

# 基于模糊熵的空间语义图像检索模型研究<sup>\*</sup>

蒋华伟

(河南工业大学 信息科学与工程学院, 郑州 450001)

**摘要:** 根据模糊熵理论和改进的空间信息分布,提出了颜色空间特征语义图像检索模型。阐述基于语法规则的颜色空间特征语义描述方法,构造从低层颜色空间特征到高层语义之间的映射,根据这些模糊语义值进行图像检索。实验结果表明,该模型能够有效地对图像高层语义进行刻画,由此实现的模型不仅能获得高效和稳定的检索结果,获得与人类视觉感知较好的一致性,该算法还能很好地消除低层图像空间特征和高层语义之间的语义鸿沟。

**关键词:** 模糊信息熵; 基于内容的图像检索; 空间特征语义; 语言变量

**中图分类号:** TP392; TP311.52      **文献标志码:** A      **文章编号:** 1001-3695(2010)01-0353-05

**doi:** 10.3969/j.issn.1001-3695.2010.01.106

## Research of spatial semantic image retrieval model based on fuzzy entropy

JIANG Hua-wei

(School of Information Science & Engineering, Henan University of Technology, Zhengzhou 450001, China)

**Abstract:** According to the theory of fuzzy entropy and the improved spatial information distribution, the paper put forward the novel model of color spatial features semantic retrieval. Set the color spatial semantic features description method forth based on grammar rule, constructed the mapping from low-level color spatial features to high-level semantic features successfully, and proposed the fuzzy retrieval algorithm according to the extracted semantic features. The experimental results show that the model is effective in characterizing image high-level semantic content, the approach not only has an excellent retrieval precision and a good accordance with human visual perception, but also has strong significance for reducing the semantic gap between the image spatial features and semantic features.

**Key words:** fuzzy information entropy; content-based image retrieval; spatial feature semantic; linguistic variable

## 0 引言

随着 Internet 和多媒体技术的快速发展,以及网络共享数据的涌现,使得人们的信息来源不再局限于原有的简单文本数据,而不断扩展到图形、图像等多种不同类型的媒体对象,这使得基于传统范式(normal forms, NF)的结构化语句查询,在数据库管理系统(DBMS)检索中遇到了前所未有的挑战,即单纯采用文本的数据查询方式难以胜任当今大量出现的基于内容的图像等多种媒体数据的检索。虽然,目前国内外进行了关系数据库模式和语义本体之间的转换研究<sup>[1]</sup>,但由于关系数据库模式与本体之间在建模思想、实现方法、应用场景、表达能力等方面的较大差距,两者不存在完美的兼容关系,使得转换方法和工具不完备,转换效果较差。

这样从 20 世纪 90 年代初期以来,逐渐形成了基于内容的大量多媒体索引和检索技术。在图像领域中,最有代表性的是基于内容的图像检索(CBIR),如 QBIC(query by image content)、Virage、Photobook 和 MARS(multimedia analysis and retrieval system)等。在基于内容的图像检索中,图像特征索引技术主要研究的是图像视觉和形象特征,包括颜色、纹理、对象形状等方面<sup>[2]</sup>。还有基于范例检索(example based retrieval)、基于草图检索(sketch based retrieval)和随机浏览及其组合的方

式。CBIR 的主要特点是利用图像本身包含的客观视觉特征,即体现在视觉感知的图像相似性,它不需要人来解释,也不需要或仅需要少量的人工干预,因而,在某些自动检索场合有所应用。

利用图像的底层视觉和形象特征为索引,对图像进行检索具有计算简单、性能稳定的特点,但目前这些特征都有一定的局限性,因为 CBIR 主要使用颜色、纹理、形状等视觉特征进行图像检索,随着研究的深入,人们发现 CBIR 系统存在以下三方面的不足:

