

# 图像语义自动标注及其粒度分析方法

张素兰<sup>1,2</sup> 郭平<sup>1</sup> 张继福<sup>2</sup> 胡立华<sup>2</sup>

**摘要** 缩小图像底层视觉特征与高层语义之间的鸿沟, 以提高图像语义自动标注的精度, 进而快速满足用户检索图像的需求, 一直是图像语义自动标注研究的关键. 粒度分析方法是一种层次的、重要的数据分析方法, 为复杂问题的求解提供了新的思路. 图像理解与分析的粒度不同, 图像语义标注的精度则不同, 检索的效率及准确度也就不同. 本文对目前图像语义自动标注模型的方法进行综述和分析, 阐述了粒度分析的思想、模型及其在图像语义标注过程中的应用, 探索了以粒度分析为基础的图像语义自动标注方法并给出进一步的研究方向.

**关键词** 图像内容检索, 自动图像语义标注, 粒度分析, 粒计算模型

**引用格式** 张素兰, 郭平, 张继福, 胡立华. 图像语义自动标注及其粒度分析方法. 自动化学报, 2012, 38(5): 688–697

**DOI** 10.3724/SP.J.1004.2012.00688

## Automatic Semantic Image Annotation with Granular Analysis Method

ZHANG Su-Lan<sup>1,2</sup> GUO Ping<sup>1</sup> ZHANG Ji-Fu<sup>2</sup> HU Li-Hua<sup>2</sup>

**Abstract** To bridge the semantic gap between low-level visual feature and high-level semantic concepts has been the subject of intensive investigation for years in order to improve the accuracy of automatic image annotation and satisfy the users' needs of quick image retrieval. Granular analysis is a hierarchical and important data analyzing method, which provides a new idea and method for solving the complicated problem. The accuracy of automatic image annotation and the efficiency of image retrieval are varying with the granularity size of image understanding and analysis. In this paper, the state-of-art models of automatic semantic image annotation are overviewed, then the idea and models of the granular analysis with its application in the process of automatic semantic image annotation are discussed, and the granular analysis based automatic image annotation methods are investigated as well as the promising research directions are given.

**Key words** Content-based image retrieval (CBIR), automatic semantic image annotation, granular analysis, granular computing model

**Citation** Zhang Su-Lan, Guo Ping, Zhang Ji-Fu, Hu Li-Hua. Automatic semantic image annotation with granular analysis method. *Acta Automatica Sinica*, 2012, 38(5): 688–697

早期的图像语义检索是基于图像内容的检索 (Content-based image retrieval, CBIR)<sup>[1]</sup>, 由于相同的视觉特征本身存在不同的语义, 而且因光照、遮挡等各种因素的影响, 底层视觉特征与高层语义存在明显的鸿沟. 同时, 随着网络和数码电子技术的发展, 各种媒体图像数据以 TB 级速度增长, 面对如此海量高维的跨媒体图像数据集, 传统的手工图像语义标注方法费时、费力且精度不能得到保障, 无法有效满足日益增长的用户检索需求. 因此, 如何结合人工智能、数据挖掘方法和技术, 从浩瀚的图像数据集中提取和分析相关的有用知识, 研究一种有效的图

像语义自动标注方法, 提高图像语义标注精度以快速满足用户检索需求, 成为当前一个非常有前景的研究领域.

目前, 研究者提出了很多图像语义自动标注模型<sup>[1]</sup>, 采用的主要是统计方法、机器学习和计算智能等技术. 统计方法侧重研究图像数据点、样本点的客观规律, 机器学习和计算智能技术主要通过研究这些客观规律, 挖掘隐藏在其中的语义知识和规律, 以对未标注图像做出准确的语义标注. 然而, 尽管对图像语义标注模型的研究取得了相当丰富的成果, 但在视觉底层特征的高维约简、多媒体高层语义表示及融合、分类模型的优化等方面仍存在一些问题, 使得图像底层视觉特征与高层之间的语义鸿沟 (Semantic gap) 仍不能得到很好的解决. 图像语义自动标注的本质是建立某种分类函数, 或者说建立一种有效的映射技术, 将未标注的图像准确地映射到相应的语义类别中. 因此, 研究有效的映射方法, 提高图像标注的精度, 以快速满足用户检索图像的需求, 成为众多研究者关注的重点.

收稿日期 2011-04-11 录用日期 2011-06-29  
Manuscript received April 11, 2011; accepted June 29, 2011  
国家自然科学基金 (61073145, 90820010) 资助  
Supported by National Natural Science Foundation of China (61073145, 90820010)

本文责任编辑 刘一军

Recommended by Associate Editor LIU Yi-Jun

1. 北京理工大学计算机学院 北京 100081 2. 太原科技大学计算机科学与技术学院 太原 030024

1. School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081 2. School of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024

信息粒度化 (Information granularity) 的思想由 Zadeh 于 1979 年提出<sup>[2]</sup>, 采用一种分而治之、层次求解的方法, 为复杂问题的求解提供了新的思路, 目前已在智能数据分析、海量数据挖掘、复杂问题处理和图形图像处理等方面得到了广泛的应用<sup>[3]</sup>. 图像语义标注的精度和效率与图像理解和分析的层次, 即粒度有着密切的关系, 粒度不同, 标注的精度和性能可能就不同. 本文综述了一些主要的图像语义自动标注模型和方法, 阐述了粒度分析方法在图像语义自动标注过程中的应用, 在此基础上, 给出了进一步深入研究的方向.

## 1 图像语义自动标注

图像语义自动标注过程主要涉及图像的底层特征提取及表示、图像语义表示和图像语义标注模型的建立等方面. 图像特征提取即提取图像的颜色、纹理、形状和空间分布等底层视觉特征. 目前, 图像特征提取的方法有基于整幅图像 (Global image)<sup>[4]</sup>、均匀网格 (Grid)<sup>[5]</sup>、区域 (Region) 或块 (Block)<sup>[6]</sup> 以及对象<sup>[7]</sup> 等方法. 基于整幅图像和均匀划分网格技术从全局或者固定的空间布局上提取特征, 不需进行图像分割, 特征提取速度快, 对标注图像的场景语义效果较好, 但是对象的标注和检索效果不理想<sup>[8]</sup>. 基于区域、块和对象的特征提取方式需对图像进行分割, 从分割后的区域中局部提取底层特征, 该方法对光的反射 (Reflection)、部分可见 (Partial visibility) 以及遮挡 (Occlusion) 等具有鲁棒性, 有利于对象的标注, 但是分割技术和分割粒度的粗细却影响最终的标注效果及其复杂性<sup>[8]</sup>.

