

DOI:10.3969/j.issn.1673-4785.201505009
网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20150630.1544.002.html>

基于模糊关联规则和决策树的图像自动标注

李志欣^{1,2}, 李灵芝¹, 张灿龙^{1,2}

(1.广西师范大学 广西多源信息挖掘与安全重点实验室, 广西 桂林 541004; 2.广西信息科学实验中心, 广西 桂林 541004)

摘 要:传统的基于关联规则算法的图像自动标注存在“锐利边界”问题,使分类存在模糊性、不准确性。且随着多媒体技术的飞速发展,图像信息数据迅速增长,海量的图像数据会形成大量冗余的关联规则,这将导致分类效率大大降低。针对这 2 个问题,文中提出基于模糊关联规则和决策树的图像自动标注模型。该模型首先获得关联训练图像低层特征和高层语义的模糊关联规则,再利用决策树方法删减冗余的模糊关联规则,基于决策树删减后的模糊关联规则,大大减小了算法的计算复杂度。实验在 Corel 5k 和 IAPR-TC12 两个基准数据集上进行,并从精度、召回率、F-measure 以及产生的规则数量几个度量措施上进行比较。与其他几种前沿的图像自动标注方法的结果对比表明,该方法在图像的标注精度和标注效率上有很大的提高。

关键词:锐利边界;模糊分类;图像自动标注;模糊关联规则;决策树

中图分类号: TP391 **文献标志码:** A **文章编号:** 1673-4785(2015)04-0636-08

中文引用格式:李志欣,李灵芝,张灿龙. 基于模糊关联规则和决策树的图像自动标注[J]. 智能系统学报, 2015, 10(4): 636-644.
英文引用格式:LI Zhixin, LI Lingzhi, ZHANG Canlong. Automatic image annotation based on fuzzy association rules and decision trees[J]. CAAI Transactions on Intelligent Systems, 2015, 10(4): 636-644.

Automatic image annotation based on fuzzy association rules and decision trees

LI Zhixin^{1,2}, LI Lingzhi¹, ZHANG Canlong^{1,2}

(1.Guangxi Key Lab of Multi-source Information Mining & Security, Guangxi Normal University, Guilin 541004, China; 2. Guangxi Experiment Center of Information Science, Guilin 541004, China)

Abstract: The traditional automatic image annotation based on association rules exists the problem of sharp boundary, which makes classification more fuzzy and inaccurate. Moreover, with the rapid development of multimedia technology, the size of image data increases quickly. Massive image data will produce a lot of redundant association rules, which greatly decreases the efficiency of image classification. In order to solve these two problems, this paper proposes an automatic image annotation approach based on fuzzy association rules and decision trees. The approach firstly obtains fuzzy association rules which represent the fuzzy correlations between low-level visual features and high-level semantic concepts of training images. Then, decision tree is adopted to reduce the redundant fuzzy association rules. As a result, computational complexity of the algorithm is decreased to a large degree. Experiments were done on Corel5k and IAPR-TC12 datasets. The evaluation measures are compared from the aspects of precision, recall, F-measure and the number of rules. The experimental results show that the proposed method acquires higher accuracy and efficiency in comparison with several state-of-the-art automatic image annotation approaches.

Keywords: sharp boundary; fuzzy classification; automatic image annotation; fuzzy association rules; decision tree

收稿日期: 2015-05-06. 网络出版日期: 2015-06-30.
基金项目: 国家自然科学基金资助项目(61165009, 61262005, 61363035, 61365009); 国家 973 计划资助项目(2012CB326403); 广西自然科学基金资助项目(2012GXNSFAA053219, 2013GXNSFAA019345, 2014GXNSFAA118368).
通信作者: 李志欣. E-mail: lizx@gxnu.edu.cn.

随着多媒体技术的飞速发展,图像信息数据迅速增长,传统的人工图像标注^[1-2]已不能满足海量的图像数据库标注要求,如何实现有效标注和快速存取,已经成为多媒体领域一项重大研究课题。基

于内容的图像检索 (content based image retrieval, CBIR),用低维的视觉特征如颜色、形状、纹理以及空间位置等来检索图像信息,但是 CBIR 系统存在低维特征与高维语义间的“语义鸿沟”,针对该问题,图像自动标注研究如何更加有效地标注和快速存取图像。