- a) 低层可视特征与高层语义概念之间存在巨大鸿沟;
- b) 与人类视觉机制具有明显的不一致性;
- c) 颜色(直方图)、纹理、形状仅仅描述了图像的全局特性,丢失了空间分布信息。

例如,对于两个颜色直方图相似的图像来说,如果颜色的空间分布差别很大,图像的内容会很不相同,显然,还需要利用图像的颜色、纹理、形状特征结合空间分布特征来进行图像检索研究。

另外,为了解决图像简单视觉特征与用户检索语义概念之间的语义鸿沟,在 CBIR 技术领域进行包含语义的检索方法的研究也是十分必要的<sup>[3]</sup>,这样就开始了语义图像检索的研究,以实现更为贴近用户理解能力的自然而简洁的查询方式,来提

收稿日期: 2009-05-08; 修回日期: 2009-06-12      基金项目: 河南省教育厅自然科学基金资助项目(2008A520005); 博士基金资助项目(2006BS007)

作者简介: 蒋华伟(1970-),男,河南济源人,副教授,博士,主要研究方向为图像理解及信息检索(lhwcad@sohu.com)。

高图像检索的精度。因此,如何结合图像的空间分布关系从底层特征中自动提取语义特征,成为基于内容的图像检索领域的难题。

### 1 相关研究工作

随着图像语义特征在多媒体信息检索中的作用不断加强,近几年的研究逐渐转向基于区域和目标对象的高层图像语义描述特征的提取,尝试由图像的底层特征来推知高层次语义,从而使用高层语义特征计算图像相似程度,对此,已有许多学者在图像高层语义特征提取技术研究方面进行了大量的探索性工作。例如基于“a word is worth a thousand pictures”,Buxton<sup>[4]</sup>提出了用语言文词来描述图像语义的方法。类似地 Hermes 等人<sup>[5]</sup>在信息系统中使用相似性技术直接从户外场景中推导出场景的自然语言描述,然后利用文本检索技术来检索。文献[6]中 Mojsilovic 等人通过实验获得理解图像相似性的语义类别,用多维尺度(MDS)和分级聚类(HC)分析感知数据,基于这些分析建立理解图像相似性的语义分类方法,以及用这些数据去发现能很好描述每个分类的低级特征,并设计图像相似矩阵来表达结论。模拟图像将要分类特性以及测量它们的相似性。在文献[7]中 Sheikholeslami 等人试图通过组合图像的不同特征来提高基于内容的图像检索效率,但由于纹理、颜色、形状特征的计算方法不同,需要用不同的相似测量方法,导致综合不同特征的检索是非常困难的,便提出基于语义的聚类索引方法——SemQuery,用于不同图像特征的可视查询,基于不同特征将数据库中图像分类,每个语义图像簇包含一系列的描述图像不同特征的子簇,如果一个图像属于一个语义簇的所有不同子簇,则它包含在该语义簇中,还设计神经网络模型去融合不同特征的基本查询结果,以及提出查询处理策略来支持不同特征的可视化查询。另外在文献[8]Fung 等人提出了知识学习的图像块语义方法,将图像语义分为原始语义和场景语义两个层次,从而可以使用文本检索算法实现图像相似性匹配。还有在文献[9]中提出了一种图像语义综合概率描述模型(IPSM),该模型由多层图像特征和语义间的映射组成,用于描述和提取模型中语义分类和图像特征,在此基础上构建了语义图像检索方法,为提高检索精度,又结合相关反馈提出语义相关反馈算法。

上述方法虽然在某些方面能起到较好的检索效果,但这些研究或由于需要作大量的实验,或存在图像分割比较困难,或是聚类方法可能带来较大的偏差等,有可能出现或完全丢失了图像颜色的空间分布信息,因此检索效果和效率有待改进和提高。

