

基于 Sift 特征的商标检索算法

林传力, 赵宇明

(上海交通大学图像处理与模式识别研究所, 上海 200240)

摘 要: 基于内容的商标检索利用全局形状特征和子图像的空间位置关系进行特征匹配, 存在检索精度不高、抗噪声能力不强和相似子图像可能发生误检、漏检等不足。该文提出一种利用 Sift 特征进行商标检索的新方法。实验结果表明, 与现有检索方式相比, 该方法准确全面, 并对遮挡、扭曲、噪声等干扰具有更强的抵抗力。

关键词: 商标检索; 基于内容的图像检索; Sift 算法; 查准率-查全率

Trademark Retrieval Algorithm Based on Sift Feature

LIN Chuan-li, ZHAO Yu-ming

(Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai 200240)

【Abstract】 Content based trademark retrieval algorithms use overall shape features or spatial relationships between components to match with features stored in database. However they're inadequate in anti-noise and always suffer premission or false matches when handling trademarks with similar components. This paper proposes a new retrieval method using the Sift descriptor. Experimental results show that this method is more accurate, more complete in retrieving similar trademarks and is more robust against noise compared with existing methods.

【Key words】 trademark retrieval; Content Based Image Retrieval(CBIR); Sift; Recall-Precision criterion

1 概述

商标是商品或服务的标识,是企业信誉和可靠度的象征。新商标必须具有足够的独特性以避免与已注册的商标混淆或冲突^[1]。因此,高效、准确地从商标库中检索出潜在的近似商标成为一个重要研究方向。传统方式是采用Vienna分类码,但存在人工分配分类码过于繁重、无法对抽象图像给出有意义的分类等缺点。

基于内容的商标检索(Content Based Image Retrieval, CBIR)方法可分为 2 类:(1)利用全局或子图像的形状特征而未考虑空间关系。如Kim等利用基于区域的描述子Zernike矩进行检索^[2],此外还有利用Fourier描述子、Hu不变矩检索等。此类方法的问题在于它仅能识别出具有相似轮廓的图像,而忽略了内部结构,可能造成误检。(2)同时考虑形状特征和子图像的空间位置关系。如文献[3]先根据子图像相似性进行粗检索,然后利用位置字符串对空间结构进行分析。Eakins的Artisan系统提取图像的区域轮廓并归入不同的轮廓族(boundary family),然后从整张图像、轮廓族和单独轮廓 3 个层次提取形状特征^[4]。此方法性能优于第 1 类,但对具有高相似度的子图像的情况可能出现漏检的情况。如图 1,待查商标显然违背“独特性”的要求,应被检索出。

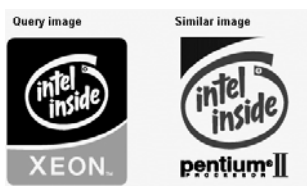


图 1 具有高度相似的子图像

本文提出利用 Sift 算子进行商标检索。该算法提取图像

中局部灰度变化突出的特征点,并将其相对方向、邻域的相对位置归结为 Sift 描述子。Sift 特征对位移、尺度、方向、微小的视角变化具有良好的不变性,且对图像轮廓也具有一定的描述能力,可用来查找相似商标。

2 Sift 算法回顾

Sift是David Lowe于 2004 年提出的局部特征描述子^[5]。独特性好,信息量丰富,并对多数变换具有很强的不变性,在同类描述子中具有最强的健壮性^[6]。

为达成尺度不变性,Sift使用高斯差分尺度空间(DoG),它由相邻尺度的高斯差分核与输入图像卷积而成。DoG不仅是对归一化的LoG($\sigma^2 - 2G$)的线性近似,而且简化了尺度空间的计算。

Sift 特征的提取有如下 4 个步骤:

(1)尺度空间极值点检测。对输入图像先进行增量式的高斯卷积以建立 DoG 空间,然后在 3 个相邻尺度的邻域中搜索极值点,初步得到特征点的位置;

(2)对特征点的精确定位。利用 DoG 的二次 Taylor 展开式 $D(x)$ 插值得到特征点位置、尺度的亚像素值并滤除低对比度点和边缘响应较强的不稳定点;

(3)为特征点分配方向值,达到方向不变性。创建梯度方向直方图,选择主峰值作为特征点的方,达到主峰值 80% 以上的局部峰值视为辅助方向,以增强特征的健壮性;

