

文章编号:1001-9081(2015)01-0224-07

doi:10.11772/j.issn.1001-9081.2015.01.0224

基于快速鲁棒特征集合统计特征的图像分类方法

王 润^{1*}, 吕学强¹, 张 凯², 李 卓¹

(1. 北京信息科技大学 网络文化与数字传播北京市重点实验室, 北京 100101; 2. 北京市朝阳区市政市容管理委员会 交通运行协调指挥中心, 北京 100021)
(*通信作者电子邮箱 wangshu.skyrider08@gmail.com)

摘要:针对现有利用快速鲁棒特征(SURF)进行图像分类的方法中存在的效率低、正确率低的问题,提出一种利用图像SURF集合的统计特征进行图像分类的方法。该方法将SURF的各个维度及尺度信息视为各自独立的随机变量,并利用拉普拉斯响应区分不同数据。首先,获取图像的SURF向量集合;然后,分维度计算SURF向量集合的一阶中心绝对矩、带权一阶中心绝对矩等统计特征,并构建特征向量;最后,结合支持向量机(SVM)进行图像分类。在Corel 1K 图像库上的实验结果表明,该方法查准率较 SURF 直方图方法和三通道 Gabor 纹理特征方法分别提高 17.6% 和 5.4%。通过与 HSV 直方图特征进行高级特征融合,可获得良好的分类性能。与 SURF 直方图结合 HSV 直方图方法、三通道 Gabor 纹理特征结合 HSV 直方图方法、基于视觉词袋(BoVW)模型的多示例学习方法相比,查准率分别提高了 5.2%, 6.8%, 3.2%。

关键词:快速鲁棒特征;图像分类;统计特征;随机变量;支持向量机

中图分类号: TP18; TP391 **文献标志码:**A

Image classification approach based on statistical features of speed up robust feature set

WANG Shu^{1*}, LYU Xueqiang¹, ZHANG Kai², LI Zhuo¹

(1. Beijing Key Laboratory of Internet Culture and Digital Dissemination Research, Beijing Information Science and Technology University, Beijing 100101, China;
2. Municipal Transportation Operations Coordination Center,
Beijing Chaoyang District Municipal Commission of City Administration and Environment, Beijing 100021, China)

Abstract: The current method of image classification which uses the Speed Up Robust Feature (SURF) is low in efficiency and accuracy. To overcome these shortages, this paper proposed an approach for image classification which uses the statistical features of the SURF set. This approach took all dimensions and scale information of the SURF as independent random variables, and split the data with the sign of Laplace response. Firstly, the SURF vector set of the image was got. Then the feature vector was constructed with the first absolute order central moments and weighted first absolute order central moments of each dimision. Finally, the Support Vector Machine (SVM) accomplished the image classification process with this vector. The experimental results show that the precision of this approach is better than that of the methods of SURF histogram and 3-channel-Gabor texture features by increases of 17.6% and 5.4% respectively. By combining this approach with the HSV histogram, a high-level feature fusion method was got, and good classification performance was obtained. Compared with the fused method of the SURF histogram and HSV histogram, the fused method of 3-channel-Gabor texture features and HSV histogram, and the multiple-instance-learning method based on the model of Bag of Visual Word (BoVW), the fused method of this approach and HSV histogram has better precision with the increases of 5.2%, 6.8% and 3.2% respectively.

Key words: Speed Up Robust Feature (SURF); image classification; statistical feature; random variable; Support Vector Machine (SVM)

0 引言

随着计算机网络的发展和手机、照相机等图像获取设备的日益增加,图像的数量迅速增长,人们对有效进行图像管理的需求愈发强烈。基于人工标注的方法存在耗时长、开销大、受主观因素影响大等缺点,因此基于内容的图像自动分类技术受到研究者的广泛关注。但图像受到摄像设备硬件参数、拍摄时的光照、角度、取景方式、后期加工处理以及不同个体

对图像内容理解存在差异等各方面因素的影响,图像自动分类的难度较大。

早期研究采用的方法是将图像的一种全局特征作为分类依据。近年来主要分为两类方法:一类基于视觉词袋(Bag of Visual Word, BoVW)方法;一类基于特征融合方法。前者如文献[1–2],通常采用分割或分块策略,对分割后的小块(Blob)提取该块的图像特征,通过聚类组成视觉词汇,经过机器学习过程后利用视觉词汇进行分类;后者如文献[3–4],

收稿日期:2014-07-31;修回日期:2014-09-17。 基金项目:国家自然科学基金资助项目(61271304);北京市教委科技发展计划重点项目暨北京市自然科学基金B类重点项目(KZ201311232037);北京市属高等学校创新团队建设与教师职业发展计划项目(IDHT20130519)。

作者简介:王润(1986–),男,山东昌乐人,硕士研究生,CCF会员,主要研究方向:图像标注; 吕学强(1970–),男,山东鱼台人,教授,博士,CCF会员,主要研究方向:中文与多媒体信息处理; 张凯(1985–),男,河南焦作人,工程师,硕士,主要研究方向:多媒体信息处理; 李卓(1983–),男,河南信阳人,讲师,博士,主要研究方向:移动互联网。

该类方法通过不同策略将不同特征融合,从而提升整体分类效果。

随着局部特征提取方法的日益增多,由于该类特征描述了更多的图像细节信息,且在旋转、光照、尺度不变等特性上较全局特征具有优势,因此被广泛地引入到图像分类任务中。Li等^[5]的研究表明,基于稠密采样方法的局部特征不受图像内容影响,图像分块数目不受限制,更适合词袋模型。文献[5~7]均采用了该类局部特征。但文献[8]指出,此类特征在不变性方面表现较差。而基于角点检测、斑点检测、区域检测类的方法虽然能够保证特征具有多种不变性,但由于无法确保在图像的所有区域都能有效地提取局部特征点,因而与词袋模型结合时效果不如稠密采样类方法。当前,使用斑点检测类特征的图像分类方法通常被限制在特定场景下,如文献[8~10]。