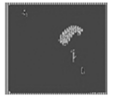

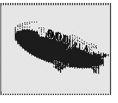
在多数图像自动标注 (automatic image annotation, AIA) 系统中,采用全局特征、基于块的局部特征或是基于区域的局部特征表示图像。Jeon 等^[3]使用区域分割方法,假设图像语义用区域特征产生的词汇来描述,结合以上 3 种特征表述方法,提出全局和局部的特征表示方法来标注图像。Wang 等^[4]提出了一种结合全局、区域及上下文特征表示的改进模型,通过计算它们的联合概率并结合以上 3 种特征表示标注图像。Duygulu 等^[5]提出翻译模型标注图像,该方法是一种生成式模型标注的代表性方法,它通过学习联合概率将关键词与图像的区域联系起来,将标注过程转化成一个将区域翻译为关键词的过程。Monay 等^[6]提出基于概率潜在语义分析 (probabilistic latent semantic analysis, PLSA) 模型的图像自动标注方法 PLSA-WORDS。李志欣等在此基础上先后提出融合语义主题的图像自动标注^[7]及混合生成式和判别式模型的图像自动标注模型 (hybrid generative/discriminative model, HGDM)^[8]。HGDM 首先用连续 PLSA 模型处理图像视觉特征,然后构建用来学习图像语义类别的分类模型链,综合神经网络、多类 SVM 以及 K 近邻分类器模型,利用 EM 算法计算一个权重参数,根据该参数来选择精确度更高的分类模型,最终得到目标的语义。茹立云等^[9]提出一种基于 boosting 学习的图像自动标注系统,假设一组具有同一语义的图像能够由一组视觉模型来表示,即颜色和纹理特殊组合的 2 维多分辨率隐马尔可夫模型,然后使用 boosting 算法实现关键词与模型的关联。Sumathi 和 Hemalatha 提出一种创新的混合分层图像标注模型^[10],该方法运用低维图像特征及其特征间距离找到图像的最近邻,然后用 SVM 方法得到图像标注。张静等^[11]提出一种新的模型提取图像前景和背景语义,运用视觉特点分析多个判别方法获得前景语义概念,用区域语义分析方法标注背景图像。Makadia 等^[12]提出用 JEC (joint equal contribution) 的方法进行图像标注, JEC 利用全局低维图像特征和基本距离度量的简单结合寻找给定图像的最近邻,然后使用一种贪心的标签传递机制将关键词赋予对应的图像,取得了很好的标注精度和检索性能。

但是,在以上标注模型中都没有考虑到图像处理中存在的模糊和不确定性的属性 (锐利边界),如图像边缘、边界、区域等定义,对于临界的点,将其确定为某一类都会影响标准的准确性,存在的这些模糊属性将干扰图像处理结果的精确性。故本文提出一种基于隶属度的模糊分类方法,该方法结合模糊关联规则 (fuzzy association rules, FARs) 和决策树方法来自动标注图像,方法的优点在于:一方面,在训练阶段,根据提出的评价标准获得模糊关联规则,该关联规则决定图像特征和语义关联的程度,更加直观的获取了低维图像特征和高维图像语义间的联系,很好地解决了“语义鸿沟”问题;根据隶属度函数将低维图像特征映射到模糊特征向量,旨在处理“锐利的边界”的问题。另一方面,运用决策树算法来解决冗余的模糊关联规则,过滤掉不必要的和弱的关联规则,大大减小了算法的时间复杂度,提高了标注的准确性和效率。

1 图像分割及特征提取

在图像自动标注中,图像用全局特征或局部特征表示,把图像分割成不同子区域。图像分割是图像处理的关键步骤,图像分割的好坏将影响到图像标注的精度。分割方法包括边缘检测、边缘跟踪、区域增长等。区域增长的图像分割通常能将具有相同特征的联通区域分割出来。并且能提供很好的边界信息和分割结果。针对本文“锐利边界问题”,为了体现对图像的边缘信息点的处理,文中采用区域增长方法分割图像。区域生长一般分 3 个步骤:1) 选择合适的生长点;2) 确定生长准则;3) 确定生长停止条件。表 1 给出了来自 Corel 5k 数据集的部分图像分割前后对比。

表 1 图像分割的几个实例

Table 1 Several examples of image segmentation			
原始图像	分割后图像	原始图像	分割后图像
			
			

图像表示和特征提取是标注算法中一个重要和决定性的步骤,图像进行区域分割后,从分割区域中提取低维视觉特征,特征向量呈现图像不同的属性和特征。图像特征表示的相关文献有 MPEG-7 标准^[13]、颜色聚合向量^[14],Lowe 在 2004 年提出的

SIFT 特征^[15]等,文中采用图像的颜色、边缘和纹理特征来表示。

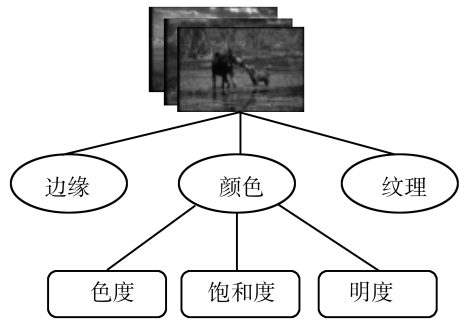


图 1 图像的视觉特征表示

Fig.1 Image representation of visual features

颜色是图像的最重要的特征之一,RGB、LUV、HSV、HMMD 是使用频繁的色彩空间,图像特征描述器包括颜色直方图、颜色矩、颜色聚合向量等。纹理特征能粗糙的捕捉图像的特点,在图像处理和计算机视觉中,纹理分析模型采用高斯分布、马尔可夫随机场和 Gabor 滤波器获取图像数据。边缘特征对于提取一些灰暗的图像特别重要,在 4 个方向(0°、45°、90°和 135°)使用 Canny 边缘检测方法检测边缘线。HSV 模型跟人类视觉感知密切相关,本文使用 HSV 来定义颜色空间,通过包含 9 个容器的一维直方图来统计图像颜色的特征分布,它把颜色空间量化成不同的小容器并且计算属于每个区间颜色像素点的频率。再综合 9 维的纹理特征向量和 4 维方向向量