(4)生成特征描述子。将特征点 16×16 邻域中各点的坐标按高斯加权归入 4×4 的位置网格,将其与特征点的相对方向加权得到 8-格方向直方图,从而获得 128 维的描述子。

作者简介: 林传力(1984 -),男,硕士研究生,主研方向:图像处理,二维图像匹配;赵宇明,副教授

收稿日期: 2008-04-15 **E-mail:** nic0306@163.com

3 Sift 算子在商标检索中的应用

Sift 算子被广泛应用于目标识别、机器人定位等场合，Sivic 和 Zisserman 也将其应用到基于内容的视频检索系统 Video Google，为 Sift 的应用和发展开辟了更广阔的空间。

商标检索的目的可归结为：给定一个待查商标，确定它是否和商标库中任何一个足够相似以至造成混淆。利用 Sift 特征，可准确查找出数据库中潜在的相似特征，并具有很强的健壮性，适用于商标检索。相对于目标检测和通用的 CBIR，商标检索和 Sift 特征都有自身的特殊性：

(1) 应对商标图像进行降噪处理，以避免产生无意义的特征，对检索的速度和准确度造成干扰。

(2) 特征点的数量对检索也有很大的影响：特征点过多不仅会使检索速度变慢而且也不必要；特征点过少又容易造成漏检。因此，构建商标数据库时应对图像产生的特征点数量进行归一控制。

在文献[5]推荐的参数下，一张 150×150 像素的普通商标图像可产生约 100 个稳定的特征点。不同尺寸、不同类型的图像产生特征点的数量也不同：同样内容的图像，尺寸越大特征点越多；大小相同的图像，纹理越复杂特征点个数越多，而以图形为主的图像则只在边缘和转角产生特征点。因此，在建立特征数据库阶段，应首先对图像尺寸进行统一调整，然后按复杂度进行分类，不同的类别使用不同的 Sift 参数，以保证数据库中所有图像具有稳定的、数量大体一致的特征。

(3) 商标数据库很庞大，国家商标局现有的商标数量已超过 150 万，本文实验中采用的商标库也包含 2 200 余个样本。同时 Sift 特征又属于高维特征，因此，检索所需的计算量也成为应考虑的问题。本文在检索阶段采用近似的最近邻搜索算法(Best Bin Search, BBF)^[7]。使用该算法可在不显著降低匹配精度的情况下有效提高检索速度。

本文构建的商标检索系统框架如图 2 所示。

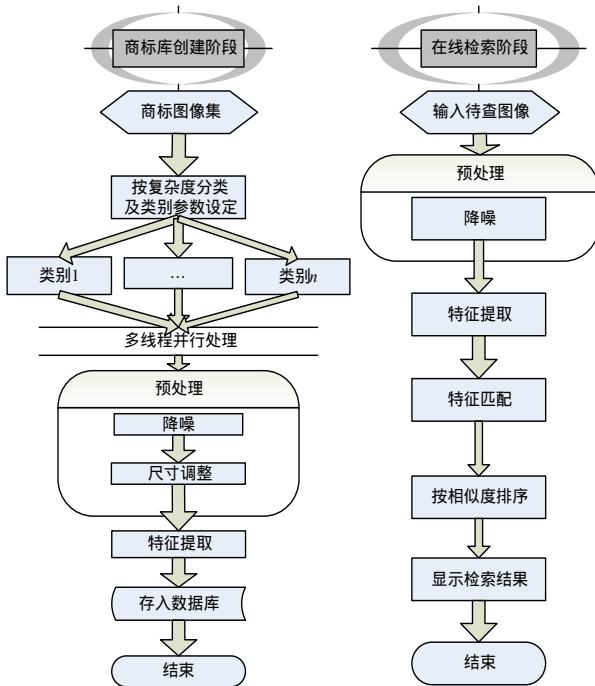


图 2 商标检索系统框架

3.1 匹配与相似性度量

设检索图像为 A ，商标库中的任一图像为 B ，从 A 中提取

的 Sift 特征点集合为 $F_a = \{f_1^{(a)}, f_2^{(a)}, \dots, f_{N_a}^{(a)}\}$ ， B 的特征集合为 $F_b = \{f_1^{(b)}, f_2^{(b)}, \dots, f_{N_b}^{(b)}\}$ ，其中， N_a, N_b 为 A, B 的特征点个数。以 Euclidean 距离作为 Sift 描述子的距离函数，其公式可表示为

$$d(f_1, f_2) = \sqrt{\sum_{i=1}^n (f_1^{(i)} - f_2^{(i)})^2} \quad (1)$$

将最近邻距离 d_1 与次近邻距离 d_2 的比率和预先设定的阈值比较以确定是否匹配到特征点^[5]。具体匹配过程如下：

(1) 用 KD-Tree 为 F_b 的全部元素建立索引；

(2) 根据式(1)，使用 BBF 搜索算法得到 F_a 中每个元素 k_i 在 KD-Tree 中的近似 k 近邻 (k 取 2)。设返回的 2 个最近邻特征点为 f_1, f_2 ；