快速鲁棒特征(Speed Up Robust Feature, SURF)是Bay等^[11]于2006年提出的一种局部特征提取算法,该算法是斑点检测方法,在尺度、旋转、噪声、光照等方面具有很强的不变性,且该算法计算效率较高,因而被广泛应用于目标识别等领域。但该算法对每张图片可提取的特征点数目不一致,因此很难直接用于图像分类或检索。现有使用SURF的针对通用图像库的图像分类方法中,SURF直方图方法^[12]可以解决图像特征点数目不一致的问题,但正确率较低,且该方法生成的特征向量维数较高,降低了其实用价值。

因此,本文提出基于统计方法的快速鲁棒特征集合矩(SURF-Moment, SURF-M)特征,并将其用于图像分类。该特征具有良好的稳定性,对构成结构和纹理特征相似的图像具有良好的分类效果。首先提取图像的SURF组成该图像的SURF向量集合,统计每一维的一阶中心绝对矩和带权一阶中心绝对矩,按拉普拉斯响应分成不同集合后统计每一维的一阶中心绝对矩,以此形成特征向量,最后通过LIBSVM^[13]进行多类别分类。在Corel 1K图像库上的实验结果表明,该特征可用于图像分类,且在进行特征融合后分类正确率可达86.8%。

1 SURF 提取和 SURF 直方图方法

1.1 SURF 特征点位置的确定

对于一幅图像 $I(x, y)$, 将图像的尺度空间表示为一个函数 $L(x, y, \sigma)$:

$$L(x, y, \sigma) = I(x, y) * G(x, y, \sigma) \quad (1)$$

其中 $G(x, y, \sigma)$ 为尺度可变的高斯函数。

SURF 提取算法在建立尺度空间中采用矩形滤波器^[14]代替二阶高斯滤波器,通过增大矩形滤波模板尺寸的方法建立图像尺度空间,并借助积分图像^[15],使图像与高斯二阶微分模板的卷积运算转化为对积分图像的加减运算,减少了运算量。

SURF 利用 Hessian 矩阵检测极值点。对图像 I 中点 (x, y) , 其在尺度 σ 写下的 Hessian 矩阵 H 表示为:

$$H = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix} \quad (2)$$

其中 $L_{xx}(x, y, \sigma)$ 等表示高斯二阶偏导在点 (x, y) 处与图像 I 的卷积。

为准确逼近高斯核函数,矩阵 H 的判别式 $\det(H)$ 表示为:

$$\det(H) = D_{xx}D_{yy} - (wD_{xy})^2 \quad (3)$$

其中: D_{xx} 、 D_{yy} 、 D_{xy} 分别是对 L_{xx} 、 L_{yy} 、 L_{xy} 的近似; w 为权重,一般取值为 0.9。Hessian 矩阵的迹的符号则被记录下来,并作为一个特征被用于加速匹配。

在采用 Hessian 矩阵求出极值后,在 $3 \times 3 \times 3$ 的立体邻域范围内进行非极大抑制,将比 26 个邻域点的值都大或都小的点作为特征候选点;然后通过在尺度空间和图像空间中进行插值运算,得到稳定的特征点位置。

1.2 SURF 特征值的获取

描述 SURF 特征点时,需要首先确定其主方向。首先,以特征点为中心,检测 6σ 邻域范围以内的点在 x 、 y 方向上的边长为 4σ 的 Haar 小波响应,并给这些响应值赋以高斯权重系数,使靠近特征点的响应贡献大,从而得到一系列向量;然后,计算以 60° 范围内的响应之和形成的新向量,并以 60° 为单位进行窗口滑动,遍历整个圆形区域,选取最长的向量的方向作为特征点的主方向。

确定特征点的主方向后,首先,以特征点为中心将坐标轴旋转到主方向,选取边长为 20σ 的正方形区域,将其划分为等大的 4×4 个子区域,在每个子区域内,计算 25σ 范围内的 Haar 小波响应,分别标记为 d_x 、 d_y ;然后,对这些响应赋以高斯权重;最后,将各个子区域的响应及响应的绝对值相加,生成描述符的特征向量。每个子区域共生成 4 个值: $|d_x|$ 、 $|d_y|$ 、 $|d_x| + |d_y|$, 每个特征点由 4×4 个子区域构成,则形成 64 维特征向量(SURF-64)。如在计算特征值时将 d_x 、 d_y 区分正负分别统计,则每个子区域可获取 8 个特征值,形成 128 维特征向量(SURF-128)。

1.3 SURF 直方图方法

通过随机直方图^[16]方法,可将 SURF 向量集合转为一个指定长度的高维向量。通过这种方法获取的特征向量可以通过支持向量机(Support Vector Machine, SVM)进行分类。

SURF 直方图算法步骤描述如下:

- 1) 从每幅图片提取 n 个特征点,每个特征点为 64 维;
- 2) 对每个特征点,随机取 B 位,根据随机阈值向量投影到汉明空间,得到 B 位二进制数组,并将其转化为一个十进制数 H ;
- 3) 构建长度为 2^B ,初始值为 0 的向量,在该向量的第 H 位处加 1;
- 4) 将步骤 1) ~ 3) 重复 M 次,将生成的 M 个向量相加,得到长度为 2^B 的向量;
- 5) 将步骤 1) ~ 5) 重复 N 次,将生成的 N 个向量拼接为一个 $2^B \times N$ 维向量;
- 6) 计算各维与各维总和的比例,即获得每幅图片的 $2^B \times N$ 维 SURF 直方图。

2 基于 SURF 集合统计特征的图像分类方法

随机直方图方法利用了局部敏感哈希(Locality-Sensitive Hash, LSH)方法^[17]的原理。局部敏感哈希方法的核心思想是,图像的特征值是对图像内容的近似,通过一个哈希函数,使得邻近的特征值在经过映射后依然邻近,从而达到降低维数的目的。因为局部敏感哈希方法产生的是对特征值的近似,并非对图像内容的直接近似,因此单一哈希函数会有可能使得分类正确率较低,通常需要采用多个哈希函数提高正确率。