表示的边缘特征,从每幅图像提取一个 22 维特征向量。特征表示如图 1 所示。

2 图像自动标注模型

关联规则算法广泛运用在数据挖掘^[16]和分类^[17]中,研究证明关联规则方法比支持向量机(support vector machine, SVM)、朴素贝叶斯、神经网络等方法具有更高的分类准确率,但关联规则算法处理“锐利边界”问题性能较弱。为了解决这个问题,Kuok 等提出了模糊技术来挖掘关联规则^[18],研究了很多基于 FARs 的分类模型,并广泛运用于数据挖掘中,同时,模糊关联规则在图像处理领域也有相关研究。文献[19]提出了一种基于自适应区间划分的模糊关联遥感图像分类方法。算法根据遥感图像分类的特点,利用模糊 C 均值聚类算法自适应地建立连续型属性模糊区间,使用新的剪枝策略对项集进行筛选从而避免生成无用规则,采用一种新的规则重要性度量方法对多模糊分类规则进行融合,从而有效地提高分类效率和精确度。文献[20]提出运用一种分层的模糊关联规则用于图像分类。但以上方法都存在模糊关联规则库过于庞大的问题。

本文提出的标注方法首先对视觉特征模糊化,再提取特征与语义之间的关系,形成模糊关联规则。最后用决策树方法处理冗余的规则,使标注性能更高,标注结果更准确。系统框架如图 2 所示。

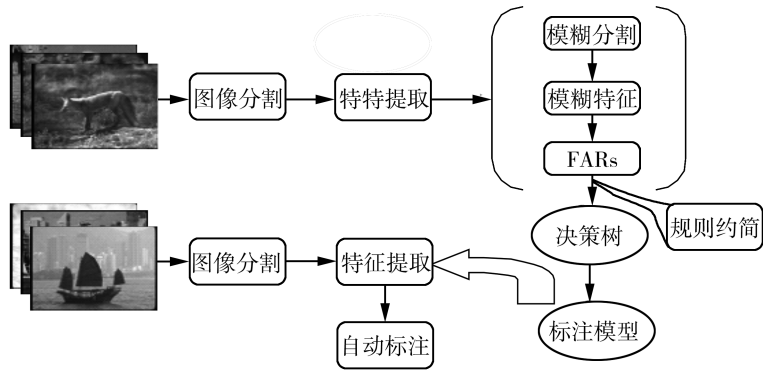


图 2 基于 FARs 和决策树的图像自动标注框架

Fig.2 Automatic image annotation framework based on FARs and decision tree

2.1 模糊关联规则

实验数据集分成训练和测试 2 个部分,训练数据集用来构建模型,测试数据集用来实现多标记图像的标注。在训练阶段,首先提取图像低维特征,根据模糊分割获得原始低维特征到模糊特征向量的映射,再计算模糊支持度和模糊置信度产生模糊关联规则,最后,使用决策树对关联规则库进行后期处理,删减冗余的关联规则。

2.1.1 模糊分割

在图像处理中,存在着模糊和不确定性的定义,如图像边缘、边界、区域和纹理等定义,模糊属性将干扰图像处理结果的精确性。

将颜色、纹理、边缘特征映射到模糊特征向量中,图 3 给出由三角隶属度函数转换数值型低维视觉特征到语义概念的模糊集,隶属度取值[0,1]。

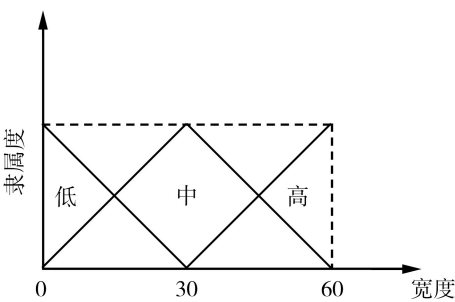


图 3 三角隶属度函数($K = 3$)

Fig.3 Triangular membership functions ($K = 3$)

图 3 中横坐标表示图像低维特征属性值,纵坐标表示其在模糊集中的隶属度, K 是每个低维视觉

特征属性映射到模糊分区数。

2.1.2 挖掘模糊关联规则

给定一个图像数据集 D , 其低维的特征向量集 $A = \{a_1, a_2, \dots, a_m\}$, m 代表特征的维度, 模糊向量集 $A_f = \{A_1^1, A_1^2, A_1^{p_1}, \dots, A_2^1, A_2^2, A_2^{p_2}, \dots, A_m^1, A_m^2, A_m^{p_m}\}$, 模糊集合 $S = \{1, 2, \dots, p_1, 1, 2, \dots, p_2, \dots, 1, 2, \dots, p_m\}$, 其中特征向量 a_j 的模糊集为 p_i , $u_i^p(a_j)$ 是特征 a_j 属于模糊集 p_i 的隶属度值。下面给出一个简单的二维图像特征的模糊化, $A = \{a_1, a_2\}$, 展示挖掘模糊关联规则的过程。模糊集 $MFs = \{low, mid, high\}$, 根据图 3 的三角隶属度函数, 可以得到原始特征向量和相应的模糊特征向量, 训练集是已经标注的图像, 标注语义用简单的字母表示, 详细的信息如表 2 所示。