图像的底层特征表示方式主要有直方图、区域特征、边缘特征和视觉词包 (Bag-of-visual-words, BOV) 等. 直方图是对随机变量分布的一种近似表现形式, 同时也是对密度估计的描述<sup>[9]</sup>. Swain 等最先提出了使用颜色直方图的表达方法作为图像内容检索特征<sup>[10]</sup>, 该图像表达方式简单, 但是整幅图像提取的颜色直方图不能体现颜色之间的空间关系, 因此目标检测效果较差. 尺度不变特征转换 (Scale invariant feature transform, SIFT)<sup>[11]</sup> 是区域特征表达的主要方式, SIFT 描述子对图像的尺度、旋转、亮度等具有不变性, 在时域和频域内都能较好地定位, 对遮挡、噪声等具有很强的鲁棒性. 边缘特征表达主要以形状上下文为主, 与 SIFT 描述子相似, 但它是基于边缘而非区域的<sup>[9]</sup>. BOV 表示法是一种基于检测图像局部关键点特征的表达方式, 源于自然语言处理中的词包 (Bag-of-words, BOW) 模型, 其思想主要是通过提取图像中局部基元特征并进一步对其进行聚类而获得图像视觉词典, 再根据图像中各个视觉单词出现的统计分布, 得到表述图像的

视觉单词统计直方图. 由于 BOV 表示法检测图像局部的关键点, 对遮挡、旋转、背景和光照改变等具有鲁棒性, 因而成为视觉物体的一种主要表示方式, 在图像检索、场景分类和语义标注等领域取得了良好的应用效果. 目前, 国内外研究者对 BOV 进行了广泛的研究, 主要集中在局部基元提取、视觉单词生成和图像分类等方面<sup>[12-18]</sup>.

图像语义表示通过用某种表示方法、技术或工具 (Wordnet、Mpeg-7、本体、图和语义网络等) 将图像的语义表示出来. Wordnet 是一种基于认知语言学的英语词典, 描述了词语之间存在的重要的语义关联关系 (同义或反义、整体或部分、上层或下层), 很多图像处理领域专家用此衡量图像语义标注词之间的相关性, Jin 等通过利用 Wordnet 删除一些冗余的标注词<sup>[19]</sup>, 进一步精化图像语义标注词, 进而提高语义标注的精度. Mpeg-7 是一个描述多媒体内容的标准化框架, 用于描述多媒体信息的颜色、纹理、形状等底层语义. 本体 (Ontology) 是一种共享概念模型的明确的形式化规范说明, 通过描述概念及概念之间的关系来表示概念的语义, 由于本体可以有效地表达知识、查询知识或对不同领域的知识进行语义消解, 因此也常用于表示图像的语义. 图和语义网络是人工智能两种有效的知识表示工具, 用于表示图像高层语义之间的复杂联系. 卢汉清等将图学习图像复杂的语义关联关系用于改善图的标注过程<sup>[20]</sup>. Lu 等使用马尔科夫网络模型捕捉图像语义之间复杂的联系, 通过使用上下文、一致性和多样性等线索提出了一种新的基于图的交互式图像分类架构<sup>[21]</sup>.

图像语义自动标注模型的建立过程实际上是依据已标注的训练图像集, 构造一种分类模型, 或者说建立一种图像底层特征与高层语义之间的映射关系、对应关系, 然后根据该映射关系, 自动地将待标注图像的高层语义标注出来. 图像语义自动标注模型建立的一般过程如图 1 所示.

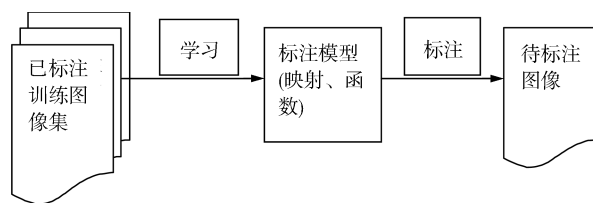


图 1 图像语义自动标注模型建立的一般过程

Fig. 1 The general process of automatic semantic image annotation

## 2 图像语义自动标注模型

图像语义标注模型的建立一直是图像研究领域的重点, 吸引了许多国内外研究学者的关注. 目前, 研究者利用统计方法和机器学习技术设计出

各种不同的图像语义自动标注模型, 主要有翻译模型 (Translation model)、潜在语义分析 (Latent semantic analysis, LSA) 模型、概率潜在语义分析 (Probabilistic latent semantic analysis, PLSA) 模型、隐狄利克雷分配 (Latent Dirichlet allocation, LDA) 模型、相关模型、支持向量机 (Support vector machine, SVM)、贝叶斯 (Bayes) 和高斯混合模型 (Gaussian mixture model, GMM) 等统计方法以及聚类、关联、示例和相关反馈 (Relevance feedback) 等机器学习技术, 如表 1 所示。

## 2.1 基于统计方法的图像语义自动标注模型

基于统计方法的图像语义自动标注模型一般分为生成式 (Generative) 和判别式 (Discriminative) 两种, 生成式的图像标注模型主要通过学习已标注的训练样本图像的视觉特征和语义概念之间的联合概率, 实现对未标注图像的语义标注。判别式的图像标注模型的主要思想是将每个语义标注词看作一个独立的类, 并为每个类学习相应的判别函数 (分类器), 依据已标注图像的语义类建立判别函数, 根据该函数对需标注的图像进行语义标注<sup>[8]</sup>。

### 1) 生成式

基于生成式的图像标注模型主要有翻译模型、LSA、PLSA、LDA 和相关模型等。翻译模型将标注过程看成一个从图像 (区域) 集到标注词集的翻译过程, 或者理解为从图像视觉词到语义词的翻译过程。Duygulu 等提出翻译模型<sup>[22]</sup>, 对分割后的图像区域特征进行聚类, 将连续特征变成离散视觉词集, 通过寻找标注词和图像特征之间的关系对待标注图像进行标注。这种方法标注的精度与区域采用什么样的聚类算法来提取离散的视觉特征点有关系, 而且视觉特征与语义的结合也比较松散。

LSA, 也称为 LSI (Latent semantic index), 是 Deerwester 等于 1990 年提出的一种新的检索方法<sup>[23]</sup>, 与一般的向量空间模型 (Vector space model) 表示方法不同, 该方法将词 (Terms) 和文档 (Documents) 映射到潜在的语义空间, 进而消除原始向量空间中的一些噪声数据, 提高了信息检索的精度。该模型在图像中主要用于识别可视文本特征子空间中

的语义子空间, 其主要思想是将高维向量映射到潜在的语义空间使其降维。

Hofmann 提出了一种新的统计模型 PLSA<sup>[24]</sup>, 该模型始于对自然语言和文本学习的研究过程中, 通过建立词语和文本间的映射关系, 处理文本中词语的多义性 (一词多义) 和相似性 (一义多词)。其思想是对每一文本-词语观察值  $(d_i, w_j)$  关联一个隐变量  $z_k$ , 通过求解隐关系  $p(w_j|z_k)$  的凸组合以求得  $(d_i, w_j)$  的概率分布<sup>[9]</sup>, 即

$$p(w_j|d_i) = \sum_{k=1}^K p(w_j|z_k)p(z_k|d_i) \quad (1)$$

但是, 当训练数据中存在噪音或训练数据太少时, PLSA 有时会出现过拟合的现象。为避免过拟合, PLSA 一般使用期望最大化 (Expectation-maximization, EM)<sup>[25]</sup> 算法对隐变量模型进行最大似然估计。由于 PLSA 采用了生成式潜在类模型实现一种概率混合分解, 具有坚实的统计推理基础, 故已成功应用于图像目标识别和场景语义标注, 为避免过拟合问题, LDA 模型<sup>[26]</sup> 在 PLSA 模型的基础上增加了超参层, 建立了隐变量  $z$  的概率分布, 其在图像语义标注和检索中也得到了成功应用。

