滑。

(3)具有解析解:稀疏编码的快速实现^[9]采用交替迭代更新法,它是一个高效的算法,但没有解析解,相比于拥有解析解的问题,其速度当然不是很快;而 LLC 方法有解析解,可以通过下式得到:

$$\begin{aligned} \tilde{s}_i &= (C_i + \lambda \text{diag}(\mathbf{d})) \setminus \mathbf{1} \\ c_i &= \tilde{c}_i / 1^T \tilde{c}_i \end{aligned} \quad (3)$$

式中, $C_i = (\mathbf{B} - \mathbf{1}\mathbf{x}_i^T)(\mathbf{B} - \mathbf{1}\mathbf{x}_i^T)^T$ 表示协方差矩阵。由此可见,实际上 LLC 的实现是很快的。

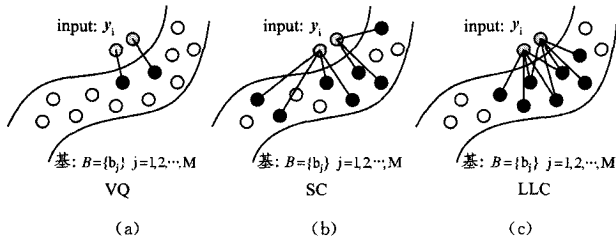


图 2 3 种编码方法的重构示意图

3.3 线性 SPM

设 z_i 表示图像 I_i , 则本文采用的线性 SPM(LSPM)核为:

$$k(z_i, z_j) = \mathbf{z}_i^T \mathbf{z}_j = \sum_{s=0}^{\ell} \sum_{t=1}^{\ell} \langle z_i'(s, t), z_j'(s, t) \rangle \quad (4)$$

式中, $\langle z_i, z_j \rangle = \mathbf{z}_i^T \mathbf{z}_j$, $z_i'(s, t)$ 是图像 I_i 在尺度 ℓ 下第 (s, t) 个分块上的系数的 maxpooling 统计。那么对于一个二分类问题, SVM 决策函数变为:

$$f(\mathbf{z}) = (\sum_{i=1}^n \alpha_i z_i + b)^T \mathbf{z} + b = \mathbf{w}^T \mathbf{z} + b \quad (5)$$

虽然基于直方图的线性 SPM 核的性能不好,但是基于稀疏编码统计的线性 SPM 核总是能获得很高的识别分类准确率,这归因于 3 个方面:(1)相比于矢量量化,稀疏编码大大减少了量化误差,有更好的重构性;(2)由于自然图像 patch 具有稀疏性,因此稀疏编码特别适合于自然图像数据;(3)通过 maxpooling 计算的特征更能显著地表现局部信息。

本文实验表明,对几个常用数据集直接采用 LLC 进行局部特征提取,然后结合线性 SPM 和线性 SVM 能获得较好的识别分类效果,同时大大减少了单张图像识别平均耗时。

3.4 优化方法

LLC 中求解 \mathbf{s} 与 \mathbf{B} 仍然需要使用交替迭代的方法,由于 \mathbf{s} 具有解析解,因此优化方法的重点在于 \mathbf{B} 的求解,文献[13]给出了快速编码 LLC 近似算法与在线增量基优化算法。

(1)快速编码 LLC 近似算法

求解式(1)实际上是执行特征选择:对于每一个描述子,选择局部基形成局部坐标系统。更快的 LLC 近似算法加速了编码过程,不是求解式(1),该方法简单地采用 y_i 的 K ($K < D < M$) 个最近邻(K Nearest Neighbors, KNN)^[16] 作为局部基 \mathbf{B}_i , 这样无需式(1)中的局部抑制项,得到公式:

$$\begin{aligned} \min_{\mathbf{s}} \sum_{i=1}^N \| y_i - \mathbf{B}_i \tilde{s}_i \|^2 \\ \text{s. t. } 1^T \tilde{s}_i = 1, \forall i \end{aligned} \quad (6)$$