表 2 原始特征向量与模糊特征向量

Table 2 Original feature vector versus fuzzy feature vector

图像编号	原始特征向量		Fuzzy(a_1)			Fuzzy(a_2)			标注语义
	a_1	a_2	$u_{low}^{a_1}$	$u_{mid}^{a_1}$	$u_{high}^{a_1}$	$u_{low}^{a_2}$	$u_{mid}^{a_2}$	$u_{high}^{a_2}$	
1	10	33	0.67	0.33	0	0.67	0.33	0	A, C, E, F
2	10	54	0.83	0.17	0	0.17	0.83	0	A, D
3	12	18	0.6	0.4	0	0.4	0.6	0	A, C, E
4	33	54	0	0.9	0.1	0	0.2	0.8	A, B, C
5	51	48	0	0.3	0.7	0	0.4	0.6	B, C, D, F
6	45	46.5	0	0.5	0.5	0	0.45	0.55	A, B, E
7	30	10.5	0	1	0	0.65	0.35	0	C, D
8	57	15	0	0.1	0.9	0.5	0.5	0	A, C
9	45	9	0	0.5	0.5	0.7	0.3	0	A, C, E
10	23.4	39	0.22	0.78	0	0	0.7	0.3	B, D, F
11	27	49.5	0.1	0.9	0	0	0.35	0.65	B, C, D
12	20.1	33	0.33	0.67	0	0	0.9	0.1	A, B, D, F

步骤 1 糊支持度(fuzzy support,FS)和模糊置信度(fuzzy confidence,FC)是模糊关联规则最常见的衡量标准,模糊支持度表示数据集中包含特征 a_i 以及语义 C_j 的百分比,模糊置信度衡量一条规则的精确度,用来计算数据集中包含特征 a_i 以及语义 C_j 的图像占含特征 a_i 的图像的百分比。关联规则(condition) $\rightarrow y$; condition 表示模糊特征向量的组合, y 表示图像语义。

例如 a_i is low $\rightarrow C_j$:

$$FS(a_i \text{ is low} \rightarrow C_j) = \frac{\sum_{a_i \in c_j} u_l(a_i)}{|T|} \tag{1}$$

$$FC(a_i \text{ is low} \rightarrow C_j) = \frac{\sum_{a_i \in c_j} u_l(a_i)}{\sum_{a_i \in c} u_l(a_i)} \tag{2}$$

式中: $|T|$ 是数据集 D 的图像数目,若 FS 等于 0.5,

表示所有实例中存在 50% 图像包含特征 a_i 并且图像语义为 C_j ; 若 FC 等于 0.5, 表示图像集中有 50% 被标注为 C_j 。根据式(1)、(2), 表 3 给出了从表 2 中挖掘出的 FARs(列出部分)及其 FS 和 FC。

步骤 2 构建 FARs。定义有效的模糊关联规则: 首先, 遍历所有模糊特征向量, 根据其 FS 大于或等于最小模糊支持度阈值 mFS 得到频繁项集, 并且关联规则的置信度大于或等于最小模糊置信度 mFC。然后, 基于 Apriori 算法获取 FARs。详细描述如算法 1 和算法 2 所示。

算法 1 取频繁项集 F

Input: 训练集 T , mFS.

Output: 频繁项集 F .

Process:

1) $C_1 \leftarrow$ Fuzzy partition(T)


```
2)  $F_1 \leftarrow \{ \text{fuzzy frequent 1 item sets} \}$ 
3) for (  $k = 2$ ;  $F_{k-1} \neq \emptyset$ ;  $k++$ ) do
4)  $C_k = \text{Candidate}(F_{k-1})$ ;
5) for each transaction  $t \in T$  do
6) for each candidate  $c \in C_k$  do
7)  $F_k = \text{Checking}(C_k, \text{mFS})$ 
8) return  $F = \bigcup_k F_k$ 
```

算法 2 获取 FARs

Input: 频繁项集 F , mFC.

Output: FARs.

Process:

```
1) for every feature attribute  $c \in C$ , do
2) while  $F_k \neq \Phi$ , do
3) for every frequent itemsets  $f \in F_k$ 
4) if (  $\text{fuzconf}(f \rightarrow c) \geq \text{MFC}$  ) then
5) Output the rule  $f \rightarrow c$  with  $\text{conf} = \text{fuzconf}(f \rightarrow$ 
```

表 3 模糊关联规则及其 FS、FC

Table 3 Fuzzy support and fuzzy confidence of association rules

编号	规则	Fuzzy support	Fuzzy confidence
1	a_1 is low $\rightarrow A$	0.23	0.88
2	a_1 is low $\rightarrow B$	0.05	0.24
3	a_1 is low $\rightarrow C$	0.11	0.50
4	a_1 is low $\rightarrow D$	0.12	0.54
5	a_2 is low $\rightarrow A$	0.20	0.79
6	a_2 is low $\rightarrow B$	0	0
7	a_2 is low $\rightarrow C$	0.24	0.95
8	a_2 is low $\rightarrow D$	0.07	0.27
9	a_1 is mid $\rightarrow A$	0.30	0.55
10	a_1 is mid $\rightarrow B$	0.34	0.62
11	a_1 is mid $\rightarrow C$	0.37	0.68
12	a_1 is mid $\rightarrow D$	0.32	0.58
13	a_2 is mid $\rightarrow A$	0.34	0.70
14	a_2 is mid $\rightarrow B$	0.25	0.51
15	a_2 is mid $\rightarrow C$	0.25	0.51
16	a_2 is mid $\rightarrow D$	0.29	0.60
17	a_1 is high $\rightarrow A$	0.17	0.74
18	a_1 is high $\rightarrow B$	0.11	0.48
19	a_1 is high $\rightarrow C$	0.18	0.81
20	a_1 is high $\rightarrow D$	0.06	0.26
21	a_2 is high $\rightarrow A$	0.12	0.48
22	a_2 is high $\rightarrow B$	0.16	1
23	a_2 is high $\rightarrow C$	0.17	0.68
24	a_2 is high $\rightarrow D$	0.14	0.55