为了利用图像颜色空间分布特征进行图像检索,文献[10]中 Pass 等人发现颜色直方图缺乏空间信息,在比较整体图像空间信息的直方图方法后,提出颜色相关矢量的方法 CCV(color coherence vector)作为图像的索引,根据它是否属于或不属于大块的相似颜色区的一部分,把每个像素分为相关或不相关情况,相比直方图,CCV's 具有较好的区分性能。文献[11]Huang Jing 提出了一个用于图像标注和索引的图像共生矩阵的新图像特征,这种特征提取了颜色的空间相关性,共生矩阵健壮的容差能力更适应由镜头缩放、不同视点等改变所导致的表征和象征。但这些算法的计算量相对比较大,制约了它们在图像检索中应用。新近的方法是基于颜色空间特征的检

索<sup>[12,13]</sup>,根据自动检测和区域的可视特征进行区域的图像检索系统研究,用无人监督粗糙检测的方法量化颜色分布的分类(LDQC)获得竞争聚类分类算法,并提出一种新的可变颜色的区域精确描述,即适应渐变颜色的分布(ADCS),来改善被检索区域的知觉相似性。这些方法虽然在一定程度上有效地提取了图像颜色的空间分布信息,但仍然存在颜色的量化问题及对光照变化敏感问题,更为重要的是缺乏特征语义对图像空间信息的智能理解,不能完全适合基于内容的图像检索需要。对于空间特征语义的研究,在文献[14]中有所提及,Zhuang 等人根据人工智能知识表示,提出了模糊布尔等模型,并给出距离度量算法,通过知识表示、推理学习机制和机器学习技术,采用统计的方法自动建立图像空间多维信息特征的索引结构,获得与颜色直方图特征对应的描述,来实现图像语义检索。

近年来,随着信息技术的快速发展,信息熵被广泛应用于图像处理中,如文献[15]对图像的融合难题采用联合熵和交互信息量来评价;文献[16]探讨了图像检索中前期工作,并采用模糊熵来处理相关问题;文献[17]延伸了 Chang 等人的方法,提出两种基于相对熵的阈值方法,即局部相对熵阈值(LRE)和相关平均熵阈值(JRE),它们产生的效果比 Chang 等人的方法好。文献[18,19]中提出了基于图像信息分布及空间熵的图像检索算法,用信息熵来实现图像的检索;在文献[19]中的算法采用信息分布熵对颜色等空间特征进行描述,在分析空间信息熵情况后,提出更适应人类视觉的加权空间分布熵提取颜色等分布特征的算法。

在已有的研究中缺乏将空间特征模糊信息熵与语义理解的有机结合;还有颜色的空间分布熵与图像中不同颜色的像素在不同区间的概率分布相关,即存在空间分布不同的颜色可能具有相同的空间分布熵情况。为解决这些问题,在上述相关研究基础上,同时考虑到人类视觉特点及熵的特性,本文提出空间语义特征图像检索模型,并利用改进的颜色空间信息分布熵方法表示图像空间特征;另外,设计了基于语言变量的图像空间语义特征描述方法,构造从低层视觉特征到高层语义特征的映射,根据这些模糊语义值进行图像检索。实验结果表明,该算法性能稳定,具有更好的检索效果。

## 2 图像空间分布模糊熵

### 2.1 图像熵的数学描述

Shannon 熵把信息定义为不确定性概念,使信息与概率统计关联起来,并把熵作为一个度量信息状态不确定性的尺度,提出了信息熵的概念。对于给定灰度级范围 $\{0,1,2,\dots,l-1\}$ 的图像直方图,其熵为