SURF 直方图使用了随机直方图方法,即通过统计在二进制化的特征值中某些随机选定的位置出现的 0,1 数字的分布情况,并将这种分布信息转化为一个固定维数的向量。这一方法有效地解决了每幅图像可以获取的 SURF 的特征点数目不相同的问题,使图片的 SURF 向量集合可以使用 SVM 分类器进行分类,并且在分类时获得一定效果。

但该方法考察的是过于细致的二进制数的某些特定位的数字分布规律。这种分布对于高维的特征而言,其区分能力的强弱很难判定,并且数字的不同二进制数位在实际分类过程中起到的作用大小并不一致,以上因素限制了该方法的分类效果。通过提高 SURF 直方图方法的 B 、 M 、 N 值可以有限地提高其效果,但会增大构建特征向量时的运算量和用于分类的特征向量的维数,降低其实用性。

2.1 SURF 特征点的特点

由 SURF 提取的方法可知,SURF 反映的是图像不同尺度下局部极值点和与之邻近的一定范围内的点之间相互关系的信息。它所含有的信息表示图像在不同尺度下的纹理特征或结构特征。

图 1 中表示 Corel 1K 图像库中大象分类下一张图片提取的 SURF 特征点。其中圆圈大小表示该特征点的尺度,圆圈内线段表示该特征点的主方向。

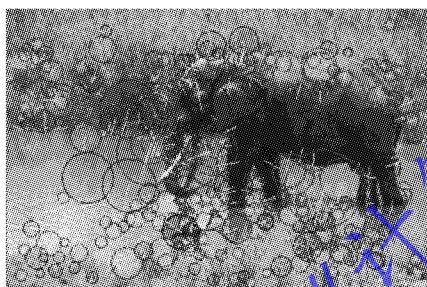


图 1 图像的 SURF 特征点

由 SURF 特征点获取的算法可知,SURF 特征点选取过程中经过了边缘抑制过程,因此其选择的特征点主要由纹理产生。如图 1 所示,SURF 特征点可以分为两类:1) 即相同物体区域内,由于物体纹理自身所具有的局部极值点而产生的特征点;2) 由于不同物体区域交汇而产生的边缘所产生的特征点。前者如草丛的纹理,通常会体现出一定规律性;后者如大象与草丛之间、水与大象之间的边缘等,是否存在规律性取决于边缘是否存在规律性。因为同类的图片中具有相似的纹理构成,且在画面中主要物体大小相似时,其斑点尺度也相似,因此对于纹理信息丰富的图片,越相似的图片在 SURF 提取的向量集合中所体现的统计特征也越相近。

通过观察图 1 可知,在特征点的中心区域,像素变化较为剧烈;而边缘区域趋于平缓。这是由 SURF 特征点的获取方法决定的。因为图像中取得的特征点周围的兴趣区域在经过高斯模糊、确定主方向、去除光照影响等过程之后,同种物体内部由纹理等因素产生的特征点所对应的区域会具有一定相似性,即 Haar 小波特征相似;由物体间边缘所产生的特征点则不具备这样的相似性,但此类特征点较少时,可将其视为噪声。

如果将图像认为是一个由有限个 SURF 特征点组成的集合,则在同一分类下已有足够多的图像时,同一分类的某一张图像,可以认为是由一定数量的已知的图像碎片(即 SURF 特征点)构成的。据此可假设,在同类图片所构成的特征向量

集合中,各个维度均可看作各自独立的随机变量,且该随机变量的分布可由已知的同类图片中获取的特征向量集合描述。这一假设可由统计分析验证。

2.2 SURF 特征值的统计特性

由辛钦大数定律,对一独立同分布随机变量序列 $X = \{a_i, 1 \leq i \leq n\}$,若 X 的数学期望存在,则服从大数定律,即对任意 $\varepsilon > 0$,如式(4) :

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{1}{n} \sum_{i=1}^n a_i - \mu\right| < \varepsilon\right) = 1 \quad (4)$$

该定律表明,由 X 产生的一系列独立同分布的随机变量序列的平均值依概率收敛为 X 的数学期望值,且该性质可以推广到 X 所具备的其他概率特性。

一阶中心绝对矩是一种常见统计量,表示样本与样本数学期望之间的平均偏离程度。设随机变量为 a ,则其一阶中心绝对矩 $m(a)$ 表示为式(5) :

$$m(a) = E|a - E(a)| \quad (5)$$

基于上文假设,对同一类图像中的 SURF 的每一维,其所对应的随机变量均有数学期望,因此各维数据对应的随机变量均满足一阶中心绝对矩存在的条件。

表 1 中所示为各分类图片提取 SURF 之后按文件随机排列后,分别取前 10%、前 25%、前 50% 与 100% 的特征值时,各维数据的一阶中心绝对矩。由于空间有限,只显示其中建筑、巴士、大象、花、食物、马等 6 个类别的 41~46 维。