2.2 决策树

决策树方法是一种基本的分类和回归方法,用于分类的决策树主要优点是具有可读性,分类速度快。决策树学习的思想主要来源于 Quinlan 提出的

著名的 ID3 算法、C4.5 以及 Breiman 提出的 CART 算法。

FARs 的不足在于模糊关联规则数目随着图像数据集的扩大变得庞大,将产生大量冗余的关联规则,导致标注性能降低。约简模糊关联规则成了研究的热点,文献[21-22]提出用遗传算法解决这个问题。

本文提出运用决策树算法^[23]实现最佳模糊关联规则提取,在决策树的每个节点处根据一个或多个属性值进行划分,这种方式直观且分类精度高、可读性好、分类速度快。通过算法可以过滤掉冗余的和弱的关联规则,实验证明该方法能取得很好的标注效果。

通过获得的模糊关联规则,规则的前件用来构建决策树的新的属性,若一条关联规则如下: $R_1: (a_1 \text{ is low}, a_2 \text{ is mid}) \rightarrow A, a_1 \cap a_2$ 即决策树的候选属性,这样的候选属性构建的决策树是否使标注效果更好,这是不确定的,这里提出了新的评价标准。首先,选择候选属性 AN_j 作为决策树的根节点,根节点的信息量计算如式(3), T 表示数据库图像实例数目, k 表示图像的类别数量, c_j 代表类 j 的图像数量。

$$G_{\text{root}} = - \sum_{j=1}^k \frac{|c_j|}{|T|} \lg \frac{|c_j|}{|T|} \tag{3}$$

用生成的候选属性进行数据分割时,使 $AN_j = 1$ 的数据记录的数目用式(4)表示,在这些符合 $AN_j = 1$ 的数据记录中,类别属性 C 的值为 c_i 的记录数目如式(5),但是类别属性 C 的值为 c_i 的记录数目是未知的,由于新的规则是利用近似精确规则来生成的,即当规则前件出现时,规则的后件仅在少数例外的情况下不出现。从而这里忽略了这类数据包含的信息量。

$$T_1 = T \frac{\text{sup}(R_i)}{\text{conf}(R_i)} \tag{4}$$

$$T_2 = T \cdot \text{sup}(R_i) \tag{5}$$

这样符合 $AN_j = 1$ 的数据记录所包含的信息量如式(6),而在这些不符合 $AN_j = 1$ 的数据记录数目为式(7):

$$G_1 = \frac{T_1}{T} \left[- \frac{T_2}{T_1} \lg \frac{T_2}{T_1} \right] \tag{6}$$

$$T_3 = T - T_1 = T \frac{\text{conf}(R_i) - \text{sup}(R_i)}{\text{conf}(R_i)} \tag{7}$$

设在这些不符合 $AN_j = 1$ 的数据记录中,类别属性值 $C_k(k = 1, 2, \dots, n)$ 的记录数目为 T_{3k} ,则这些记录所包含的信息量为

$$G_2 = \frac{T_3}{T} \left[- \sum_{k=1}^n \frac{T_{3k}}{T_3} \lg \frac{T_{3k}}{T_3} \right] \quad (8)$$

综上所述,新属性的近似信息增益为 $\text{Gain}(\text{AN}_j) = G_{\text{root}} - G_1 - G_2$, 其中,若 $\text{Gain}(\text{AN}_j) > 0$, 表示属性 AN_j 有效,应该保留,否则,删除该条规则。详细描述如算法 3 所示。

算法 3 根据新规则删减 FARs

Input: FARs

Output: 决策树

Process:

- 1) Generate new candidate attribute in the fuzzy association rules F ;
- 2) for every candidate attribute AN_j do
- 3) if $\text{Gain}(\text{AN}_j) > 0$ then
- 4) Keep AN_j in use;
- 5) else abandon AN_j ;
- 6) Delete the FARs which including the AN_j .