$$H_T = - \sum_{i=0}^{l-1} p_i \ln(p_i) \tag{1}$$

其中: $p_i$  为第  $i$  个灰度出现的概率。

设阈值  $t$  将图像划分为不同的类,则图像的总熵为

$$H(t) = \ln(P_t(1 - P_t)) + H_t/P_t + (H_T - H_t)/(1 - P_t) \tag{2}$$

其中: $P_t = \sum_{i=1}^t p_i, H_t = - \sum_{i=1}^t p_i \ln(p_i)$ 。

### 2.2 改进的空间分布熵

对于颜色空间直方图,不同颜色都将会在图像空间有一个信息分布直方图与之对应,当使用颜色空间分布图进行图像相似性检索时,会出现如下问题:a)空间直方图分布信息需要足

够的空间来存储;b)计算量大和占用资源较多,导致了图像查询效率低下。因此可结合熵的概念,用信息分布熵来描述颜色的空间分布特征。

根据上面信息熵的定义,结合图像颜色的空间分布直方图,给出颜色  $c$  在  $L$  重空间分布熵:

$$H'(c) = - \sum_{l=0}^{L-1} p_c(l) \ln(p_c(l)) \quad (3)$$

式(3)显示,空间信息分布熵表达了给定颜色总量在图像空间的平均分布趋势,空间信息分布熵越大,对应的该颜色总量在图像空间中分布得越分散;否则,这种颜色总量在图像空间中分布得越紧密。

但上面的信息熵会在图像颜色空间分布情况不同时,出现一样的概率分布结果,即简单的信息熵仅与图像空间中不同颜色特征是否出现有关,即与该颜色信息总量(像素)和图像总的信息量的比值相对应,它表达的是不同颜色在整体空间的分布特征,没有考虑局部子空间的不同分布。显然这种信息分布与人的视觉习惯有着较大差异,因为,人的视觉受颜色像素的空间分布疏密影响,对于给定图像中某一颜色来说,该颜色空间分布集中时,则会引起视觉的较大刺激,颜色空间分布不集中时,则不会引起视觉的较大刺激。为此,在借鉴文献[19]的基础上,采用改进的空间信息分布熵,即通过聚类把整个图像空间分为一定的区间( $L$ ),然后对于聚类算法使用爬山法确定子空间聚类,在这些区间分成若干小区间( $N$ ),获得不同信息粒度下的层次聚类,由此来研究适应人的视觉习惯的颜色空间信息分布。

在每个区间上具体处理时,采用模拟人的视觉特征权重系数  $W$  来反映该空间中不同颜色对图像的作用。对权重系数的选取考虑如下:由于子空间(质心)与颜色聚类质心(中心)的远近(距离)来决定权重系数,即子空间离颜色聚类质心较近的给较大的权重,距离颜色聚类质心较远的给较小的权重。这是因为距离颜色聚类质心比较近的子空间,在区间上该种颜色整幅图像中分布相对紧凑,对于人眼的刺激作用较大,一般选取较大的权重系数;相对而言,当距离颜色聚类质心比较远的子空间,区间上该种颜色整幅图像中分布相对疏散,对于人眼的刺激作用较小,一般选取较小的权重系数。

由此可以获得颜色  $c$  在整个图像空间的分布熵:

$$S(t) = - W \sum_{l=0}^{L-1} \sum_{n=0}^{N-1} p_{lc}(n) \ln(p_{lc}(n)) \quad (4)$$

其中: $p_{lc}(n)$ 为某大区间  $L$  内颜色  $c$  在小区间  $N$  的分布, $W$  为该区间内的权重系数。

### 3 图像空间特征语义

基于人的视觉感知和心理认知经验,在搜索图像时并不是根据图像的低层视觉特征(如颜色、形状、纹理以及空间关系等),而是根据搜索目标在意识中的语义描述进行。图像空间语义检索模型综合了底层特征语义、对象语义以及抽象概念来获得图像对象及其相互之间的空间关系,得到高层推理和语义描述。下面给出语言变量、空间语义特征、语义规则和句法规则设计的概述。

#### 3.1 图像特征语言变量和语义描述

描述图像空间语义的语言变量与数学上的含义有所不同,它是以自然语言或人工语言中的词或词组作为值的变量,而不

是使用数字来精确表达,因为用字或词组会更加灵活。例如,谈到身高时,当说“某某身矮”比说“某某 126cm”虽然更模糊一些,但前者表达了更加丰富的信息,从这个意义上讲,“身矮”就可以认为是变量身高的一个语言值,而变量身高就称做语言变量。