相关模型标注思想假设图像 (或局部特征) 和语义标注词之间并非一一对应, 而是以一定的概率存在对应关系, 因此, 需要首先通过计算图像特征向量与语义标注词的联合概率分布, 然后对其进行标注。Jeon 等提出了一种跨媒体相关模型 (Cross-media relevance model, CMRM) 的图像标注和检索方法<sup>[27]</sup>, 图像被假设为由 Blob 组成的词汇表所表示, 即把图像分割为区域, 对类似的区域进行 K-均值聚类, 将每个聚类作为图像词汇, 称为“Blob”, 通过计算 Blob 和语义标签的概率, 并结合 PLSA 模型实现了图像的语义标注。为了避免 CMRM 对聚类数目过于敏感的缺点, Lavrenko 等对 CMRM 进行了改进, 提出连续空间相关模型 (Continuous-space relevance model, CRM)<sup>[28]</sup>, 将图像划分成区域, 不再使用聚类表示 Blob, 而是用一个连续特征向量值表示。

表 1 图像语义自动标注模型

Table 1 The automatic semantic image annotation model

方法和技术	类别	主要模型
统计	生成式	翻译模型、潜在语义分析模型、概率潜在语义分析模型、隐狄利克雷分配模型、相关模型
统计	判别式	支持向量机、贝叶斯、高斯混合模型
机器学习	归纳	聚类、关联、示例
机器学习	示教	相关反馈
机器学习	类比	网络检索

## 2) 判别式

常见的判别式模型有 SVM、Bayes 和 GMM 等<sup>[29]</sup>. SVM 因其良好的泛化能力、可处理非线性和高维小样本等特点, 在图像语义标注中得到广泛的应用, 该方法将与给定标注词相关的图像作为正例, 其他图像作为该类的负例, 从而进行多个二元分类器的学习. Li 等<sup>[30]</sup> 使用对象作为场景中的属性来表示图像, 并利用 SVM 进行场景分类, 取得了良好的分类效果.

Bayes 方法通过运用 Bayes 公式求解后验概率, 实现图像语义标注. Dong 等提出了一种利用 Bayes 进行软标注的方法<sup>[31]</sup>, 对每个训练的语义标签, 赋予图像一个信任级别, 这个信任标签向量可用于按照关键词的搜索, 将相关图像排序. Shi 等使用层次多项式混合模型对每个语义概念进行建模, 将先验层次知识与图像的多级概念结构表示方法相结合, 最后使用 Bayes 学习框架进行图像语义标注<sup>[32]</sup>.

由于图像数据的视觉多样性, 模型中往往存在很多参数, 需要大量的训练图像数据进行参数学习, 因此很多研究者使用 GMM 模型方法标注图像语义. Barnard 等提出的图像语义自动标注模型先用 GMM 对每个语义类的分布进行建模, 再用 EM 方法对模型参数进行学习<sup>[33]</sup>. Perronnin 利用 GMM 建模整个视觉词典, 将每个高斯成分作为一个视觉单词, 提出了一种通用的用于一般可视化分类的自适应调整视觉词的方法<sup>[12]</sup>. 一般地, 为了取得好的图像标注效果, 也常将两种统计方法混合使用<sup>[34]</sup>.

## 2.2 基于机器学习策略的图像语义自动标注

学习是人类智能的根本特征, 人类通过学习不断提高和改进自己的能力. 1983 年, Simon 提出了学习的概念, 认为学习是能够让系统在执行同一任务或同类的另一任务时比前一次执行得更好. 可见, 机器学习是一个有特定目的的知识获取过程, 通过获取知识、积累经验、发现规律, 使系统性能得到改进、实现自我完善、自适应环境<sup>[35]</sup>. 图像语义自动标注过程不仅使用了大量的统计学习方法, 还融入了很多的机器学习策略, 如归纳、示教和类比等.

### 1) 归纳 (Induction)

归纳学习的思想是从给出的实例环境中取得若干与某概念有关的例子, 经归纳推理得出一般概念性的一种学习方法. 在图像语义标注中, 常用的归纳法主要有无监督 (Unsupervised) 学习的聚类、关联和有监督 (Supervised) 学习的示例学习. 聚类主要用于图像语义标注过程中图像低层特征的提取, 其思想是利用某种聚类方法将图像 (区域) 聚合成某些有意义的类, 使同一类内的图像 (图像区域) 相似度

尽可能地大, 不同类间的相似度尽可能地小, 然后利用统计或者机器学习的方法为每一类加一个类标签, 以实现对各聚类区域的语义标注, 常见的图像聚类方法有 K-均值方法及其变型<sup>[36]</sup>. 夏利民等采用改进的 K-均值无监督图像分割算法将图像分割成不同的区域, 提出利用信息瓶颈聚类方法对分割后的区域进行聚类, 进而建立图像语义概念和聚类区域之间的相互关系<sup>[37]</sup>. 通常, 图像区域特征点的选取影响标注的精度, 因此聚类算法的好坏直接影响标注的性能. Frey 等提出一种新的基于近邻传播 (Affinity propagation, AP) 的聚类算法<sup>[38]</sup>, 该算法的思想是根据  $N$  个数据点之间的相似度进行聚类, 不需要事先定义聚类个数, 而是将所有数据点都作为代表点, 看作潜在的类别中心, 称之为 Exemplar, 通过迭代过程不断更新每一个点的吸引度和归属度值, 直到产生  $m$  个高质量的 Exemplar, 同时将其余数据点分配到相应的类别中. 由于该算法采用动态求解聚类中心, 因此聚类效果较好. Dueck 等将该算法思想用于人图像脸的识别, 取得了良好的分类效果<sup>[39]</sup>. Yang 等<sup>[40]</sup> 提出应用 AP 算法改进图像语义标注建模过程, 提高了对高斯混合模型参数估计的效率. 关联思想是利用已标注好的训练图像数据集, 建立图像的视觉特征和语义关键词之间的关联关系, 然后利用这一关系预测未标注的图像的语义. Li 等用分类的方法在图像的视觉特征和语义关键词之间建立关联关系, 依据关联关系构建分类器, 并将分类器用于未标注图像的标注过程<sup>[41]</sup>. 但是, 这种基于视觉特征的分类方法通常将具有相同视觉特征的区域归为一类, 即使区域的语义完全不同, 也用相同的关键词标注, 因此标注的精确度较低. 为避免这些不足, 可以在区域归类之前, 依据一些先验信息赋予其初始语义, 以提高标注精度.

示例学习是一种有监督的学习过程, 通过从环境中取得若干与某概念相关的实例, 经归纳得出一般性概念的描述. 然而, 为避免归纳出的概念存在偏差 (歧义), 该方法要求学习环境给出足够的正例和反例, 即多示例 (Multi-instance) 学习, 由 Dietterich 等在药物活性预测的研究中提出<sup>[42]</sup>. 具体思想是将训练样本视为包, 每个包由多个示例组成, 示例没有概念标记, 但每个训练包有一个概念标记. 如果包中至少有一个示例是正例, 则该包被标记为正包, 如果包中没有任何一个示例是正例, 即所有示例都是反例, 则该包被标记为反包. 张敏灵对基于神经网络的多示例学习进行了深入的研究, 并在自然场景图像检索领域中取得了良好的测试性能<sup>[43]</sup>.