该公式将计算时间复杂度从 $O(M^2)$ 减少到 $O(M+K^2)$, 其中 $K \ll M$, 因此速度加快很多。

(2)在线增量基优化算法

考虑式(1),实际上训练描述子数量 N 通常是很大的,这样在每一次迭代中加载所有的 LLC 系数进行计算所消耗的内存是巨大的。为此,文献[13]提出在线增量基优化算法,其基本思想是:先用 K-means 算法初始字典 \mathbf{B} , 然后递增地循环加载训练描述子来更新 \mathbf{B} , 在每一次迭代中,加载一个 y_i 或一小批(batch)的描述子,接着使用当前的 \mathbf{B} 求解式(1)获得相应的系数,得到的系数通过梯度下降法来更新 \mathbf{B} 。假设训练描述子 $\mathbf{Y} \in \mathbb{R}^{D \times N}$, $\mathbf{B}_{\text{init}} \in \mathbb{R}^{D \times K}$, 该算法具体流程如下。

Step1:初始化 \mathbf{B}

将 K-means 聚类得到的 \mathbf{B}_{init} 赋给 \mathbf{B} ;

Step2:局部抑制项参数计算

用零向量初始化 \mathbf{d} , 再通过式(2)计算 \mathbf{d} , 最后归一化到 $(0, 1]$ 之间;

Step3:编码

固定基 \mathbf{B} , 结合式(3)来求解式(1)得到系数 \mathbf{s}_i ;

Step4:除去偏差

只留下绝对值大于设定的阈值(文献[13]中设定为 0.01)的系数对应的基向量 \mathbf{B}_i , 结合式(3)求解式(5), 得到除去偏差后的系数 \tilde{s}_i ;

Step5:更新 \mathbf{B} , 迭代结束

通过以下公式更新基向量:

$$\Delta \mathbf{B}_i \leftarrow -2 \tilde{s}_i (\mathbf{x}_i - \mathbf{B}_i \tilde{s}_i), \mu \leftarrow \sqrt{1/\ell}$$

$$\mathbf{B}_i \leftarrow \mathbf{B}_i - \mu \Delta \mathbf{B}_i / \|\tilde{s}_i\|_2$$

$$\mathbf{B}(:, \text{id}) \leftarrow \text{proj}(\mathbf{B}_i)$$

当迭代次数小于等于 N 时,继续迭代;否则停止迭代,循环结束。

4 实验结果与分析

本文将在 3 个常用数据集 Caltech-101^[17]、Caltech-256^[18] 和 15 Scenes^[10] 上测试本文方法的平均识别准确率和单张图像识别的平均耗时,并与其他两种图像识别分类方法进行比较。

(1)ScSPM 方法^[12]:对 SIFT 特征进行 SC 处理,然后结合线性 SPM 和线性 SVM;

(2)LLC 方法^[13]:对 SIFT 特征进行 LLC 处理,然后结合线性 SPM 和线性 SVM;

(3)本文 LLC 方法:直接对图像进行 LLC 处理来提取系数,然后进行线性 SPM 计算,最后用简单的线性 SVM 进行分类器训练与预测。

当对 LLC 处理之后获得的系数进行线性 SPM 计算时,采用的是 maxpooling 池化方法。

将图像转换成灰度图并转换到 double 类型空间进行预处理后,获取样本向量时 patch 的大小选为 15, 每张图像抽取的 patch 的数量设置为 1000; 在 LLC 处理时,仅仅采用 LLC 近似算法, K 近邻的数量设置为 5。

本文方法的重点在于获得很好的图像识别准确率的同时很大程度地减少了单张图像识别耗时,因此,此处需要给出单张图像识别耗时的定义。给定大小为 N 的测试样本集,提取测试系数的耗时为 T_1 , 线性 SPM 计算的耗时为 T_2 , LSPM 预测耗时为 T_3 , 其他程序读写耗时为 t , 则本文测试样本的单张识别耗时定义为:

$$T = (T_1 + T_2 + T_3 + t) / N \quad (7)$$

实验环境:64 位 Windows8 操作系统, CPU 主频 2.50

GHz, 4GB 内存, Matlab2013。

由于本文方法毕竟不是基于 SIFT 这类人工特征,而且因实验环境的限制,我们从 Caltech-101、Caltech256 和 15 Scenes 3 个数据集中分别选取 5 个类别来进行实验分析。