根据人类对空间关系的视觉感知和心理学研究发现:HSV颜色空间更符合人眼的视觉特征,它对应人眼色彩视觉特征的三要素,即色调(hue)、饱和度(saturation)和亮度(value)。因此,选择颜色直方图空间模型的分布特征作为图像语义特征,给出视觉系统更易于感知的空间模式特征,即色调、饱和度和亮度,在计算机处理时要进行RGB到HSV空间的线性变换。在基于空间的图像检索模型中,主要用的是色调直方图模型空间分布特征作为语义特征,用这三个特征当做语言变量来描述图像的空间语义。基于文献[20]给出下面的定义以及空间语义和规则设计,先以语言变量色调为例来阐述图像语义的语言变量描述方法。

对于色调空间模式的直方图分布,一般描述空间分布的色调词是“红色”“橙色”“黄色”等。针对上面检索描述中的空间信息特征的语言变量可以用一个五元组( $T, W(t), R, G, M$ )来表示,具体的语言变量可以用空间语义特征来形式化地描述如下:

a)  $T$  是语言变量的名称,用于描述空间特征的属性,色调即为此语言变量。

b)  $W(t)$  是语言变量  $T$  取值的语言值的集合,包括基本和扩展词集两部分。前一部分可根据研究对象而产生,在HSV颜色空间并考虑人的视觉特征,对于  $T$  基本词集可以定义为集合{红、橙、黄、绿、蓝、紫};后一部分可由句法规则生成。

c)  $R$  是表征描述研究对象的实体。在图像检索中, $R$  可以是图像本身(像素值)或图像的一个近似(图像特征向量)。

d)  $G$  是句法规则。用它生成丰富查询语义(扩展词集)的语言值,如对“红色”后加上“分散”就可以生成新的语言值“红色分散”。对句法规则的具体叙述见3.3节。

e)  $M$  是语义规则。根据图像空间特征构造语义  $R$  的模糊集,即获得论域  $R$  中每个特征符合该语言值的语义隶属度。

#### 3.2 空间语义规则

在空间语义特征检索模型中,关键是对语言变量  $T$  的表述,它由语义规则来完成设计变量  $T$  的语言值,获得图像  $I$  中空间特征的语义隶属度。这里的语义隶属度指的是计算机视觉上图像空间向量(色调、饱和度等)与语言变量  $T$  的语言值描述的符合程度,标准化隶属度区间为 $[0, 1]$ ,即根据人的视觉经验,通过构造语义规则,进行监督训练学习,获得对图像颜色空间特征的理解,并计算判断得出两者的符合程度(语义隶属度值)。

对语义规则的设计用数学方法具体描述为,给定训练集  $T_c = \{(I_j, \text{sub}_j) | j = 1, 2, \dots, n\}$ ,可构造图像空间分布特征向量与语言变量  $T$  的映射  $g$  如下:

$$g: T \rightarrow \text{sub}$$

其中: $I_j$ 表示图像, $T_i$ 表示空间分布特征向量, $\text{sub}_j$ 表示图像  $I_j$  对相应语言值的隶属度,取值范围 $[0, 1]$ 。

根据文献[19,21]简化改进后,对  $L$  个灰度级的图像,取隶属度函数  $g(t_{ij})$  为分段线性函数。

$g_c(T)$  为  $T$  对空间颜色特征  $C$  的隶属度函数,则由论域  $R$

中属于  $C$  的元素的隶属度构成了模糊集合  $C$ , 即  $C = \{g_c(1), g_c(2), \dots, g_c(n)\}, (n=1, 2, \dots)$ 。给定一幅  $m \times n$  维图像可描述为

$$I = \begin{bmatrix} g(t_{11}) & g(t_{12}) & \dots & g(t_{1n}) \\ g(t_{21}) & g(t_{22}) & \dots & g(t_{2n}) \\ \hat{u} & \hat{u} & \dots & \hat{u} \\ g(t_{m1}) & g(t_{m2}) & \dots & g(t_{mn}) \end{bmatrix} \quad (5)$$