### 2) 示教式 (指导式, Guidance)

示教式学习是由外部环境向系统提供一般性的指示或建议, 使系统不断完善的机器学习方法. 相关反馈是一种典型的指导式学习方法, 最先用于信

息检索中,其思想主要是尽可能通过迭代反馈和查询改善准确捕捉用户的需求.在缺乏一种可靠的将人的主观感知特征与图像的高层语义建模的构架下,用户的反馈提供了一种基于特定案例查询语义的途径.但是,用户的相关反馈结果仅仅产生与每个高层概念相关的少量的标注图像.为弥补这些不足,产生了很多机器学习技术,如单类学习(One-class learning)、激励学习(Active learning)和多向学习(Manifold learning)等<sup>[1]</sup>.为解决从小样本测试集学习的问题,Wu等提出了一种判别式EM算法,使用数据库中的未标注数据选择更多的判别特征<sup>[44]</sup>.Zhou等从相关反馈中利用多元可视化特征比较图像,提出了一种最优学习策略<sup>[45]</sup>.

### 3) 类比 (Analogy)

类比学习方法是通过对相似事物比较进行创造性学习的一种方法.互联网的膨胀提供了巨大的联机图像资源,因此在缺少已标注的训练图像数据集的情况下,可以使用大量的网络图像数据,来提高图像标注模型的鲁棒性<sup>[46]</sup>.类比学习的图像标注思想是在已标注好的网络图像集中寻找与需标注图像底层特征相似的图像,利用其语义标注词对未标注图像进行标注.Wang等<sup>[47]</sup>提出了一种新的脱离模型的图像标注方法,即基于数据驱动通过挖掘搜索结果来标注图像的方法.通过发现视觉和语义与其相似的搜索结果,从搜索结果的文本描述中识别显著词的挖掘过程.该方法由于脱离模型约束,因此不需要训练数据提供可用于标注的无限词汇,同时对于离群数据(Outliers,也称为孤立点)也具有鲁棒性.但是,由于检索到相似的图像数据可能很多,冗余和噪声数据也可能很多.所以,可以通过用户或者专家经验先将挖掘后的语义标注词聚类后再删除冗余的词汇,或利用Wordnet删除一些冗余的标注词以进一步改善图像标注的性能.

为解决语义标注的“瓶颈”问题,大多数标注模型通常将统计方法与机器学习策略技术相结合.Chen等将多示例学习技术应用于基于区域的图像语义标注中.首先将每幅图像表示成一个区域集,并将其映射为一个包特征空间中的点,然后利用SVM对包特征空间进行训练,从而取得了良好的标注效果<sup>[48]</sup>.Yang等将基于区域的图像语义自动标注问题转化为多示例学习问题<sup>[49]</sup>,通过设计非对称的SVM对图像语义自动标注的多示例学习进行研究.Carneiro等提出了一种监督多类标注(Supervised multi-class labeling, SML)方法,应用多示例学习原理计算每个语义类的图像特征概率密度,并使用Bayes分类器计算测试图像的后验概率,用于标注图像<sup>[50]</sup>.

## 3 粒计算理论及其在图像语义标注过程中的应用

人们在解决问题时往往将复杂的问题分解为简单的问题,或者将难以解决的问题转化为简单问题,这都反映了人们从不同的粒度观点去分析和解决问题.自1979年Zadeh在基于模糊集(Fuzzy set)的基础上首次提出信息粒度化的概念以来,信息粒度化的思想就吸引了众多研究者.Hobbs等于1985年提出了“粒度(Granularity)”,分析了粒的分解和合并以及产生不同大小粒的模型<sup>[51]</sup>.1996年,Lin提出了要进行粒计算(Granular computing, GRC)的研究,并发表了与粒计算相关的论文<sup>[52]</sup>.Yao结合邻域系统将粒计算应用于知识挖掘等领域,提出由所有划分构成的格可用于求解统一分类问题,为数据挖掘求解问题提供了新的途径<sup>[53]</sup>.我国较早的粒计算研究学者Zhang等在出版的专著<sup>[54]</sup>中引用粒度概念并发表了一系列相应的论文,从代数格的角度研究如何产生不同粗细的粒度空间,并讨论不同粒度空间上的粒之间的函数关系<sup>[55]</sup>.目前,粒计算主要应用于机器学习、数据挖掘、智能数据处理和粒逻辑等方面<sup>[56-57]</sup>.

### 3.1 粒计算

粒计算是人工智能领域的研究热点,采用一种分而治之和层次求解的思想,其理论和模型能使计算机从多层次、多角度分析和解决问题,实现多层次、多粒度间灵活自如地切换,为复杂问题的求解提供了新的思路<sup>[56]</sup>.粒计算解决问题的思想是先将问题抽象出一个个粒子,即粒化、分解的过程,然后将粒子按照某种求解或推理规则转化或者合并以得到整个问题的解,即粒的计算过程.一般的粒计算模型<sup>[3]</sup>有词计算(Computing with words)<sup>[58]</sup>、粗糙集(Rough set)<sup>[59]</sup>、商空间(Quotient space)<sup>[60]</sup>和概念格(Concept lattice)<sup>[61]</sup>等.词计算用词语代替数进行计算和推理,是一种基于自然语言的粒计算,在信息粒度、语言变量和约束概念上存在一定的理论和方法,主要解决模糊集合论的数值化隶属度函数表示法的不足,使其能更符合人类的逻辑思维.粗糙集用上近似(Upper approximation)和下近似(Lower approximation)两个精确集(Crisp set)近似表示一个概念.商空间理论由张钹等提出<sup>[54, 60]</sup>,以商集作为不同粒度世界的数学模型,由于问题的不同粒度表示对应于不同的等价关系,因而商空间粒度计算对论域进行不同的划分.基于概念格的粒度分析方法实际上是一种层次分析方法,将概念格的结点概念作为粒子,通过概念的内涵或外延的合取(或者析取)求解问题.

### 3.2 粒计算模型在图像语义标注过程中的应用

由于粒度分析是解决复杂问题的有效方法, 而且图像理解又存在图像低层特征维数高、高层语义词之间关系复杂以及图像低层视觉特征与高层语义映射之间有鸿沟等问题, 因此如何利用粒化思想将标注问题有效转化为求解较简单的问题, 即采用粒度分析的方法, 建立有效的图像语义自动标注模型, 提高图像语义检索的效率, 成为图像领域专家的一个关注点. 目前, 研究者在图像语义标注的各阶段都使用了粒分析思想.

#### 1) 特征提取与分割

图像特征提取采用的分割技术是图像标注过程中体现粒度思想求解最多的方面. 在目标检测中, 首先将图像分割成不同的区域, 待检测的目标通常可以认为是这些区域的某种组合. 过细的分割会增加检测的难度和复杂度, 过粗的分割会导致错误的检测结果<sup>[62]</sup>. 可见, 基于不同粒度的图像区域, 提取特征的效果、图像标注的精度以及检索的效率不同, 即图像分割技术和分割粒度的粗细影响最终的标注性能及其复杂性. 修保新等给出了一种基于模糊信息粒化思想的图像边缘检测方法<sup>[63]</sup>. 图像分割中也常利用粗糙集来提高图像分割的效果<sup>[64]</sup>, Malyszko 等提出一种基于粗糙熵的图像分割阈值技术, 将基于熵的阈值与粗糙集的结果结合起来, 给出了一种基于粒度多层粗糙熵进化阈值算法, 该算法可产生比 K-均值更好的图像分割聚类效果<sup>[65]</sup>. Pal 等利用图像粒度概念定义了图像的粗糙熵, 提出了一种利用粗糙集提取图像中的对象方法<sup>[66]</sup>. 刘仁金等深入分析了图像分割问题中的商空间粒度原理, 构造图像分割的商空间粒计算模型和分割方法, 实现了复杂纹理图像的成功分割<sup>[67]</sup>.