(3) 根据 d_1/d_2 确定 f_1 是否 k_i 的有效匹配；

(4) 对所有 k_i 重复上述过程，并统计在 F_b 中 $k_1-k_{N_b}$ 匹配到的所有特征点数 N 。

最后将图像 A 和图像 B 的相似度定义为 N ，作为对检索结果排序的依据。

3.2 系统性能评价

本文采用规格化的 Recall-Precision 准则^[4]作为系统性能评价标准。该准则是在传统的 Recall-Precision 准则中增加了对返回结果排序的考量，因此，对检索性能的衡量更加精确，可用于评价任何返回有序结果的检索系统。评价结果反映了本次检索的查全率和查准率，结果为 1 表示最优检索，为 0 则表示最差检索。其定义如下：

$$R_n = 1 - \frac{\sum_{i=1}^n R_i - \sum_{i=1}^n i}{n(N-n)} \quad (2)$$

$$P_n = 1 - \frac{\sum_{i=1}^n (\lg R_i) - \sum_{i=1}^n (\lg i)}{\lg \left(\frac{N!}{(N-n)!n!} \right)} \quad (3)$$

其中， R_n, P_n 分别是查全率和查准率； R_i 是第 i 个相关图像在检索结果中的排序； n 是所有相关图像的个数； N 是整个商标数据库的规模。

4 实验结果与分析

4.1 实验设计

实验选用 24 张商标作为检索图像(query image)，编号为 Q_1, Q_2, \dots, Q_{24} ，并为每张设计了 8 个~12 个基准图像。实验采用 4 种方式生成基准图像，并据此将检索划分为 4 类，包括：

(1) 将 $Q_1 \sim Q_8$ 进行尺度、旋转、视角、镜像变换(及其组合)，此类检索记作类 A ；

(2) 将 $Q_9 \sim Q_{16}$ 进行遮挡、扭曲、噪声等形变或干扰(及其组合)，记作类 B ；

(3) 将 $Q_{17} \sim Q_{24}$ 的子图像与其它无关内容进行组合，记作类 C 。

商标库包含无关商标 2 000 张及基准图像 200 张，总数为 2 200 张。实验的置信水平设为 0.01，即每次返回最相似的 $2\ 200 \times 0.01 = 22$ 个商标，并按相似度将商标由高到低排序。

4.2 结果与分析

图 3 是用 Sift 描述子、子图特征结合空间关系、Hu 不变矩 4 种方法对 A 类 Q_4 的检索结果。可见，Sift 描述子返回了全部正确的基准图像，且都排在最前面，而其他 2 种方法均存在错误的返回图像，说明 Sift 算子对尺度、旋转、视角等变化(及其组合)的不变性很强，对此类相似商标具有最佳的检索能力。

图 4 是 4 种方法对 B 类 Q_9 的检索结果, Sift 描述子同样得到完全正确的结果, 表明它对遮挡、扭曲、噪声干扰等具有很强的健壮性。子图特征结合空间关系方法对 B 类的检索能力略优于 Hu 不变矩, 但二者都对遮挡没有不变性。

图 5 是 C 类 Q_{21} 的检索结果。Sift 算子取得了更高的查准率和查全率, 而 Hu 不变矩由于描述的仅仅是全局形状, 对此类检索能力很差。C 类检索反映了对相似子图像的检索能力。

上述 3 种方法的 3 类检索平均 R_n, P_n 如表 1 所示。所有检索图像的查全率/查准率(PVR)曲线如图 6 所示。可见, Sift 方法的 PVR 指数高于另外 2 种, 其整体检索性能也更高。