其中:  $g(t_{ij})$  定义为图像  $I$  中第  $(i, j)$  像素具有某种特征的程度,  $i=1, 2, \dots, m, j=0, 2, \dots, n, 0 \leq g(i, j) \leq 1$ 。对于  $I_{mn}$  图像平面, 其模糊信息熵为

$$H(C) = \frac{1}{mn \ln(2)} \sum_{i=1}^m \sum_{j=1}^n H(g_c(t_{ij})) \quad (6)$$

其中:  $H(g_c(t))$  为 Shannon 函数。

### 3.3 空间句法规则

由 3.1 节可知变量  $T$  (如色调) 所取的语言值包括基本和扩展词集两部分, 前一部分在考虑人的视觉特征基础上, 可对研究对象枚举产生语言描述词; 后一部分的语言值是通过句法规则产生, 即通过句法规则作用在基本词集上形成新的语言值, 得到丰富基本词集和增强语言表达能力的目的。

对于图像空间, 若某种颜色分布的信息熵越大, 则表明该颜色的信息 (像素点) 在图像空间的分布越分散, 相反则很集中。这样对色调中某一颜色, 根据句法规则 (如“红色”) 后面加上程度限定词语“分散”就构成新的语言值 (“红色分散”), 它主要是根据人的视觉对颜色的灵敏度和感知经验, 结合一定的规则设计而生成。

在句法规则设计时, 为了更具严谨性和科学性, 本文采用扩展的巴科斯—瑙尔范式 (extended Backus-Naur form, EBNF) 来生成语法指定的值, 对语言变量色调的 EBNF 表达如下:

```
Query = "Hau" | "Reinforce" ;
Hau = "红色" | "橙色" | "黄色" | "绿色" | "蓝色" | "紫色" ;
Reinforce = "分散" | "集中" ;
```

根据 EBNF 范式可知, Query 的词义由“Hau”“Reinforce”可选构成, 在由“Hau”和“Reinforce”共同组成时, 它与普通的使用主谓结构形成句子的方法基本相同。Hau 是基础语言值集, Reinforce 是在 Hau 基础上形成的扩充词集 (起强调补充作用), 即它结合 Hau 的语义通过语义规则扩展生成的。Reinforce 中的每一个词对应一个语义生成算法。

## 4 图像空间特征语义检索

根据上面的语言变量的设计规则, 可以描述和抽取颜色空间模型中的另外两个特征属性 (对于亮度, 可以借鉴物理定义的客观量光强的计算)。不失一般性, 假设在图像检索中采用如下多特征属性的多个语言值来描述图像的空间语义特征。

采用  $T_1, T_2, T_3$  作为描述图像空间语义的语言变量值,  $M(T_1), M(T_2), M(T_3)$  是相应的特征空间语义规则,  $g(t)$  表示图像  $I$  对应语言值  $T_i$  的语义隶属度 (色调等三维语言向量)。

对于给定语言向量, 其中的每个分量对应的语言值构成的模糊集合形式为  $A = \{g(t), I\}$ ,  $I$  表示数据库中的图像,  $0 \leq g(t) \leq 1$ , 是图像  $I$  中对应的语言值的隶属度。这样图像空间语义检索过程就变成语义隶属度值在模糊集上的具体运算过程。

由于采用模糊集进行检索时, 需要一定的逻辑表达知识, 查询不方便, 这样可把模糊语义值通过隶属度函数转换成  $[0,$

1] 之间的确定值。针对空间特征的每个语言值给出逻辑操作为: 设定逻辑词组函数  $\eta_i(a, g, I, \Psi)$ , 若  $g_a(t)$  大于阈值, 则该函数取为真; 否则为假。  $\Psi$  为取值范围  $[0, 1]$  的分割点。