#### 2) 视觉内容描述粒度

图像的视觉内容, 如图像像素的空间关系和亮度、形状的规则程度和纹理的类型等, 都可从粒度角度形象化表示<sup>[68]</sup>. Hildebrand 等以自然语言的方式, 即词语来描述图像的颜色, 提出了一种基于知识的模糊颜色处理方法<sup>[69]</sup>. Hirota 等将模糊信息粒化思想应用于图像压缩领域, 用模糊关系描述静态灰度图像, 提出了一种基于模糊关系计算的图像压缩方法<sup>[70]</sup>. Zheng 等提出了相容粒度空间模型<sup>[71]</sup>, 并在图像纹理识别和数据挖掘中取得了成功. 李清勇等针对 Tamura 纹理模型, 提出了一种基于语言变量的图像纹理语义特征描述方法, 并通过遗传程序构造从低层视觉特征到高层语义特征的映射<sup>[72]</sup>.

#### 3) 语义粒度<sup>[73]</sup>

图像语义标注词由于本身词义存在上下(或包含)关系, 因而存在层次关系, 即粒度关系. 如一幅图像标注成泛化概念上的交通工具还是具体细化的

自行车或汽车, 对用户造成的理解程度不一样, 同时用户检索时获取所需求的图像集也不一样. 目前, 层次语义标注大多采用 Wordnet 的概念层次结构, 综合分析检索关键词在不同抽象层次上的语义信息, 计算图像间的语义相似距离, 以解决因用户主观理解不同而图像语义标注词不一致所产生的歧义问题. 陈世亮等<sup>[74]</sup>针对基于关键词的图像检索方式存在因用户理解差异而使得导致对图像语义理解的歧义问题, 利用 Wordnet 词典中单词间的同义关系、上下位关系, 提出了一种基于多层语义相似性度量的图像检索方法.

#### 4) 标注粒度

语义标注模型建立的过程即分类函数的建立, 也吸引了不少研究者. 其粒度思想是先建立粒度不同的分类模型, 然后利用粒度理论合成最终的标注模型. 张向荣等将商空间粒度计算引入 SAR 图像的分类中, 首先提取不同的纹理特征, 利用 SVM 得到不同的分类结果, 再基于商空间粒度合成技术实现信息融合, 从而提高 SAR 图像分类精度<sup>[75]</sup>. 许相莉等将商空间粒度计算理论引入图像检索领域, 针对彩色图像在不同粒度下的表现得到不同粒度下的颜色特征, 获得不同的商空间, 然后根据商空间粒度合成原理对已得到的商空间进行合成<sup>[76]</sup>. Fan 等通过使用具有标注主题的 SVM 从已标注的图像区域搜索最优参数, 学习一组测试函数, 为了产生图像语义概念, 使用有限的混合模型近似相关敏感概念显著对象的类分布, 提出一种自适应 EM 算法来确定最优模型结构和参数模拟, 进而实现了多层次语义标注<sup>[77]</sup>. 许宏丽等利用小波变换的多尺度特性对图像特征分布曲线进行不同尺度的小波变换, 去除一些小的分类和可能的噪声干扰, 从而得到不同粒度下的层次聚类标注模型<sup>[78]</sup>. 这些聚类方法一般均先利用各粒度的图像底层特征进行分类, 再利用粒度原理进行合成, 进而提高语义标注的精度, 但是这些粒度分析方法在融合分类语义时, 各底层特征与高层语义结合比较松散.

### 3.3 粒计算标注模型与其他标注模型的分析比较

基于统计方法的标注模型一般通过统计概率的方法标注图像的语义, 因此标注的精度高, 但是对大量高维图像数据标注的复杂性也较高. 基于机器学习的标注模型大多因采用了学习策略, 图像标注的复杂性较低, 但其标注的精度也不是很理想. 与这两种模型相比, 粒计算标注模型更注重综合考虑图像标注的精度和复杂性, 将层次分析的思想应用于图像语义标注的各阶段. 不仅可以大量高维图像语义标注问题进一步简化, 降低图像语义标注的复杂性, 而且不同粒度级的标注精度也可以满足用户从不同角度检索语义的需求, 最终提高图像语义标注

的性能。

#### 4 存在的问题和进一步的研究方向

图像语义自动标注模型的建立主要涉及图像底层视觉特征提取、视觉内容表示、语义联系、底层视觉特征和高层语义映射的建立等方面,存在的主要问题有:

##### 1) 高维底层特征的维数约简和语义鸿沟

高维底层视觉特征一直是影响图像语义标注效率的重要因素,如何利用人工智能和数据挖掘工具,有效地降低视觉特征维数,提取主要特征维,并且不丢失主要视觉特征与高层语义的联系,进一步缩小底层视觉特征与高层语义的鸿沟,提高图像标注效率和精度,是一个值得研究的主题。

##### 2) 图像标注的粒度

视觉图像语义分类用计算机模拟人类对视觉图像的语义理解,而从哲学角度看,“人们观察、度量、定义和推理的实体都是粒度”<sup>[79-80]</sup>,因此借助粒计算方法分析图像语义标注问题更接近人类认知问题的层次。图像语义标注的各阶段都存在粒度问题,例如,大多数语义标注方法的图像特征提取时,通过对原始图像进行分割,获取像素点阵,因此,图像的区域分割存在粒度问题,即“粗细”的问题。又如,图像特征表示方式的视觉词典的大小存在着粒度问题,视觉词集大则标注不适应噪声数据的处理,且标注复杂性高,视觉词集小,泛化能力强,但标注效果不好。可见,利用一种有效的粒分析方法分析并约简词集可进一步提高图像语义标注性能。同样,一些高层语义词之间也存在层次问题,即图像高层语义也存在“粒度”问题,某些语义词就比其他语义词所表示概念的泛化程度强。例如,一幅图像其语义表示为自然景色、建筑物或者楼房,其中自然景色、建筑物为粗粒度语义,而楼房则为细粒度语义表示。而且图像语义标注的映射同样存在粒度问题,粒度较细的分类函数可以较准确地反映图像的内容,但复杂性高,而粒度较粗的分类函数其检索的效率不高。因此,如何利用粒计算方法建立一种有效的图像语义自动标注模型从而提高图像语义检索效率是一个值得研究的内容。

3) 稀疏样本标注和网络图像海量信息有趣知识的挖掘、冗余及噪声

当已标注的样本数据很少时,所建立的自动标注模型标注精度不高,然而随着网络资源中图像数据资源的剧增,很容易搜索到与待标注图像相似的已标注图像。因此,利用数据挖掘技术,通过挖掘隐藏在待标注图像和已标注的网络图像的语义词之间的关系,是一种提高图像语义标注精度的有效途径。然而在对挖掘到与其相似的图像语义标注词进行重

组时,由于图像相似度量方法的选择影响需要重组的图像的语义,势必会产生一些冗余和噪声数据。例如标识一幅以海边沙滩为主题的图像,挖掘到与其语义相似并重组后的语义词为沙滩、人、海、雪,其中,雪为噪声数据。因此,基于网络图像海量信息有趣知识挖掘的标注过程所产生的噪声和冗余数据的剪枝问题也是一个值得研究的主题。