图 3 A 类 Q_4 检索结果

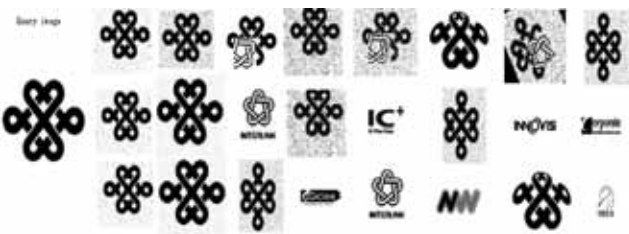


图 4 B 类 Q_9 检索结果



图 5 C 类 Q_{21} 检索的返回结果

表 1 各方法的 3 类检索平均 R_n, P_n

方法类别	R_n			P_n		
	M_1	M_2	M_3	M_1	M_2	M_3
A	0.98	0.88	0.85	0.98	0.84	0.79
B	0.94	0.68	0.63	0.90	0.61	0.42
C	0.87	0.61	0.36	0.77	0.48	0.27

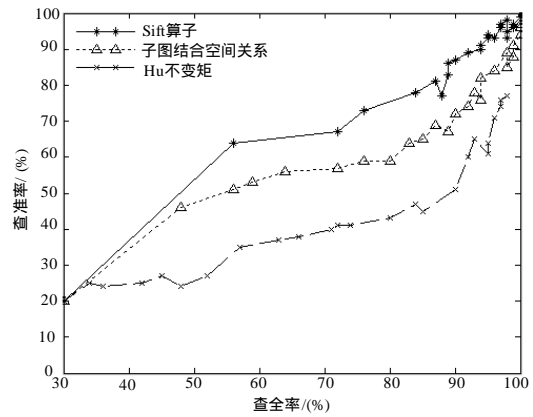


图 6 3 种方法的 PVR 指数曲线

5 结束语

本文提出一种利用 Sift 描述子进行商标图像检索算法, 它对于相似的图像或子图像具有良好的识别能力, 对尺度、镜像、旋转等具有很好的不变性, 表现出很强的抗遮挡和噪声的能力。实验结果表明 Sift 描述子的表现优于现有的 2 类方法。Sift 描述子的不足之处在于它对图像的轮廓描述能力不够强, 可能造成对某些商标的漏检, 影响查全率。对此可将其与形状特征融合, 可更符合人们对标识相似性的判断。下一步研究方向是准确有效地使 Sift 特征与形状特征协同。

参考文献

- [1] 中国商标网. 中华人民共和国商标法 [Z]. (2001-10-27). <http://sbj.saic.gov.cn>.
- [2] Kim Y S, Kim W Y. Content-based Trademark Retrieval System Using Visually Salient Feature[J]. Image Vision Compute, 1998, 16(12): 931-939.
- [3] 郭丽, 黄元元, 杨静宇. 基于形状和空间结构的商标图像检索方法[J]. 计算机软件与应用, 2005, 22(1): 93-95.
- [4] Eakins J P, Shields K, Boardman J M. ARTISAN——A Shape Retrieval System Based on Boundary Family Indexing[J]. Storage and Retrieval for Image and Video Databases, 1996, 31(4): 73-80.
- [5] Lowe D G. Distinctive Image Features from Scale-invariant Keypoints[J]. International Journal of Computer Vision, 2004, 59(1): 60.
- [6] Mikolajczyk K, Schmid C. A Performance Evaluation of Local Descriptors[C]//Proc. of International Conference on Computer Vision and Pattern Recognition. Madison, USA: [s. n.], 2003: 257-263.
- [7] Beis J S, Lowe D G. Shape Indexing Using Approximate Nearest-neighbor Search in High-dimensional Spaces[C]//Proc. of International Conference on Computer Vision and Pattern Recognition, Washington D. C. USA: [s. n.], 1997: 1000-1006.

(上接第 274 页)

参考文献

- [1] 梁爱虎. 精通 SOA: 基于服务总线的整合应用开发[M]. 北京: 电子工业出版社, 2007.
- [2] 侯战伟, 莫林. 基于 SOA 的数据库中间件的研究与设计[J]. 计算机应用研究, 2007, 24(6): 285-286.
- [3] 简斌, 左荣国, 闫光荣, 等. 基于 SOA 的中小制造企业应用集成系统研究[J]. 计算机工程, 2007, 33(5): 243-245.
- [4] 丁兆青, 董传良. 基于 SOA 的分布式应用集成研究[J]. 计算机工程, 2007, 33(10): 247-248.
- [5] 孙烽. 基于 SOA 的门户系统架构研究[J]. 软件导刊, 2007, (15): 35.