表 1 不同分类不同数据规模下一阶中心绝对矩的对比

| 分 类 例/% | 维序号 | | | | | | |
|---------------|-----|--------|--------|--------|--------|--------|--------|
| | 41 | 42 | 43 | 44 | 45 | 46 | |
| 建筑 | 10 | 0.1301 | 0.1094 | 0.1055 | 0.1064 | 0.0469 | 0.0748 |
| | 25 | 0.1268 | 0.1103 | 0.0987 | 0.1054 | 0.0475 | 0.0767 |
| | 50 | 0.1280 | 0.1100 | 0.0970 | 0.1039 | 0.0476 | 0.0763 |
| | 100 | 0.1264 | 0.1090 | 0.0976 | 0.1018 | 0.0481 | 0.0750 |
| 巴士 | 10 | 0.1209 | 0.1154 | 0.1070 | 0.1059 | 0.0418 | 0.0717 |
| | 25 | 0.1204 | 0.1147 | 0.1074 | 0.1055 | 0.0422 | 0.0729 |
| | 50 | 0.1208 | 0.1171 | 0.1080 | 0.1070 | 0.0426 | 0.0735 |
| | 100 | 0.1207 | 0.1163 | 0.1075 | 0.1066 | 0.0435 | 0.0742 |
| 大象 | 10 | 0.1201 | 0.1025 | 0.1021 | 0.0919 | 0.0577 | 0.0777 |
| | 25 | 0.1208 | 0.1040 | 0.1043 | 0.0901 | 0.0576 | 0.0748 |
| | 50 | 0.1217 | 0.1037 | 0.1052 | 0.0903 | 0.0575 | 0.0749 |
| | 100 | 0.1228 | 0.1041 | 0.1051 | 0.0913 | 0.0572 | 0.0748 |
| 花 | 10 | 0.1418 | 0.1285 | 0.1011 | 0.0972 | 0.0572 | 0.0671 |
| | 25 | 0.1426 | 0.1277 | 0.0993 | 0.0981 | 0.0578 | 0.0676 |
| | 50 | 0.1384 | 0.1248 | 0.0947 | 0.0955 | 0.0589 | 0.0695 |
| | 100 | 0.1407 | 0.1264 | 0.0969 | 0.0969 | 0.0583 | 0.0687 |
| 食物 | 10 | 0.1222 | 0.1114 | 0.1030 | 0.0969 | 0.0516 | 0.0745 |
| | 25 | 0.1236 | 0.1116 | 0.1067 | 0.0977 | 0.0492 | 0.0741 |
| | 50 | 0.1243 | 0.1114 | 0.1068 | 0.0975 | 0.0496 | 0.0738 |
| | 100 | 0.1244 | 0.1126 | 0.1050 | 0.0983 | 0.0499 | 0.0742 |
| 马 | 10 | 0.1059 | 0.0890 | 0.0954 | 0.0806 | 0.0569 | 0.0750 |
| | 25 | 0.1067 | 0.0921 | 0.0963 | 0.0825 | 0.0564 | 0.0772 |
| | 50 | 0.1053 | 0.0898 | 0.0960 | 0.0809 | 0.0565 | 0.0751 |
| | 100 | 0.1034 | 0.0876 | 0.0955 | 0.0795 | 0.0564 | 0.0732 |

通过比较表 1 中各分类下前 10%、前 25%、前 50% 与 100% 的特征点的一阶中心绝对矩,可知 SURF 特征值的各个维度的一阶中心绝对矩具有一定稳定性。即各个维度的数据均符合一定的分布规律。由此,可认为 SURF 特征值的各维

均符合随机变量的特征,且其分布可由已知数据确定。

令第*i*分类中的SURF特征值的第*k*维为一随机变量 $surf_k$,其概率密度函数 $f_i(surf_k)$,其中 $f_i(surf_k)$ 可由同分类已知部分的数据近似获得。将待分类图片 I_{next} 表示为由*n*个SURF构成的集合 S_{next} , $S_{\text{next}} = \{s_1, s_2, \dots, s_n\}$,其中 s_p 表示第*p*个SURF向量。设 d_{jk} 为第*p*个SURF向量的第*k*维, Q_k 为 S_{next} 集合中所有元素第*k*维数值的集合,则 $Q_k = \{d_{1k}, d_{2k}, \dots, d_{nk}\}$,则 Q_k 符合第*i*分类第*k*维的概率密度函数 $f_i(surf_k)$ 。

由于 Q_k 对应的随机变量 $surf_k$ 必然具有收敛性,因此 Q_k 的一阶中心绝对矩作为一数学期望值为一定值,且在*n*值较大时可以以平均值近似其期望值。

图2显示为建筑、巴士、大象、花、食物、马等6个分类间,由各分类全部特征点数据获取的第41~46维的一阶中心绝对矩的对比。

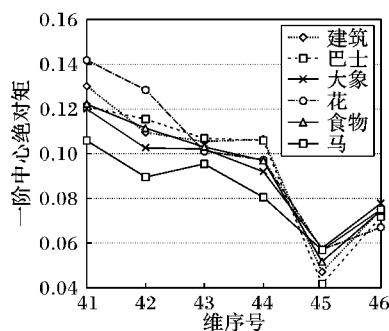


图2 各分类各维数据的一阶中心绝对矩的对比

由图2可知,各分类间存在一定的差异,且在部分维度上有较大差异。如果各分类间数据具有显著的差异,如数据中心相差较大且数据分布较为集中,则其分类效果会较好。

2.3 SURF特征值数据与拉普拉斯响应的统计特性

用于加速匹配的Hessian矩阵的迹的符号表明了亮度变化的整体方向,即黑背景下的白斑与白背景下的黑斑,因此通常被称为拉普拉斯正/负响应。在SURF特征点进行匹配时,只有Hessian矩阵的迹的符号相同的特征点才有可能匹配。同样,在特征点中,如果特征点产生的原因是相同的,即同为某些特定纹理或边缘,则其Hessian矩阵的迹的符号也有较大的可能是相同的。

通过特征数据对拉普拉斯正负响应训练朴素贝叶斯分类器,记录其查准率,以此作为拉普拉斯正负响应与两类数据之间关系强弱的表征,可得表2(仅显示其中6类)。

表2 各分类以特征值数据对拉普拉斯响应进行分类的查准率

| 分类 | 查准率 | 分类 | 查准率 | 分类 | 查准率 |
|----|-------|----|-------|----|-------|
| 建筑 | 0.828 | 大象 | 0.850 | 食物 | 0.826 |
| 巴士 | 0.827 | 花 | 0.827 | 马 | 0.848 |

通过对表2中数据可知,所有类别都具有较高的查准率,即拉普拉斯正负响应可以较明确地区分两类性质不同的数据,因此按照该符号将SURF向量集合分成两部分统计有助于更好地利用SURF特征值信息。

2.4 SURF特征点尺度信息的统计特性

SURF特征点的尺度信息同样是重要信息。通过对SURF特征点的尺度信息按前10%、25%、50%及100%数据规模统计其一阶中心绝对矩,可得表3(仅显示其中6类)。

由表3可知,不同数据规模下,不同类别的一阶中心绝对

矩同样具有和SURF特征值各维度相似的性质。即类别内不同数据规模的一阶中心绝对矩具有一定的稳定性,类别间具有一定差异。以上性质可以用于分类,但如尺度信息作为特征的单独一维,则其信息可能在分类器训练过程中为维数较高的其他特征淹没。