4.1 Caltech-101

Caltech-101 数据集包含 9144 张图片,共 101 个类别,其中包括家具、动物、运动器材、车辆等,每一类图像的数量最少是 31,最多是 800,部分示例图像如图 3 所示。



图 3 Caltech-101 数据集的部分图像示例图

本文对从 Caltech-101 选取的 5 个类别进行训练时,训练数量分别为 10、20 和 30,测试数量不超过 30,训练基大小为 1024,SPM 使用 $4 \times 4, 2 \times 1, 1 \times 1$ 子区域。实验中,样本尺寸被重置为像素,将本文方法和另外两种方法的实验结果做对比,具体如表 1 和表 2 所列。

表 1 Caltech-101 数据集上的图像识别平均准确率(%)

Method	10	20	30
ScSPM	84.08	89.48	92.25
LLC	82.75	88.77	91.94
本文方法	81.43	87.04	90.65

表 2 Caltech-101 数据集上的单张图像识别平均耗时(s)

Method	10	20	30
ScSPM	2.67	2.92	3.21
LLC	2.23	1.94	2.01
本文方法	0.37	0.42	0.50

从表 1 和表 2 可以看到,本文方法在 Caltech-101 的 5 个类别上的识别平均准确率接近 ScSPM 和 LLC 的准确率,但却大大减少了单张图像识别平均耗时,仅需 0.5s(30 个训练样本)。

4.2 Caltech-256

Caltech-256 数据集包含 30607 张图片,共 256 个类别。相比于 Caltech-101,Caltech-256 由于类与类之间和每类中图像主体位置变化明显,因此对其识别分类更有难度,部分示例图像如图 4 所示。



图 4 Caltech-256 数据集的部分图像示例图

Caltech-256 中每类图像数量最少是 80,本次实验中选取的 5 个类别的样本训练数量分别设为 15、30 和 45,实验结果

同样与另外两种方法作对比,具体见表 3 和表 4。

表 3 Caltech-256 数据集上的图像识别平均准确率(%)

Method	15	30	45
ScSPM	64.03	68.17	69.80
LLC	62.94	67.98	68.98
本文方法	61.47	65.46	67.12

表 4 Caltech-256 数据集上的单张图像识别平均耗时(s)

Method	15	30	45
ScSPM	2.85	2.92	2.86
LLC	2.03	2.01	2.05
本文方法	0.64	0.50	0.48

从表 3 和表 4 可以看到,本文方法在 Caltech-256 的 5 个类别上获得了 69.12% 的识别平均准确率,接近 ScSPM 和 LLC 的识别平均准确率,但单张图像的识别平均耗时为 0.48s(45 个训练样本),相比 ScSPM 的 2.86s 和 LLC 的 2.2s 有显著的降低。

4.3 15 Scenes

15 Scenes 数据集是由几个研究者^[10,11,19]收集的,共包含 4485 张图片,15 个类别,这些类包括厨房、森林和街道等,每一类的数量在 200 和 400 之间,部分示例图像如图 5 所示。



图 5 15 Scenes 数据集的部分图像示例图

实验中选取的样本训练数量分别为 40、60 和 80,实验结果具体见表 5 和表 6。

表 5 15 Scenes 数据集上的图像识别平均准确率(%)

Method	40	60	80
ScSPM	91.43	91.95	92.35
LLC	90.21	90.51	91.19
本文方法	88.67	89.69	90.39

表 6 15 Scenes 数据集上的单张图像识别平均耗时(s)

Method	40	60	80
ScSPM	2.86	2.72	3.21
LLC	2.24	2.25	2.21
本文方法	0.57	0.61	0.52

从表 5 和表 6 可以看到,本文方法在 15 Scenes 的 5 个类别上获得了 90.39% 的识别平均准确率,与 ScSPM 和 LLC 的准确率只相差了 0.8~2.8 个百分点,然而单张图像的识别平均耗时为 0.52s(80 个训练样本),仅为 ScSPM 的 3.12s 和 LLC 的 2.21s 的 1/5~1/3,显著地减少了单张图像的识别平均耗时。