逻辑词组函数之间的逻辑操作同一般的数学方法, 即: 与 ( $\wedge$ )、或 ( $\vee$ )、非 ( $\sim$ )。这样相应的高层空间语义检索就可由词组的逻辑运算表达式方便求解。例如, 对于查询语句“红色分散的并且红色纯或者红色亮的空间”, 转换成语言变量的表达式就是:  $\eta = \eta_1(\text{“红色分散”}, g, I, \Psi) \wedge \eta_2(\text{“红色纯”}, g, I, \Psi) \vee \eta_3(\text{“红色亮”}, g, I, \Psi)$ 。

## 5 实验与分析

本文所进行的基于空间信息熵的语义图像检索实验主要包括如下方面: a) 基于语义的图像检索系统的效率; b) 检索效率与语义规则的关系。

一般情况下, 按照相关语义描述和抽取方法, 需要事先把图像分类到事先指定的特征语义类别中, 接着把语义类别的描述转换为相应的图像颜色空间语义特征, 这种语义描述功能不仅有限, 而且获得的空间语义是确定的表述, 缺乏模糊描述的灵活性。本文根据 HSV 颜色空间模型, 采用三个空间特征, 其中  $T_1, T_2, T_3$  分别表示图像色调、饱和度和亮度。对这三个不同变量词集中的各个语言值, 通过隶属度函数使每个图像都应该与之对应产生映射, 为了更好协调图像检索效果和数据库的储存功能的关系, 因而本文采用的方法是: 把隶属度函数产生的基本语言值保存到数据库中, 扩展词的语言隶属度值直接由特定的转换算法来实现, 再结合数据库中基本语言值构成逻辑表达式进行图像检索。

### 5.1 图像语义检索效果

在评价图像检索算法效率中, 本文采用精确度 (precision) 和检索率 (recall) 两个指标。

精确度为查询结果中和例子图像相关的目标图像数与检索结果返回的图像总数之比:

$$\text{precision} = R/N \quad (7)$$

检索率定义为检索结果队列中检索到的目标图像数与全部目标图像数之比:

$$\text{recall} = R/M \quad (8)$$

其中:  $R$  是查询结果中与例子图像相关的目标图像数,  $N$  是检索结果返回的图像总数,  $M$  是图像库中与例子图像相关的目标图像总数。对于不同图像检索算法, 在相同的检索率条件下, 检索的精确度越高表明该算法的检索效果越好。

对于文中采用模糊集和语言变量所构成的图像检索系统, 在实验中自建图像库共包含 2 200 幅包括建筑物、鲜花、汽车、动物、自然景物等不同类型的彩色图像, 在对彩色图像处理时, 色调有红色、橙色、黄色、绿色、蓝色、紫色, 饱和度以及亮度为可选值域。选取 RGB 颜色空间, 并将图像均匀量化为 HSV 空间。使用 3.3 节中的语义规则来获取语义信息, 构成图像空间特征语义, 由第 4 章中方法获得语义特征值或者语义特征值的逻辑组合并进行特征检索, 由上面查准率的定义对检索效果则进行计算, 为客观评价选取平均值当做图像查准率的指标。图 1 给出了检索请求“红色分散的并且红色纯或者红色亮的空间”的检索结果。

从图 1 可知, 空间语义特征能够表达图像颜色空间的分布情况, 同时由隶属度函数构成的模糊值在特征语义表达上具有

简便、灵活特性,这对提高图像检索非常有效。另外,在进行第4章中方法给出的语言描述构成的逻辑表达式时,结果显示精确度和检索率并不是同步增长的,与真正使用的自然语言还有一定的距离,这是以后研究和实验需要改进和完善的地方。