3 实验结果分析

3.1 数据集

实验采用 Corel 5k^[5] 和 IAPR TC-12^[12] 2 个数据集来进行图像标注测试,这 2 个数据集在近年来的图像标注研究中被广泛应用于标注算法性能的比较。自从被提出用于图像标注实验后,已经成为图像实验的标准数据集。

Corel 5k 图像集涵盖多个主题的 5 000 幅图片,Corel 5k 由 50 张 CD 组成,包含 50 个语义主题。每个 CD 包含 100 张大小相等的图像,每张 CD 代表一个语义主题,例如天空、非洲、海滩、建筑、城市等。Corel 5k 图像库通常分成 3 个部分: 4 000 张图像作为训练集,500 张图像作为验证集用来估计模型参数,其余 500 张作为测试集评价算法性能。

使用验证集寻找到最优模型参数后 4 000 张训练集和 500 张验证集混合起来组成新的训练集。该图像库中的每张图片被标注 1~5 个标注词,训练集中总共有 374 个标注词,在测试集中总共使用了 263 个标注词。

IAPR TC-12 数据集应用在自动图像标注并扩展到多媒体信息检索影响的评估上,该数据集包括 20 000 幅分割图像,包括不同的动物,城市和许多其他类别的图像。在特征提取阶段,每个区域的特征向量具有 99 535 个属性,所述特征由已提取的颜色空间 LAB、纹理和空间位置表示。

3.2 参数设置

方法中最小模糊支持度(mFS)、最小模糊置信度(mFC)和分区大小 K 3 个参数都将影响图像标注性能。为了获得最优参数,在 Corel 5k 和 IAPR TC-12 数据集下,分别设置不同 mFS 和 mFC 的参数值,如图 4 所示,当 $\text{mFS}=0.05$, $\text{mFC}=0.85$ 时,标注的精度最高,为 92.5%。从结果可以看到 mFS 越大,对于标注的性能更敏感。因此实验中选择更小的 mFS。另一方面,mFS 越高,标注的性能更加健壮。从表中可以看到最好的标注准确率 92.5%满足的条件是 $(\text{mFS}, \text{mFC}) = (0.05, 0.85)$ 。同样,把这 2 个最优参数用于后续分区最佳大小 K 的确定实验中。图 5 给出了当分区数分别为 3、4、5、6、7、8 时,标注的性能比较。从结果看出, K 为 6 满足最佳效果。综上,实验将每个低维图像特征属性划分为 6 个模糊分区。由原来的 22 维低维图像特征,模糊分割后,得到 22×6 维的模糊特征向量。

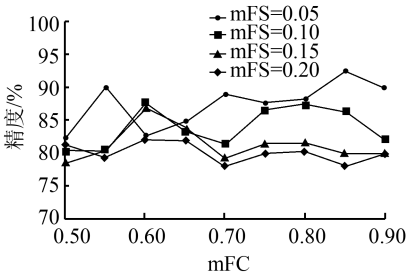


图 4 不同 mFC 和 mFS 下的图像标注精度

Fig.4 Precisions of image annotation with different mFC and mFS

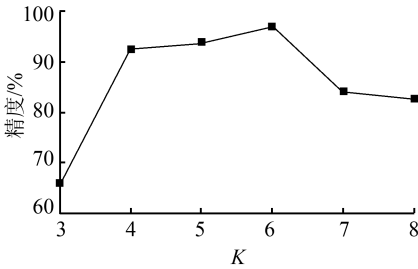


图 5 不同 K 值下的图像标注精度 ($\text{mFS}=0.05$, $\text{mFC}=0.85$)

Fig.5 Precisions of image annotation with different K ($\text{mFS}=0.05$, $\text{mFC}=0.85$)

3.3 评价标准

实验在机器系统 Windows 7, 软件 Visual Studio 2012 下测试运行,使用 Intel 2.66 GHz Pentium4 CPU。实验给出了标注的评价标准、精确率、召回率以及综合评价指标 F-measure,公式如下:

$$\text{precision} = a/b \quad (9)$$

$$\text{recall} = a/c \quad (10)$$

$$F\text{-measure} = \frac{2 \times \text{precision} \times \text{recall}}{\text{recall} + \text{precision}} \quad (11)$$

式中: a 表示正确标注为某类的图像数量, b 表示所有自动标注为该类的图像数量, c 表示测试集中真实标注包含该类的所有图像数量。

表 4 和表 5 给出本文标注方法的准确率、召回率及 F-measure,从表 4 中看出,该方法可以高效且准确地标注测试图像集。从表中看出,类别“建筑”在 Corel 5k 和“景色”在 IAPR TC-12 数据集中标注准确率尤为突出。

表 4 Corel 5k 中标注的评价标准

Table 4 Evaluation measure of annotation on Corel 5k

类别	precision	recall	F-measure
非洲	0.81	0.77	0.85
沙滩	0.87	0.80	0.90
建筑	0.93	0.89	0.93
山	0.87	0.86	0.89
大象	0.92	0.88	0.95
平均值	0.88	0.84	0.90

表 5 IAPR TC-12 中标注的评价标准

Table 5 Evaluation measure of annotation on IAPR TC-12

类别	precision	recall	F-measure
运动	0.79	0.75	0.82
人	0.87	0.89	0.94
风景	0.91	0.90	0.88
动物	0.88	0.81	0.87
平均值	0.86	0.84	0.88

3.4 对比实验

3.4.1 Corel 5k 上的实验结果

为了更有力证明该方法的准确性和可靠性,实验做了与其他相关标注方法的对比,对比方法有翻译模型^[9]、JEC^[12]、PLSA-WORDS^[6]、HGDM^[8]和 CFAR^[24],PLSA-WORDS 是基于 PLSA 的标注模型,HGDM 是一种基于混合模型的标注方法,CFAR 提出一种模糊关联规则语义标注模型。