4.4 讨论

从 4.1 节—4.3 节这 3 组实验可以清楚地看到,本文方法在类别不多的情况下能获得较好的识别准确率,同时显著地减少了单张图像识别耗时,这充分说明了该方法的有效性和鲁棒性。

为了进行更综合的分析,评估了 LLC 近似算法中 K 最近邻数量和训练基大小对本文方法性能的影响。

(1)图 6 和图 7 示出了分别使用 2、5、10、20 和 40 个最近邻数量下本文方法的性能。

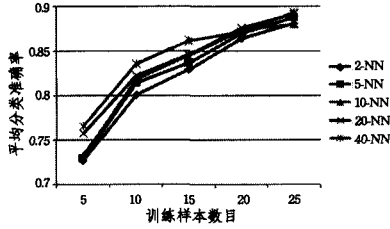


图 6 不同最近邻数量时的识别平均准确率

分析图 6 可知,使用更多的最近邻数量的性能会相对更好,其平均识别准确率相对更高;而对于每一类最近邻数量,随着训练样本数目的增加,平均识别准确率提高。

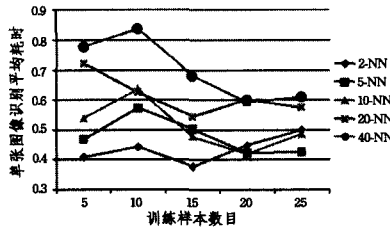


图 7 不同最近邻数量下单张图像识别的平均耗时

图 7 中,最近邻数量越大,单张图像识别平均耗时相对越多,但并不是绝对的,这是由于实验环境限制而导致的不稳定所致;其次,单张图像识别平均耗时和训练样本数目并没有直接关系。

(2)下面比较使用 3.4 节(1)中 K-means 算法的不同大小的超完备基时本文方法的性能,分别用 512、1028 和 2048 进行实验,实验结果如图 8 和图 9 所示。

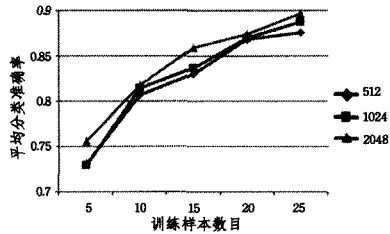


图 8 不同基大小的平均分类准确率

图 8 清楚地显示,在相同参数的情况下,训练基越大,本文方法在 Caltech-101 的 5 类图像上的识别平均准确率越高,因为基越大则能越较好地表示图像局部特征;此外,在每一类训练基中,训练样本数目的增加都相应地提高了本文方法的识别平均准确率。

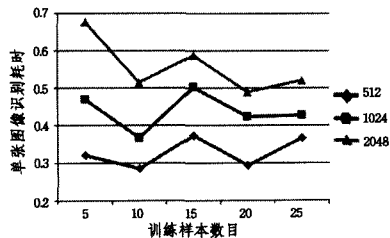


图 9 不同基大小的单张图像识别平均耗时

从图 9 可以看到,使用的训练基越大则相应的单张图像识别平均耗时就越长,这是因为使用越大的基时计算系数 s

所需要的时间就越多,而对于每一类训练基,训练样本数目的多少对单张图像识别平均耗时并没有直接影响。

结束语 稀疏编码和局部抑制线性编码虽然在计算机视觉和模式识别的一些方面获得了很好的应用,但之前的应用都是在 SIFT、HOG 等人工特征的基础上进行的,图像识别耗时并不低。本文提出了一种快速识别自然图像的 LLCL 方法,由于自然图像有着复杂的拓扑结构且具有稀疏特性,我们舍弃了传统的基于人工特征的识别方法,而是直接采取 LLC 算法进行无监督学习来提取特征,然后通过线性 SPM 计算并结合线性 SVM 进行识别分类。实验结果表明,本文方法在类别不多的情况下能获得较高的识别平均准确率,同时显著地减少了单张图像识别平均耗时,这将在很多图像识别应用中具有一定的参考价值。