##### 4) 跨媒体图像的语义标注

在数字媒体时代,随着电子计算机的普及特别是互联网的发展,各种跨媒体图像服务得到了极大发展。以网页为中心,综合文本、图形、图像、音频、视频和动画等跨媒体图像数据成为第一大类,并且数量还在急剧增长中。如何有效地融合跨媒体图像数据特征,建立一种统一的跨媒体图像语义标注模型,将会成为一个新的研究热点。

综上所述,粒度分析是一种层次的、有效的数据分析方法,在图像分割和底层特征提取方面已取得一些研究成果。但是在语义标注模型的建立方面,由于粒度分析方法在融合分类语义时,各底层特征与高层语义结合比较松散,标注精度有待提高。而且,目前很多用于图像语义标注的粒模型方法没有充分挖掘隐藏在图像数据之间的关联关系,因此标注效率不高。基于形式背景构造的概念格是一种粒度分析模型,又是一种有效的数据挖掘工具,在很多方面已得到成功应用,但是,将概念格用于图像语义标注的研究较少。在概念格结构中,概念之间的父子关系是一种典型的层次关系,每个概念可以看作最小的知识粒(概念粒),而且每个知识粒的内涵本质上也是一个最大项目集,知识粒的上下层之间存在明显的关联关系,概念格就是通过知识粒之间的泛化和特化关系表示形式背景上层结构化的知识。因此将概念格的每个结点看作一个由图像的底层特征及相关语义组成的粒子,通过粒子的合并(重构),分析隐藏在其中图像语义的关联关系、层次关系,进而建立一种有效的图像自动标注模型,提高自动标注的精度和效率,是一个值得研究的新领域。

#### 5 结束语

缩小图像底层特征与高层语义标注的语义鸿沟,提高和改善图像语义标注性能已经吸引了很多研究者。本文首先对一些主要的图像自动标注模型进行综述和讨论,分析了各自的优势和缺陷;其次,阐述了粒度分析方法的思路、模型,并对当前粒计算模型在图像语义标注各阶段中的应用和存在的问题及发展趋势做了分析和探讨。最后,探索了以粒分析方法为基础的图像语义自动标注相关研究领域的重要课题和下一步的研究方向。

## References

- 1 Datta R, Joshi D, Li J, Wang J Z. Image retrieval: ideas, influences, and trends of the new age. *ACM Computing Surveys*, 2008, **40**(2): 1–60
- 2 Zadeh L A. Fuzzy sets and information granularity. *Advances in Fuzzy Set Theory and Applications*. Amsterdam: North-Holland Publishing, 1979. 3–18
- 3 Wang Guo-Yin, Zhang Qing-Hua, Hu Jun. An overview of granular computing. *CAAI Transactions on Intelligent Systems*, 2007, **2**(6): 8–26  
(王国胤, 张清华, 胡军. 粒计算研究综述. 智能系统学报, 2007, **2**(6): 8–26)
- 4 Yavilinsky A, Schofield E, Ruger S. Automated image annotation using global features and robust nonparametric density estimation. In: Proceedings of the 4th International Conference on Image and Video Retrieval. Singapore, Singapore: Springer, 2005. 507–517
- 5 Feng S L, Manmatha R, Lavrenko V. Multiple Bernoulli relevance models for image and video annotation. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2004. 1002–1009
- 6 Bi J, Chen Y, Wang J Z. A sparse support vector machine approach to region-based image categorization. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005. 1121–1128
- 7 Zhang Q, Goldman S A, Yu W, Fritts J. Content-based image retrieval using multiple-instance learning. In: Proceedings of the 19th International Conference on Machine Learning. Sydney, Australia: Morgan Kaufmann, 2002. 682–689
- 8 Wang Mei. Research on Automatic Image Annotation Based on Multi-label Learning [Ph. D. dissertation], Fudan University, China, 2008  
(王梅. 基于多标签学习的图像语义自动标注研究 [博士学位论文], 复旦大学, 中国, 2008)
- 9 Gao Jun, Xie Zhao. *Theory and Method of Image Understanding*. Beijing: Science Press, 2009  
(高隽, 谢昭. 图像理解理论与方法. 北京: 科学出版社, 2009)
- 10 Swain M J, Ballard D H. Color indexing. *International Journal of Computer Vision*, 1991, **7**(1): 11–32
- 11 Lowe D G. Object recognition from local scale-invariant features. In: Proceedings of the 7th IEEE international Conference on Computer Vision. Kerkyra, Greece: IEEE, 1999. 1150–1157
- 12 Perronnin F. Universal and adapted vocabularies for generic visual categorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, **30**(7): 1243–1256
- 13 Quélhas P, Monay F, Odobez J M, Gatica-Perez D, Tuytelaars T. A thousand words in a scene. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, **29**(9): 1575–1589
- 14 Yang Heng, Wang Qing. A novel local invariant feature detection and description algorithm. *Chinese Journal of Computers*, 2010, **33**(5): 935–944  
(杨恒, 王庆. 一种新的局部不变特征检测和描述算法. 计算机学报, 2010, **33**(5): 935–944)
- 15 Liu Shuo-Yan, Xu De, Feng Song-He, Liu Di, Qiu Zheng-Ding. A novel visual words definition algorithm of image patch based on contextual semantic information. *Acta Electronica Sinica*, 2010, **38**(5): 1156–1161  
(刘硕妍, 须德, 冯松鹤, 刘镭, 裘正定. 一种基于上下文语义信息的图像块视觉单词生成算法. 电子学报, 2010, **38**(5): 1156–1161)
- 16 Li F F, Pietro P. A Bayesian hierarchical model for learning natural scene categories. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005. 524–531
- 17 Jiang Yue, Wang Run-Sheng, Wang Cheng. Scene classification with context pyramid features. *Journal of Computer-Aided Design and 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. 计算机辅助设计与图形学学报, 2010, **22**(8): 1366–1373)
- 18 Perronnin F, Dance C, Csúrká G, Bressan M. Adapted vocabularies for generic visual categorization. In: Proceedings of the 9th European Conference on Computer Vision. Graz, Austria: Springer, 2006. 464–475
- 19 Jin Y, Khan L, Prabhakaran B. Knowledge based image annotation refinement. *Journal of Signal Processing Systems*, 2010, **58**(3): 387–406
- 20 Lu Han-Qing, Liu Jing. Image annotation based on graph learning. *Chinese Journal of Computers*, 2008, **31**(9): 1629–1639  
(卢汉清, 刘静. 基于图学习的自动图像标注. 计算机学报, 2008, **31**(9): 1629–1639)
- 21 Lu Z W, Ip H H S. Combining context, consistency, and diversity cues for interactive image categorization. *IEEE Transactions on Multimedia*, 2010, **12**(3): 194–203
- 22 Duygulu P, Barnard K, Freitas J F G, Forsyth D A. Object recognition as machine translation: learning a lexicon for a fixed image vocabulary. In: Proceedings of the 7th European Conference on Computer Vision. Copenhagen, Denmark: Springer, 2002. 97–112
- 23 Deerwester S, Dumais S T, Furnas G W, Landauer T K, Harshman R. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 1990, **41**(6): 391–407
- 24 Hofmann T. Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning*, 2001, **42**(1–2): 177–196
- 25 Dempster A P, Laird N M, Rubin D B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 1977, **39**(1): 1–38
- 26 Blei D M, Ng A Y, Jordan M I. Latent Dirichlet allocation. *Journal of Machine Learning Research*, 2003, **3**: 993–1022
- 27 Jeon J, Lavrenko V, Manmatha R. Automatic image annotation and retrieval using cross-media relevance models. In: Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Toronto, Canada: ACM, 2003. 119–126
- 28 Lavrenko V, Manmatha R, Jeon J. A model for learning the semantics of pictures. In: Proceedings of the Neural Information Processing Systems. Vancouver, Canada: MIT Press, 2004. 553–560
- 29 Guo P, Jia Y D, Lyu M R. A study of regularized Gaussian classifier in high-dimension small sample set case based on MDL principle with application to spectrum recognition. *Pattern Recognition*, 2008, **41**(9): 2842–2854
- 30 Li L J, Su H, Lim Y, Li F F. Objects as attributes for scene classification. In: Proceedings of the 12th European Conference of Computer Vision, the 1st International Workshop on Parts and Attributes. Crete, Greece, 2010.
- 31 Dong A, Bhanu B. Active concept learning for image retrieval in dynamic databases. In: Proceedings of the 9th IEEE International Conference on Computer Vision. Nice, France: IEEE, 2003. 90–95