实验在 Corel 5k 和 IAPR TC-12 2 个数据集下进行,表 6 为 Corel 5k 数据集下所得标注结果,在性能最佳的 49 个关键词集合和训练集中所有 260 个关键词集合中评价平均精度(MP)和平均召回率(MR),从表中数据得出,本文提出的方法在性能上要优于其他多数标注模型,体现了算法的准确性和稳定性。

表 6 Corel 5k 上图像自动标注模型的性能比较
Table 6 Performance comparison of automatic image annotation models on Corel 5k

模型	召回率>0 的关键词个数	49 个性能最佳的词集上的结果		全部 260 个词集上的结果	
		MP	MR	MP	MR
TM	49	0.20	0.34	0.06	0.04
JEC	139	—	—	0.27	0.32
PLSA-WORDS	105	0.56	0.71	0.14	0.20
HGDM	137	0.78	0.83	0.28	0.32
Our approach	138	0.81	0.85	0.30	0.35

3.4.2 IAPR TC-12 上的实验结果

表 7 给出了 IAPR TC-12 数据集上实验结果,测试在所有 291 个关键词下,测试包括召回率大于 0 的关键词个数、平均精度、平均召回率。对比其他标注模型,表现出更好的标注性能。

表 7 IAPR TC-12 上图像自动标注模型的性能比较

Table 7 Performance comparison of automatic image annotation models on IAPR TC-12

模型	召回率>0 的关键词个数	MP	MR
JEC	196	0.25	0.16
PLSA-WORDS	177	0.18	0.12
HGDM	194	0.29	0.18
Our approach	199	0.32	0.21

为了验证该模型的高效性,实验在 2 个图像集下,选择不同图像数目进行测试。表 8 中,基于模糊关联规则的标注方法的平均精确度比非模糊关联规则更高,另外,本文提的方法对比 CFAR,实验得到的规则数目减少,标注精确度增加,证明决策树对于删减冗余关联规则有显著的成效。

3.5 在 Corel 5k 中的实验结果

在 Corel 5k 数据集下,从 5 000 幅图像中随机选择 60%图像作为训练集,运用本文方法训练获得标注模型,即决策树删减后的模糊关联规则库,再对余下 40%的图像进行标注。

表 9 给出随机抽取的一些图像,实验选取标注精确度最高的 5 个关键字作为最后的标注结果,表中给出在 Corel 5k 图像集下与人工标注的比较,实验结果充分证明了此方法的准确性和可靠性。




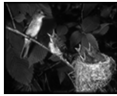


表 8 Corel 5k 和 IAPR TC-12 上不同模型的分类结果比较

Table 8 Classification results comparison of different models on Corel 5k and IAPR TC-12

类别	数据集	图像数	SVM 精确度	CFAR		Our approach	
				精确度	规则数	精确度	规则数
天空	Corel 5k	100	0.85	0.83	12	0.83	9
大海	Corel 5k	56	0.76	0.82	7	0.83	5
人	Corel 5k	78	0.80	0.81	9	0.82	7
树	Corel 5k	90	0.69	0.77	6	0.80	6
花	Corel 5k	30	0.69	0.75	4	0.77	5
运动	IAPR TC-12	86	0.71	0.80	10	0.82	8
人	IAPR TC-12	55	0.70	0.74	7	0.80	6
景	IAPR TC-12	90	0.71	0.80	9	0.79	6
动物	IAPR TC-12	60	0.72	0.82	8	0.80	7

表 9 Corel 5k 中图像自动标注结果

Table 9 Results of automatic image annotation on Corel 5k

						
本文方法	boat, sky, buildings, water	tiger, snow, stone	sun, water, seabeach	birds, tree, branchleaf	cars, road, buildings	snow, mountain, stone, sky
人工标注	boat, city, buildings	tiger, snow	sun, water, seabeach, sky	birds, branchleaf	cars, grass buildings, road	snow, mountain, stone, sky

4 结束语

本文提出基于 FARs 和决策树的图像自动标注方法,在该标注模型中,FARs 提取低维图像特征和高维语义间的联系,在现实世界中,由于面对的是海量的图像数据集,将决策树方法应用到约简 FARs 中。实验证明此标注模型不仅大大减少了标注的时间,也提高了标注的准确性。与经典的机器学习方法 SVM、boosting、neural network 等比较,优势在于运用模糊关联规则方法直观的获取了低维特征和高维语义间的联系;与传统的关联规则方法相比,增加了决策树对模糊关联规则的处理,提高了标注性能;在未来的研究方向中,重点放在图像低维属性的模糊分割上,尝试更多的隶属度函数将低维特征模糊化,获得更准确的模糊分割函数;另一方面,在对 FARs 评价的准则上,加入更多可靠的标准,以达到更高的标注精度和更好的检索效果。

参考文献:

[1]CHANG S K, HSU A. Image information systems: where do we go from here? [J]. IEEE Transactions on Knowledge

and Data Engineering, 1992, 4(5): 431-442

[2]MARKKULA M, SORMUNEN E. End-user searching challenges indexing practices in the digital newspaper photo archive[J]. Information Retrieval, 2000, 1(4): 259-285.

[3]JEON J, LAVRENKO V, MANMATHA R. Automatic image annotation and retrieval using cross-media relevance models [C]//Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, USA, 2003: 119-126.