参考文献

- [1] Xie Zhao, Gao Jun. A novel method for scene categorization with constraint mechanism based on gaussian statistical model[J]. Acta Electronica Sinica, 2009, 37(4): 733-738(in Chinese)
谢昭,高隽.基于高斯统计模型的场景分类及约束机制新方法[J].电子学报,2009,37(4):733-738
- [2] Sonka M, Hlavac V, Boyle R. Image processing, analysis, and machine vision[M]. Toronto: Thomson, 2008
- [3] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110
- [4] Cox D, Pinto N. Beyond simple features: A large-scale feature search approach to unconstrained face recognition[C]//Proceedings of the 2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops. Piscataway: IEEE, 2011: 8-15
- [5] Yang C, Zhu Y, Chen G. New classifier ensemble method based on rough set attribute reduction[J]. Application Research of Computers, 2012, 29(5): 1648-1650
- [6] Huang R, Lang F, Shi Z. Log-Gabor and 2D semi-supervised discriminant analysis based face image retrieval[J]. Application Research of Computers, 2012, 29(1): 393-396
- [7] Olshausen B A. Emergence of simple-cell receptive field properties by learning a sparse code for natural images[J]. Nature, 1996, 381(6583): 607-609
- [8] Zhu Xiao-feng, Huang Zi, Shen Heng-tao, et al. Linear cross-modal hashing for efficient multi-media search[C]//Proc of the 21st ACM International Conference on Multimedia. 2013: 143-152
- [9] Lee H, Battle A, Raina R, et al. Efficient sparse coding algorithms[M]//Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2006: 801-808
- [10] Lazebnik S, Schmid C, Ponce J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories[C]//Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2006: 2169-2178
- [11] Li Fei-fei, Perona P. A bayesian hierarchical model for learning natural scene categories[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, CA: IEEE Press, 2005: 524-531

(下转封三)

- gram [J]. *Computer Science*, 2013, 40(8): 313-315 (in Chinese)
 龚劬, 付云凤, 叶剑英, 等. 基于二维直方图重建的 Otsu 图像分割算法[J]. *计算机科学*, 2013, 40(8): 313-315
- [11] Gao Shang-bing, Zhou Jing-bo, Yan Yun-yang. A new superpixel based spectral clustering for image segmentation [J]. *Journal of Nanjing University(Natural Sciences)*, 2013, 49(2): 169-175 (in Chinese)
 高尚兵, 周静波, 严云洋. 一种新的基于超像素的谱聚类图像分割算法[J]. *南京大学学报(自然科学版)*, 2013, 49(2): 169-175
- [12] Luo Hui-lan, Zhong Rui, Kong Fan-sheng. Method of point tracking based on superpixel [J]. *Journal of Image and Graphics*, 2014, 19(3): 428-438 (in Chinese)
 罗会兰, 钟睿, 孔繁胜. 基于超像素的点追踪方法[J]. *中国图象图形学报*, 2014, 19(3): 428-438
- [13] Felzenszwalb P F, Huttenlocher D P. Efficient graph-based image segmentation [J]. *International Journal of Computer Vision*, 2004, 59(2): 167-181
- [14] Shi J, Malik J. Normalized cuts and image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 888-905
- [15] Moore A P, Prince S, Warrell J, et al. Superpixel lattices [C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*. 2008: 1-8
- [16] Wang Ai-qi, Qiu Tian-shuang. Superpixels construction method based on geodesic distance [J]. *Journal of Dalian University of Technology*, 2012, 52(4): 610-614 (in Chinese)
 王爱齐, 邱天爽. 基于测地距离的超像素生成方法[J]. *大连理工大学学报*, 2012, 52(4): 610-614
- [17] Luo Xue-gang, Lv Jun-rui, Wang Hua-jun, et al. Color image segmentation based on superpixels and reciprocal nearest neighbors clustering [J]. *Journal of Guangxi University(Natural Sciences)*, 2013, 38(2): 374-378 (in Chinese)
 罗学刚, 吕俊瑞, 王华军, 等. 基于超像素的互惠最近邻聚类彩色图像分割[J]. *广西大学学报(自然科学版)*, 2013, 38(2): 374-378
- [18] Vincent L, Soille P. Watersheds in digital spaces; an efficient algorithm based on immersion simulations [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991, 13(6): 583-598
- [19] Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(5): 603-619
- [20] Vedaldi A, Soatto S. Quick shift and kernel methods for mode seeking [M]// *Computer Vision-ECCV 2008*. Berlin: Springer-Verlag, 2008: 705-718
- [21] Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(11): 2274-2282
- [22] Nan Bing-fei, Mu Zhi-chun. SLIC0-based superpixel segmentation method with texture fusion [J]. *Chinese Journal of Scientific Instrument*, 2014, 35(3): 527-534 (in Chinese)
 南柄飞, 穆志纯. 基于 SLIC0 融合纹理信息的超像素分割方法[J]. *仪器仪表学报*, 2014, 35(3): 527-534
- [23] Austvoll I, Kwolek B. Region covariance matrix-based object tracking with occlusions handling [J]. *Lecture Notes in Computer Science*, 2010, 6374: 201-208
- [24] Khan S S, Ahmad A. Cluster center initialization algorithm for K-means clustering [J]. *Pattern Recognition Letters*, 2004, 25(11): 1293-1302
- [25] Von Luxburg U. A tutorial on spectral clustering [J]. *Statistics and Computing*, 2007, 17(4): 395-416
- [26] Karacan L, Erdem E, Erdem A. Structure-preserving image smoothing via region covariances [J]. *ACM Transactions on Graphics*, 2013, 32(6): 176
- [27] Liu M Y, Tuzel O, Ramalingam S, et al. Entropy rate superpixel segmentation [C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*. 2011: 2097-2104
- [28] Mori G. Guiding model search using segmentation [C]// *Proceedings of the IEEE International Conference on Computer Vision(ICCV)*. 2005: 1417-1423
- [29] Grundland M, Gibbs C, Dodgson N A. Stylized rendering for multiresolution image representation [C]// *International Society for Optics and Photonics, Electronic Imaging*. 2005: 280-292
- [30] DeCarlo D, Santella A. Stylization and abstraction of photographs [J]. *ACM Transactions on Graphics*, 2002, 21(3): 769-776