表3 不同数据规模下SURF特征点尺度一阶中心绝对矩的比较

| 数据规模/% | 建筑 | 巴士 | 大象 | 花 | 食物 | 马 |
|--------|--------|--------|--------|--------|--------|--------|
| 10 | 7.7312 | 7.6118 | 6.6011 | 8.3670 | 6.9337 | 6.9735 |
| 25 | 7.5234 | 7.4829 | 6.8436 | 8.2021 | 7.1331 | 7.2253 |
| 50 | 7.3756 | 7.2627 | 6.9524 | 8.1184 | 7.2528 | 7.2096 |
| 100 | 7.4334 | 7.1957 | 6.9614 | 8.1072 | 6.9337 | 7.1599 |

基于以上分析,可认为尺度信息是一随机变量。设SURF特征值某维度对应随机变量为 X ,尺度信息对应随机变量为 Y ,则 X, Y 独立,且 X, Y 具有数学期望。

设随机变量 $Z = XY$,则由 X, Y 独立,得 $E(Z) = E(X)E(Y)$,即 Z 具有数学期望,一阶中心绝对矩。由辛钦大数定律可知,随机变量 Z 在一定数量独立实验中所获得的数据的一阶中心绝对矩依概率收敛于随机变量 Z 的一阶中心绝对矩。

因此可将尺度信息作为权重信息,组成带权一阶中心绝对矩,并用于分类工作。

2.5 SURF-M特征计算方法

基于以上分析,本文提出采用一阶中心绝对矩和带权中心矩表示SURF信息中各维的统计特征,并将获取的统计特征用于图像分类工作。以下称该特征为SURF-M特征。

设图像中共提取出*n*个SURF特征点。对每个特征点取SURF-128特征值。设图像中所有SURF特征点的特征值的第*k*维组成的数列为 S_k ,设第*i*个SURF特征点的Hessian矩阵的迹的符号为 h_i , S_k^+ 是由 $h_i > 0$ 的特征点的第*k*维组成的数列, S_k^- 是由 $h_i < 0$ 的特征点的第*k*维组成的数列。由所有SURF特征点的特征值的尺度信息组成的数列为 K 。

设第*k*维的SURF特征点的一阶中心绝对矩为 $M(k)$:

$$M(k) = E(|S_k - E(S_k)|) \quad (6)$$

设第*k*维的SURF特征点中Hessian矩阵的迹的符号为正的特征点的一阶中心绝对矩为 $M^+(k)$:

$$M^+(k) = E(|S_k^+ - E(S_k^+)|) \quad (7)$$

设Hessian矩阵的迹的符号为负的特征点的特征值中第*k*维的一阶中心绝对矩为 $M^-(k)$:

$$M^-(k) = E(|S_k^- - E(S_k^-)|) \quad (8)$$

设第*k*维带权一阶中心绝对矩为 $M_w(k)$:

$$M_w(k) = E(|S_k K - E(S_k K)|) \quad (9)$$

其中,式(6)~(9)中使用了数学期望,但在实际应用中以算数平均值近似。最终将由各维计算得到的特征值组织为向量,即构成SURF-M特征。

2.6 SURF-M特征与其他特征融合的方法

在图像分类任务中,特征值是对图像按特定方式进行描述而产生的数据,因此是对图像信息的某种近似和简化。通过适当的机器学习方法可以获取其中的规律,并将此规律用于图像分类工作。不同的特征描述的信息描述是对同一图像的不同近似。不同的特征之间,由于描述角度不同,有可能出现在某些描述下不相似的实例在另外一种描述下存在相似

性。将不同描述通过适当的方法组织在一起的方法即特征融合。特征融合在合适条件下,可以更充分地利用图像的近似信息,得到更好的分类效果。

特征融合可以分为 3 种:底层特征融合、中层特征融合和高层特征融合。底层特征融合是将不同的特征向量组合成为新的特征向量;中层特征融合是训练针对不同特征的分类器并通过线性组合形成新的分类器;高层特征融合根据不同分类器的分类结果所含有的决策信息构建。

高层特征融合方式需要利用分类器输出的分到不同类别的概率,通过结合不同分类器输出的决策信息,可获得目标被分到不同类别的概率,其中最大的,即为最终分类结果。

设数据集分为 n 类,共有分类器 r 个,对图像 I ,第 i 个分类器的预测结果为 $vR_i = \langle p_{i1}, p_{i2}, \dots, p_{i(n-1)}, p_{in} \rangle$,最终分类结果为 a ,则:

$$a = \arg \max \left\{ \sum_{i=1}^r vR_i \right\} \quad (10)$$

SURF 是一种局部特征,表示的是图像局部的纹理特性与结构特征,通过对图像的局部特征进行数学统计,SURF-M 特征具备一定的对图像的全局特性的描述能力。但 SURF 提取过程中存在灰度化过程,即其描述的信息中不含图像的色彩信息。色彩信息需要通过颜色特征描述。颜色特征是一种全局特征,具有旋转,平移不变性质,对尺度变化敏感度低,具有良好的稳定性。通过与颜色特征进行特征融合可以弥补 SURF-M 特征的缺点。本文采用常见的 72 维 HSV(Hue, Saturation, Value)颜色直方图作为颜色特征,其中 H、S、V 分段数分别取 6、4、3。

通过高层特征融合,对于局部特征丰富、结构稳定的图像,SURF-M 特征因为分类效果较好而获得较大的权重;而对于颜色特征鲜明,纹理特征较少或结构不稳定的图像,颜色特征会获得较大的权重。通过相互弥补,最终可以获得较高的正确率。