- 32 Shi R, Chua T S, Lee C H, Gao S. Bayesian learning of hierarchical multinomial mixture models of concepts for automatic image annotation. In: Proceedings of the 5th Conference on Image and Video Retrieval. Tempe, USA: Springer, 2006. 102–112
- 33 Barnard K, Forsyth D. Learning the semantics of words and pictures. In: Proceedings of the 8th IEEE International Conference on Computer Vision. Vancouver, Canada: IEEE, 2001. 408–415
- 34 Bosch A, Zisserman A, Munoz X. Scene classification using a hybrid generative/discriminative approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, **30**(4): 712–727
- 35 Wang Yong-Qing. *Principle and Method of Artificial Intelligence*. Xi'an: Xi'an Jiaotong University Press, 1998 (王永庆. 人工智能原理与方法. 西安: 西安交通大学出版社, 1998)
- 36 Li Zhi-Xin, Shi Zhi-Ping, Li Zhi-Qing, Shi Zhong-Zhi. A survey of semantic mapping in image retrieval. *Journal of Computer-Aided Design and Computer Graphics*, 2008, **20**(8): 1085–1096 (李志欣, 施智平, 李志清, 史忠植. 图像检索中语义映射方法综述. 计算机辅助设计与图形学学报, 2008, **20**(8): 1085–1096)
- 37 Xia Li-Min, Tan Li-Qiu, Zhong Hong. Semantic annotation of image based on information bottleneck method. *Pattern Recognition and Artificial Intelligence*, 2008, **21**(6): 812–818 (夏利民, 谭立球, 钟洪. 基于信息瓶颈算法的图像语义标注. 模式识别与人工智能, 2008, **21**(6): 812–818)
- 38 Frey B J, Dueck D. Clustering by passing message between data Points. *Science*, 2007, **315**(5814): 972–976
- 39 Dueck D, Frey B J. Non-metric affinity propagation for unsupervised image categorization. In: Proceedings of the 11th IEEE International Conference on Computer Vision. Rio de Janeiro, Brazil: IEEE, 2007. 1–8
- 40 Yang D, Guo P. Improvement of image modeling with affinity propagation algorithm for semantic image annotation. In: Proceedings of the 16th International Conference on Neural Information Processing. Bangkok, Thailand: Springer, 2009. 778–787
- 41 Li W, Sun M S. Automatic image annotation based on wordnet and hierarchical ensembles. In: Proceedings of the 7th International Conference on Computational Linguistics and Intelligent Text Processing. Mexico City, Mexico: Springer, 2006. 417–428
- 42 Dietterich T G, Lathrop R H, Lozano-Perez T. Solving the multiple instance problem with axis-parallel rectangles. *Artificial Intelligence*, 1997, **89**(1–2): 31–71
- 43 Zhang Min-Ling. Research on Multi-instance Learning and Multi-label Learning [Ph. D. dissertation], Nanjing University, China, 2007 (张敏灵. 多示例与多标记学习的研究 [博士学位论文], 南京大学, 中国, 2007)
- 44 Wu Y, Tian Q, Huang T S. Discriminant-EM algorithm with application to image retrieval. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Hilton Head Island, USA: IEEE, 2000. 222–227
- 45 Zhou X S, Huang T S. Relevance feedback in image retrieval: a comprehensive review. *Multimedia Systems*, 2003, **8**(6): 536–544
- 46 Li L J, Li F F. OPTIMOL: automatic online picture collection via incremental model learning. *International Journal of Computer Vision*, 2010, **88**(2): 147–168
- 47 Wang X J, Zhang L, Li X R, Ma W Y. Annotating images by mining image search results. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, **30**(11): 1919–1932
- 48 Chen Y X, Wang J Z. Image categorization by learning and reasoning with regions. *Journal of Machine Learning Research*, 2004, **5**: 913–939
- 49 Yang C B, Dong M, Hua J. Region-based image annotation using asymmetrical support vector machine-based multiple-instance learning. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE, 2006. 2057–2063
- 50 Carneiro G, Chan A B, Moreno P J, Vasconcelos N. Supervised learning of semantic classes for image annotation and retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, **29**(3): 394–410
- 51 Hobbs J R. Granularity. In: Proceedings of the 9th International Joint Conference on Artificial Intelligence. Los Angeles, USA: Morgan Kaufmann, 1985. 432–435
- 52 Lin T Y. Granular computing on binary relations I: data mining and neighborhood systems, II: rough sets representations and belief functions. *Rough Sets in Knowledge Discovery I: Methodology and Applications*. Heidelberg: Physica-Verlag, 1998. 107–140
- 53 Yao Y Y. Relational interpretations of neighborhood operators and rough set approximation operators. *Information Sciences*, 1998, **111**(1–4): 239–259
- 54 Zhang B, Zhang L. *Theory and Application of Problem Solving*. New York: North Holland, 1992
- 55 Zhang L, Zhang B. The quotient space theory of problem solving. In: Proceedings of the 9th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing. Chongqing, China: Springer, 2003. 11–15
- 56 Miao Duo-Qian, Wang Guo-Yin, Liu-Qing, Lin Zao-Yang, Yao Yi-Yu. *Granular Computing: Past, Present and Future*. Beijing: Science Press, 2007 (苗夺谦, 王国胤, 刘清, 林早阳, 姚一豫. 粒计算: 过去、现在与展望. 北京: 科学出版社, 2007)
- 57 Qiu Tao-Rong, Liu Qing, Huang Hou-Kuan. A granular computing approach to knowledge discovery in relational databases. *Acta Automatica Sinica*, 2009, **35**(8): 1071–1079 (邱桃荣, 刘清, 黄厚宽. 关系数据库中知识发现的一种粒计算方法. 自动化学报, 2009, **35**(8): 1071–1079)
- 58 Zadeh L A. Fuzzy logic = computing with words. *IEEE Transactions on Fuzzy Systems*, 1996, **4**(2): 103–111
- 59 Pawlak Z. *Rough Sets: Theoretical Aspects of Reasoning about Data*. Boston: Kluwer Academic Publishers, 1991
- 60 Zhang Ling, Zhang Bo. Theory of fuzzy quotient space (methods of fuzzy granular computing). *Journal of Software*, 2003, **14**(4): 770–776 (张铃, 张钊. 模糊商空间理论 (模糊粒度计算方法). 软件学报, 2003, **14**(4): 770–776)
- 61 Wille R. Formal concept analysis as mathematical theory of concepts and concept hierarchies. In: Proceedings of the Formal Concept Analysis. Berlin, Germany: Springer, 2005. 1–33
- 62 Wang Yan, Sun Yi. Adaptive mean shift based image smoothing and segmentation. *Acta Automatica Sinica*, 2010, **36**(12): 1637–1644 (王晏, 孙怡. 自适应 Mean Shift 算法的彩色图像平滑与分割算法. 自动化学报, 2010, **36**(12): 1637–1644)