(上接第 312 页)

- [12] Yang J, Yu K, Gong Y, et al. Linear spatial pyramid matching using sparse coding for image classification [C]// *Proceedings of the 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, DC: IEEE Computer Society, 2009: 1794-1801
- [13] Wang J, Yang J, Yu K, et al. Locality-constrained linear coding for image classification [C]// *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, DC: IEEE Computer Society, 2010: 3360-3367
- [14] Bengio Y. Learning deep architectures for AI [M]. Hanover: Now Publishers Inc, 2009
- [15] Yu K, Zhang T, Gong Y. Nonlinear learning using Local Coordinate Coding [C]// *Proceedings of NIPS*. 2009: 2223-2231
- [16] Zhang H, Berg A, Maire M, et al. Svm-knn: Discriminative nearest neighbor classification for visual category recognition [C]// *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE Press, 2006: 2126-2136
- [17] Li Fei-fei, Fergus R, Perona P. Learning generative visual models from few training examples: an incremental bayesian approach tested on 101 object categories [J]. *Computer Vision and Image Understanding*, 2010, 106(1): 59-70
- [18] Holub G G, Perona P A D. Caltech-256 object category dataset; Technical Report 7694 [R]. California Institute of Technology, 2007
- [19] Oliva A, Torralba A. Modeling the shape of the scene: A holistic representation of the spatial envelop [C]// *IJCV*. 2001
- [20] Wagner A, Wright J, Ganesh A, et al. Toward a practical face recognition system: Robust alignment and illumination by sparse representation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(2): 372-386
- [21] Grosser, Rainar, Kwongh, et al. Shift-invariant sparse coding for audio classification [EB/OL]. [2014-12-01]. <http://axon.cs.byu.edu/Dan/778/papers/Sparse%20Coding/ng3.pdf>
- [22] Huang Wen-ming, Cai Wen-zheng, Deng Zhen-rong. Cerebrospinal fluid images fast recognition model based on sparse coding [J]. *Journal of Computer Applications*, 2014, 34(7): 2040-2043