3 实验结果与分析

实验采用 COREL1K 图像库,该图像库包括恐龙、花、大

表 5 不同 SURF-M 特征组合的查准率对比

| 特征代码 | 恐龙 | 花 | 大象 | 巴士 | 海滩 | 建筑 | 马 | 人 | 食物 | 山峰 | 合计 |
|--------------|------|------|------|------|------|------|------|------|------|------|-------|
| SURF-M-512 | 0.96 | 0.88 | 0.72 | 0.88 | 0.68 | 0.60 | 0.88 | 0.6 | 0.72 | 0.64 | 0.756 |
| SURF-M-384-0 | 0.96 | 0.88 | 0.68 | 0.88 | 0.64 | 0.52 | 0.84 | 0.56 | 0.76 | 0.64 | 0.736 |
| SURF-M-384-1 | 0.96 | 0.88 | 0.64 | 0.88 | 0.68 | 0.56 | 0.88 | 0.52 | 0.68 | 0.56 | 0.724 |
| SURF-M-256-1 | 0.96 | 0.84 | 0.60 | 0.92 | 0.72 | 0.44 | 0.92 | 0.4 | 0.76 | 0.52 | 0.708 |
| SURF-M-128-0 | 0.96 | 0.92 | 0.76 | 0.84 | 0.56 | 0.48 | 0.84 | 0.60 | 0.48 | 0.60 | 0.704 |
| SURF-M-128-2 | 0.96 | 0.84 | 0.68 | 0.84 | 0.68 | 0.52 | 0.92 | 0.48 | 0.56 | 0.56 | 0.704 |
| SURF-M-256-0 | 0.96 | 0.88 | 0.72 | 0.84 | 0.52 | 0.40 | 0.88 | 0.60 | 0.56 | 0.56 | 0.692 |
| SURF-M-128-3 | 0.92 | 0.84 | 0.68 | 0.88 | 0.48 | 0.32 | 0.92 | 0.44 | 0.72 | 0.48 | 0.668 |
| SURF-M-128-1 | 0.96 | 0.84 | 0.64 | 0.80 | 0.60 | 0.28 | 0.72 | 0.60 | 0.48 | 0.44 | 0.636 |

如表 5 所示,各种特征单独使用均有一定效果,其中带权重的中心绝对矩在单独使用时反而使分类效果比平均水平有所降低,通过观察实验数据,由纹理原因产生的 SURF 特征点的尺度信息与纹理特性等信息密切相关,因此带权一阶中心绝对矩特征具有一定稳定性;但由于不同物体交汇等原因产生的 SURF 特征点的尺度变化较大,且通常远大于由纹理原因产生的特征点的尺度,这种特征点会对所有维度的带权一阶

象、巴士、海滩、建筑、马、人、食物、山峰 10 个类别,每类 100 张图片,共 1000 张图片。采用 LIBSVM 进行分类。分类器在训练之前均经过数据缩放。

3.1 实验设置

为测试本文提出特征的有效性并测试其适用性,在相同的条件下,分别进行 3 组实验。

第 1 组实验采用 10 个分类,将 COREL 1K 图像库按 3:1 比例分为训练集和测试集,分别测试各种 SURF-M 特征,验证以上设想,并寻找最佳组合。

第 2 组实验采用 10 个分类,将 COREL 1K 图像库按 1:1 比例分为训练集和测试集。对本文方法和 SURF 直方图方法进行比较。两者均使用数学方法对原始特征维数不等的描述信息进行了处理,因此具有良好的可比性。同时以常见的 HSV 直方图方法和 Gabor 纹理描述方法作为参考。

第 3 组实验采用 10 个分类,将 COREL 1K 图像库按 1:1 比例分为训练集和测试集。使用 SURF-M 特征、SURF 直方图特征和 Gabor 纹理特征,与 HSV 直方图特征作高级特征融合,同时以一种基于词袋模型的方法作为参考。

3.2 分类性能比较

3.2.1 第 1 组实验

将 COREL 1K 图像库按 3:1 比例分为训练集和测试集。将以上各种公式计算得出的特征值单独或组合成特征向量,利用 SVM 分类器进行分类,得到各类图像的分类概率。不同特征组合命名如表 4 所示,其中组合一栏中的序号表示其采用的计算公式。该实验采用查准率作为评价标准,结果已按查准率降序排序,如表 5 所示。

表 4 特征代码与特征组合对照

| 特征代码 | 组合方式 | 特征代码 | 组合方式 |
|--------------|----------|--------------|------------------|
| SURF-M-128-0 | 式(6) | SURF-M-256-1 | 式(7)、(8) |
| SURF-M-128-1 | 式(9) | SURF-M-384-0 | 式(6)、(7)、(8) |
| SURF-M-128-2 | 式(7) | SURF-M-384-1 | 式(7)、(8)、(9) |
| SURF-M-128-3 | 式(8) | SURF-M-512 | 式(6)、(7)、(8)、(9) |
| SURF-M-256-0 | 式(6)、(9) | | |

中心绝对矩产生较大影响,因而部分图片的 SURF-M-128-1 特征与同类图片相差较大,导致使用该特征进行分类时效果降低。组合的方法除 SURF-M-256-0 特征外,均相对于单组特征的方法有一定提高,而使用了全部 4 组信息的 SURF-M-512 特征正确率最高。这与 SVM 分类器的分类原理有一定关系。

通过比较可知,对于局部纹理信息表现较为一致的恐龙、

花、大象、马等分类,本文方法分类效果较好,这是由于这些分类下的图片结构基本一致,图片中物体的结构与纹理信息较为相似,则各图像 SURF 特征点对应特征值的集合也更为相似,由此而获取的 SURF-M 特征更为集中,这对于 SVM 分类器是较为有利的条件。对于海滩、建筑、人、食物这 4 种分类的分类效果较差。这是由于图库中此 4 类图片本身较为复杂,有部分图片在图像构成上与同类图片有较大差异,如海滩分类下存在部分图像沙滩、岩石、人、树、天空等与海滩面不同的物体产生的 SURF 特征点占据特征值集合中主要部分的情况。这都导致由这些图像提取出的 SURF 特征点的构成与同

类图片的平均水平具有较大的差异,从而导致错误分类。这表明,本文方法会受到不具有代表性样本问题的影响。根据第 1 组实验的结论,以下各组实验中所采用的 SURF-M 特征均为 SURF-M-512。