- 63 Xiu Bao-Xin, Wu Meng-Da. Adaptability measure to fuzzy information granule on image and its application to edge detection. *Acta Electronica Sinica*, 2004, **32**(2): 274–277 (修保新, 吴孟达. 图像模糊信息粒的适应性度量及其在边缘检测中的应用. *电子学报*, 2004, **32**(2): 274–277)
- 64 Yue Xiao-Dong, Miao Duo-Qian, Zhong Cai-Ming. Roughness measure approach to color image segmentation. *Acta Automatica Sinica*, 2010, **36**(6): 807–816 (岳晓冬, 苗夺谦, 钟才明. 基于粗糙性度量的彩色图像分割方法. *自动化学报*, 2010, **36**(6): 807–816)
- 65 Malyszko D, Stepaniuk J. Adaptive multilevel rough entropy evolutionary thresholding. *Information Sciences*, 2010, **180**(7): 1138–1158
- 66 Pal S K, Shankar B U, Mitra P. Granular computing, rough entropy and object extraction. *Pattern Recognition Letters*, 2005, **26**(16): 2509–2517
- 67 Liu Ren-Jin, Huang Xian-Wu. The granular theorem of quotient space in image segmentation. *Chinese Journal of Computers*, 2005, **28**(10): 1680–1685 (刘仁金, 黄贤武. 图像分割的商空间粒度原理. *计算机学报*, 2005, **28**(10): 1680–1685)
- 68 Pedrycz W, Loia V, Senatore S. Fuzzy clustering with viewpoints. *IEEE Transactions on Fuzzy Systems*, 2010, **18**(2): 274–284
- 69 Hildebrand L, Fathi M. Knowledge-based fuzzy color processing. *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews*, 2004, **34**(4): 499–505
- 70 Hirota K, Pedrycz W. Fuzzy relational compression. *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, 1999, **29**(3): 407–415
- 71 Zheng Z, Hu H, Shi Z Z. Granulation based image texture recognition. In: Proceedings of the 4th International Conference on Rough Sets and Current Trends in Computing. Uppsala, Sweden: Springer, 2004. 659–664
- 72 Li Qing-Yong, Hu-Hong, Shi Zhi-Ping, Shi Zhong-Zhi. Research on texture-based semantic image retrieval. *Chinese Journal of Computers*, 2006, **29**(1): 116–122 (李清勇, 胡宏, 施智平, 史忠植. 基于纹理语义特征的图像检索研究. *计算机学报*, 2006, **29**(1): 116–122)
- 73 Wang Hui-Feng, Sun Zheng-Xing, Wang Jian. Semantic image retrieval: review and research. *Journal of Computer Research and Development*, 2002, **39**(5): 513–523 (王惠锋, 孙正兴, 王箭. 语义图像检索研究进展. *计算机研究与发展*, 2002, **39**(5): 513–523)
- 74 Chen Shi-Liang, Li Zhan-Huai, Yuan Liu. A novel semantics-based image retrieval method using similarity measure of multi-level semantics. *Journal of Northwestern Polytechnical University*, 2008, **26**(5): 588–591 (陈世亮, 李战怀, 袁柳. 一种基于多层语义相似性度量的图像检索方法. *西北工业大学学报*, 2008, **26**(5): 588–591)
- 75 Zhang Xiang-Rong, Tan Shan, Jiao Li-Cheng. SAR image classification based on granularity computing of quotient space theory. *Chinese Journal of Computers*, 2007, **30**(3): 483–490 (张向荣, 谭山, 焦李成. 基于商空间粒度计算的 SAR 图像分类. *计算机学报*, 2007, **30**(3): 483–490)
- 76 Xu Xiang-Li, Zhang Li-Biao, Yu Zhe-Zhou, Zhou Chun-Guang. Application of multi-granularity color features in image retrieval. *Journal of Applied Sciences*, 2009, **27**(1): 56–61 (许相莉, 张利彪, 于哲舟, 周春光. 多粒度颜色特征在图像检索中的应用. *应用科学学报*, 2009, **27**(1): 56–61)
- 77 Fan J P, Gao Y L, Luo H Z, Xu G Y. Statistical modeling and conceptualization of natural image. *Pattern Recognition*, 2005, **38**(6): 865–885

- 78 Xu Hong-Li, Xu De, Lin En-Ai. An approach of hierarchical image index based on subspace cluster. *Journal of Image and Graphics*, 2009, **14**(1): 142–147 (许宏丽, 须德, 林恩爱. 一种基于子空间聚类的图像分层索引方法. *中国图象图形学报*, 2009, **14**(1): 142–147)
- 79 Yao Y Y. Granular computing: basic issues and possible solutions. In: Proceedings of the 5th Joint Conference on Information Sciences. New Jersey, USA, 2000. 186–189
- 80 Li Dao-Guo, Miao Duo-Qian, Zhang Dong-Xing, Zhang Hong-Yun. An overview of granular computing. *Computer Science*, 2005, **32**(9): 1–12 (李道国, 苗夺谦, 张东星, 张红云. 粒度计算研究综述. *计算机科学*, 2005, **32**(9): 1–12)



张素兰 太原科技大学计算机科学与技术学院副教授, 北京理工大学博士研究生. 主要研究方向为概念格与数据挖掘, 粒度计算, 图像处理. 本文通信作者.  
E-mail: zhsulan@126.com

(ZHANG Su-Lan Associate professor at the School of Computer Science and Technology, Taiyuan University of Science and Technology, Ph. D. candidate at Beijing Institute of Technology. Her research interest covers concept lattice and data mining, granular computing, and image processing. Corresponding author of this paper.)



郭平 北京理工大学计算机学院教授. 主要研究方向为图像处理, 模式识别, 神经网络, 软件可靠性.  
E-mail: pguo@pku.org.cn

(GUO Ping Professor at the School of Computer Science and Technology, Beijing Institute of Technology. His research interest covers image processing, pattern recognition, neural network, and software reliability engineering.)



张继福 太原科技大学计算机科学与技术学院教授. 主要研究方向为数据挖掘, 模式识别, 智能信息系统.  
E-mail: jifuzh@sina.com

(ZHANG Ji-Fu Professor at the School of Computer Science and Technology, Taiyuan University of Science and Technology. His research interest covers data mining, pattern recognition, and intelligence information system.)



胡立华 太原科技大学计算机科学与技术学院讲师, 主要研究方向为概念格与数据挖掘.  
E-mail: hulihua1982@sina.com

(HU Li-Hua Lecturer at the School of Computer Science and Technology, Taiyuan University of Science and Technology. Her research interest covers concept lattice and data mining.)