[4]WANG Yong, MEI Tao, GONG Shaogang, et al. Combining global, regional and contextual features for automatic image annotation[J]. Pattern Recognition, 2009, 42(2): 259-266.

[5]DUYGULU P, BARNARD K, DE FREITAS J F G, et al. Object recognition as machine translation: learning a lexicon for a fixed image vocabulary [M]//HEYDEN A, SPARR G, NIELSEN M, et al. Lecture Notes in Computer Science, vol. 2353. Berlin: Springer-Varlag, 2002: 97-112.

[6]MONAY F, GATICA-PEREZ D. Modeling semantic aspects for cross-media image indexing[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(10): 1802-1817.

[7]李志欣,施智平,李志清,等. 融合语义主题的图像自动标注[J]. 软件学报, 2011, 22(4): 801-812.

LI Zhixin, SHI Zhiping, LI Zhiqing, et al. Automatic image

- annotation by fusing semantic topics[J]. Journal of Software, 2011, 22(4): 801-812.
- [8] LI Zhixin, SHI Zhongzhi, ZHAO Weizhong, et al. Learning semantic concepts from image database with hybrid generative/discriminative approach[J]. Engineering Applications of Artificial Intelligence, 2013, 26(9): 2143-2152.
- [9] 茹立云, 马少平, 路晶. 基于 Boosting 学习的图片语义自动标注[J]. 中国图象图形学报, 2006, 11(4): 486-491. RU Liyun, MA Shaoping, LU Jing. Boosting-based automatic linguistic indexing of pictures[J]. Journal of Image and Graphics, 2006, 11(4): 486-491.
- [10] SUMATHI T, HEMALATHA M. An innovative hybrid hierarchical model for automatic image annotation[M]//KRISHNA P V, BABU M R, ARIWA E. Global Trends in Information Systems and Software Applications, Volume 270. Berlin: Springer-Varlag, 2012: 718-726.
- [11] 张静, 胡微微, 陈志华, 等. 多模型融合的多标签图像自动标注[J]. 计算机辅助设计与图形学学报, 2014, 26(3): 472-478. ZHANG Jing, HU Weiwei, CHEN Zhihua, et al. Multi-model fused framework for image annotation[J]. Journal of Computer-Aided Design and Computer Graphics, 2014, 26(3): 472-478.
- [12] MAKADIA A, PAVLOVIC V, KUMAR S. Baselines for image annotation[J]. International Journal of Computer Vision, 2010, 90(1): 88-105.
- [13] LI Feifei, FERGUS R, PERONA P. One-shot learning of object categories[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(4): 594-611.
- [14] PASS G, ZABITH R. Histogram refinement for content-based image retrieval[C]//Proceedings of the 3rd IEEE Workshop on Applications of Computer Vision. Sarasota, USA, 1996: 96-102.
- [15] LOWE D G. Distinctive image features form scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [16] JUKIC N, NESTOROV S. Comprehensive data warehouse exploration with qualified association-rule mining[J]. Decision Support Systems, 2006, 42(2): 859-878.
- [17] HU Y C, CHEN R S, TZENG G H. Mining fuzzy association rules for classification problems[J]. Computer and Industrial Engineering, 2002, 43(4): 735-750.
- [18] KUOK C M, FU A, WONG M H. Mining fuzzy association rules in databases[J]. ACM SIGMOD Record, 1998, 27(1): 41-46.
- [19] 董杰, 沈国杰. 一种基于模糊关联分类的遥感图像分类方法[J]. 计算机研究与发展, 2012, 49(7): 1500-1506.
- DONG Jie, SHEN Guojie. Remote sensing image classification based on fuzzy associative classification[J]. Journal of Computer Research and Development, 2012, 49(7): 1500-1506.
- [20] TAZAREE A, EFTEKHARI-MOGHADAM A M, SAJJADI-GHAEM-MAGHAMI S. A semantic image classifier based on hierarchical fuzzy association rule mining[J]. Multimedia Tools and Applications, 2014, 69(3): 921-949.
- [21] SHIBUCHI H, NOZAKI K, YAMAMOTO N, et al. Selecting fuzzy if-then rules for classification problems using genetic algorithms[J]. IEEE Transactions on Fuzzy Systems, 1995, 3(3): 260-270.
- [22] PACH F P, ABONYI J. Association rule and decision tree based methods for fuzzy rule base generation[J]. International Scholarly and Scientific Research and Innovation, 2008, 2(1): 546-551.
- [23] GAO Jing, ZHAO Baoyong. New method about how to construct decision tree based on association rule[C]//Proceedings of International Workshop on Open-Source for Scientific Computation. Beijing, China, 2011: 131-135.
- [24] SILLA C N, FREITAS A A. A survey of hierarchical classification across different application domains[J]. Data Mining and Knowledge Discovery, 2011, 22(1/2): 31-72.

作者简介:



李志欣,男,1971年生,副教授,博士,主要研究方向为图像理解、机器学习、多媒体分析与检索。发表学术论文40余篇,其中SCI收录5篇,EI收录30篇。



李灵芝,女,1987年生,硕士研究生,主要研究方向为图像理解、机器学习。



张灿龙,男,1975年生,副教授,博士,主要研究方向为模式识别、图像目标跟踪。

[责任编辑:孟玮]