3.2.2 第 2 组实验

将 COREL 1K 图像库按 1:1 比例分为训练集和测试集。分别提取图像的 SURF-M 特征、SURF 直方图特征、Gabor 三通道纹理特征^[4]、HSV 直方图特征利用 SVM 分类器进行分类,得到各类图像的分类概率。该实验采用查准率作为评价标准,结果如表 6 所示。

表 6 各种方法的查准率对比

| 分类方法 | 恐龙 | 花 | 大象 | 巴士 | 海滩 | 建筑 | 马 | 人 | 食物 | 山峰 | 平均 |
|-----------|------|------|------|------|------|------|------|------|------|------|-------|
| SURF-M | 0.98 | 0.94 | 0.74 | 0.94 | 0.60 | 0.50 | 0.80 | 0.58 | 0.70 | 0.60 | 0.738 |
| SURF 直方图 | 0.98 | 0.68 | 0.62 | 0.84 | 0.38 | 0.28 | 0.70 | 0.56 | 0.14 | 0.44 | 0.562 |
| Gabor 三通道 | 0.98 | 0.86 | 0.60 | 0.72 | 0.52 | 0.40 | 0.90 | 0.70 | 0.72 | 0.44 | 0.684 |
| HSV 直方图 | 0.90 | 0.86 | 0.82 | 0.86 | 0.66 | 0.66 | 0.96 | 0.78 | 0.70 | 0.46 | 0.766 |

从表 6 中可以看出,除人分类以外,本文提出的 SURF-M 分类算法相比纹理类分类算法都有不同程度提高。与 SURF 直方图方法相比,本文方法在大多数分类上具有优势,整体分类效果高 17.6%。对于经过 Gabor 滤波器处理后获得的纹理特征,尽管相对于通常的纹理特征表述更加清晰明确,但作为全局特征,该类方法在描述纹理较为复杂、变化较多的图像时,效果不佳。与 HSV 直方图方法相比,两者各有优势,SURF-M 方法对于人、马等分类的分类效果与 HSV 直方图方法相差较大,前者是因为在“人”分类内图片不仅在图像构成

上有差异,且纹理变化也较大,而在颜色分布上一致性稍好;后者则是因为“马”分类内图片在颜色分布上一致性更好。

3.2.3 第 3 组实验

分别提取图像的 SURF-M 特征、SURF 直方图特征、Gabor 三通道纹理特征、HSV 直方图特征利用 SVM 分类器进行分类,得到每张测试图像的分类概率,并分别将 SURF-M 特征、SURF 直方图特征、Gabor 三通道特征与 HSV 直方图特征作高级特征融合。此外以文献[1]中方法在相同实验条件下所得结果作为参考。该实验采用查准率作为评价标准,结果如表 7 所示。

表 7 纹理特征与颜色特征高层特征融合方法的查准率对比

| 分类方法 | 恐龙 | 花 | 大象 | 巴士 | 海滩 | 建筑 | 马 | 人 | 食物 | 山峰 | 平均 |
|---------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| SURF-M + HSV 直方图 | 0.920 | 0.980 | 0.860 | 0.940 | 0.760 | 0.760 | 0.980 | 0.860 | 0.900 | 0.680 | 0.868 |
| SURF 直方图 + HSV 直方图 | 0.980 | 0.900 | 0.780 | 0.920 | 0.760 | 0.620 | 0.960 | 0.840 | 0.800 | 0.600 | 0.816 |
| Gabor 三通道 + HSV 直方图 | 0.980 | 0.900 | 0.620 | 0.880 | 0.620 | 0.700 | 0.980 | 0.800 | 0.860 | 0.660 | 0.800 |
| 包示例学习(参考数据) | 0.990 | 0.898 | 0.764 | 0.925 | 0.769 | 0.692 | 0.932 | 0.780 | 0.884 | 0.727 | 0.836 |

通过对比表 6 与表 7 可知,通过将纹理特征与颜色特征进行高层特征融合,各种方法的分类效果均有部分提升。其中,SURF-M 特征结合 72 维 HSV 直方图的方法分别较两种单独的方法提高了 13.0% 和 10.2%。这说明高级特征融合的效果提升较为明显。同时,由于在描述复杂图像纹理特征方面本文方法优于 SURF 直方图方法和 Gabor 三通道方法,因此在总体效果上优于其他两种融合特征。

通过将高级特征融合方法与文献[1]中方法进行比较可知,高级特征融合方法通过结合不同分类器的分类决策,可实现与基于词袋模型的方法相当的分类效果。词袋模型的主要优点是其视觉词汇是针对图像中的某些特定局部区域的,因而可以通过建立概率模型等方法,对含有特定视觉词汇的图像进行分类。因此文献[1]中方法对山峰、海滩等部分同类图像之间纹理变化较大的分类有相对较好的分类效果。而本文方法基于 SURF,其对图像的纹理和图像构成的描述更具有针对性,因而降低了词袋模型对本方法的优势。而融合方法不经过分块、按特征向量聚类、建立概率模型等容易引入错误或放大误差的过程,因而对于纹理或颜色特征稳定、与其他类别有一定差异的类别,如花、大象、人等,效果优于部分基于

词袋的方法。最终,本文方法查准率较文献[1]中方法提高了 3.2%。

4 结语

本文提出一种新的利用 SURF 进行分类的方法,将难以直接利用的描述局部特征的 SURF 向量集合通过统计方法转换为能够表示图片全局特性的 SURF-M 特征。该特征可以输入 SVM 等分类器进行分类。通过与其他特征进行比较,本文方法对图片构成结构和纹理相似的图片具有良好的分类效果。通过与其他适当的特征进行特征融合,可获得更好的分类效果。

作为一种使用统计特性的特征,SURF-M 特征在使用过程中会遇到由个别图片获得的样本不具有代表性的问题。如在抽样样本较小时,个别过大或过小的数据可能对统计特征产生较大的影响,则剔除小部分数据有助于提高分类准确率。在今后的工作中,将尝试引入点估计等数学方法,增强 SURF-M 特征的鲁棒性。

参考文献:

- [1] WANG K, YANG Y, WANG X. Automatic classification of multi-

- ple-instance image based on the bag space [J]. Journal of Image and Graphics, 2013, 18(9): 1093 – 1100. (王科平, 杨艺, 王新良. 包空间多示例图像自动分类[J]. 中国图象图形学报, 2013, 18(9): 1093 – 1100.)
- [2] SINGARAJU D, VIDAL R. Using global bag of features models in random fields for joint categorization and segmentation of objects [C]// CVPR 2011: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2011: 2313 – 2319.
- [3] HE D, QIAO Y, LI P, et al. Weed recognition based on SVM-DS multi-feature fusion [J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(2): 182 – 187. (何东健, 乔永亮, 李攀, 等. 基于 SVM-DS 多特征融合的杂草识别[J]. 农业机械学报, 2013, 44(2): 182 – 187.)
- [4] LI S. Image classification based on the low-level feature and SVM [D]. Kaifeng: Henan University, 2013. (李顺才. 基于底层特征和 SVM 的图像分类[D]. 开封: 河南大学, 2013.)
- [5] LI F, PERONA P. A Bayesian hierarchical model for learning natural scene categories [C]//CVPR 2005: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2005: 524 – 531.
- [6] GRAUMAN K, DARRELL T. The pyramid match kernel: discriminative classification with sets of image features [C]// ICCV 2005: Proceedings of the 10th IEEE International Conference on Computer Vision. Piscataway: IEEE, 2005: 1458 – 1465.
- [7] van GEMERT J C, VEENMAN C J, SMEULDERS A W M, et al. Visual word ambiguity [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(7): 1271 – 1283.
- [8] ZHOU L. Research on key technologies of scene classification and object recognition [D]. Changsha : National University of Defense Technology, 2012. (周莉. 自然场景分类与目标识别关键技术研究[D]. 长沙: 国防科学技术大学, 2012.)
- [9] JIANG Y, WANG R, WANG C. Scene classification with context pyramid features [J]. Journal of Computer-Aided Design & Computer Graphics, 2010, 22(8): 1366 – 1373. (江悦, 王润生, 王程. 采
- 用上下文金字塔特征的场景分类[J]. 计算机辅助设计与图形学报, 2010, 22(8): 1366 – 1373.)
- [10] HAN B, YANG C, GAO X. Aurora image classification based on LDA combining with saliency information [J]. Journal of Software, 2013, 24(11): 2758 – 2766. (韩冰, 杨辰, 高新波. 融合显著信息的 LDA 极光图像分类[J]. 软件学报, 2013, 24(11): 2758 – 2766.)
- [11] BAY H, TUYTELAARS T, van GOOL L. SURF: speeded up robust features [C]// ECCV 2006: Proceedings of the 9th European Conference on Computer Vision. Berlin: Springer, 2006: 404 – 417.
- [12] LIANG J, LIU Q, AI Q. Research of image classification based on fusion of SURF and global feature [J]. Computer Engineering and Applications, 2013, 49(17): 174 – 177. (梁进, 刘泉, 艾青松. 基于 SURF 和全局特征融合的图像分类研究[J]. 计算机工程与应用, 2013, 49(17): 174 – 177.)
- [13] CHANG C, LIN C. LIBSVM: a library for support vector machines [J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2(3): Article No. 27.
- [14] SIMARD P Y, BOTTOU L, HAFFNER P, et al. Boxlets: a fast convolution algorithm for signal processing and neural networks [C]// Proceedings of the 1998 Conference on Advances in Neural Information Processing Systems. Cambridge: MIT Press, 1999: 571 – 577.
- [15] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features [C]// CVPR 2001: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2001: I-511 – I-518.
- [16] DONG W, WANG Z, CHARIKAR M, et al. Efficiently matching sets of features with random histograms [C]// Proceedings of the 16th ACM International Conference on Multimedia. New York: ACM, 2008: 179 – 188.
- [17] CHARIKAR M S. Similarity estimation techniques from rounding algorithms [C]// Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing. New York: ACM, 2002: 380 – 388.

(上接第 223 页)

- [7] ZHANG Z. Determining the epipolar geometry and its uncertainty: a review [J]. International Journal of Computer Vision, 1998, 27 (2): 161 – 195.
- [8] XU Y, ZHOU Q, CONG L, et al. High-speed simultaneous image distortion correction transformations for a multi-camera cylindrical panorama real-time video system using FPGA [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2014, 24(6): 1061 – 1069.
- [9] HARTLEY R, ZISSELMAN A. Multiple view geometry in computer vision [M]. WEI H, YANG S, ZHANG Q, et al. translated. Hefei: Anhui University Press, 2002: 291 – 300. (HARTLEY R, ZISSELMAN A. 计算机视觉中的多视图几何学[M]. 韦穗, 杨尚骏, 章权宾, 等译. 合肥: 安徽大学出版社, 2002: 291 – 300.)
- [10] WANG P, XU Y. Photogrammetry [M]. Wuhan: Wuhan University Press, 2005: 109 – 112. (王佩军, 徐亚明. 摄影测量学[M]. 武汉: 武汉大学出版社, 2005: 109 – 112.)
- [11] BRADSKI G, KAEBLER A. Learning OpenCV: computer vision
- with the OpenCV library [M]. Nanjing: SouthEast University Press, 2009: 15 – 21.
- [12] CAO J, LI K, GAO C, et al. Application of local features in aerial image mosaic [J]. Journal of University of Electronic Science and Technology of China, 2013, 42(1): 125 – 129. (曹健, 李侃, 高春晓, 等. 局部特征在航拍图像拼接中的应用[J]. 电子科技大学学报, 2013, 42(1): 125 – 129.)
- [13] SONG F, LU B. An automatic video image mosaic algorithm based on SIFT feature matching [C]// Proceedings of the 2012 International Conference on Communication, Electronics and Automation Engineering Advances in Intelligent Systems and Computing. Berlin: Springer, 2013: 879 – 886.
- [14] JIANG H, ZHANG L. A color correction algorithm for stereo video coding based on histogram matching [J]. Information and Communications Technologies, 2009, 3(5): 75 – 78. (姜浩, 张蕾. 一种基于直方图匹配的立体视频颜色校正算法[J]. 信息通信技术, 2009, 3(5): 75 – 78.)