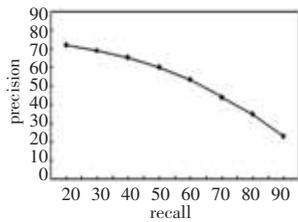


图1 检索结果中精确度与检索率的关系

## 5.2 检索效率与特征语义规则关系

本文提出语义规则对图像检索效率有着直接的影响,因此,如何构造从颜色空间特征向量到语义隶属度函数的映射,是语言变量描述空间特征语义的关键。实验中获得的语义规则的适应度与图像检索查准率的关系如图2所示。

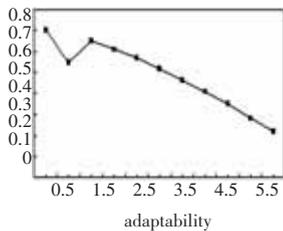


图2 查准率与适应度的关系

图2显示在适应度超过0.5时,图像检索的查准率与适应度不是同方向变化的,即查准率的增大是以适应度的减少为代价,这是因为适应度与构成颜色空间信息熵的隶属度函数密切相关;当适应度处于0.5~1.5附近时,图像检索中查准率与适应度的关系由分段函数来给定,无论适应度的增减,查准率的变化不是十分明显。这是因为图像空间特征向量对图像空间信息分布的一种近似表达,语义规则给出的语义特征是一种模糊表达,不具有文本信息查询的精确性。为了更好地符合人类的视觉感知行为,可采用广义信息熵理论来完善和解决这些问题。

## 6 结束语

本文采用模糊信息熵来描述图像颜色的空间分布特征,并结合人类的视觉特征和改进图像空间特征的提取方法,从图像的语义特征描述、语义特征提取以及模糊语义检索三方面系统地阐述了图像语义检索的方法,通过提交空间语义描述来检索空间图像。实验结果表明,用语言变量描述图像语义不仅可表达丰富的语义,而且能提供灵活的模糊匹配策略,与人类视觉具有较好的一致性,还可以有效地减少图像颜色空间分布特征与图像高层语义之间的差距。

为了进一步提高和改进空间特征语义图像检索算法的性能,以下是需要进一步研究的方向:结合图像的纹理、形状等其他低层视觉特征,研究完整的颜色空间语义检索算法;考虑更加符合人类视觉特征的权函数,结合研究相关反馈的语义检索方法,来改进算法获得更好的检索效果。

### 参考文献:

[1] 翟裕忠,胡伟,郑东栋,等. 关系数据库模式和本体间映射的研究综述[J]. 计算机研究与发展, 2008, 45(2): 300-309.

[2] RUI Y, HUANG T S, CHANG S F. Image retrieval, current techniques, promising directions, and open issues[J]. *Journal of Visual Communication and Image Representation*, 1999, 10(1): 39-62.

[3] COLOMBO C, BIMBO A D, PALA P. Semantics in visual information retrieval[J]. *IEEE Multimedia*, 1999, 6(3): 38-53.

[4] BUXTON H, MEKERJEE A. Conceptualizing images[J]. *Image and Vision Computing*, 2000, 18(2): 79.

[5] HERMES T, KLAUCK C, KREYSS J, et al. Image retrieval for information systems in storage and retrieval for image and video data bases III[C]//Proc of SPIE 2420. San Jose: [s. n.], 1995.

[6] MOJSILOVIC A, ROGOWITZ B. Capturing image semantics with low-level descriptors[C]//Proc of International Conference on Image Processing, ICIIP2001. [S. l.]: IEEE Press, 2001: 18-21.

[7] SHEIKHOESLAMI G, CHANG W, ZHANG A. Semantic clustering and querying on heterogeneous features for visual data[C]//Proc of ACM Multimedia. Bristol: ACM Press, 1998: 3-12.

[8] FUNG C Y, LOE F K. Learning primitive and scene semantics of images for classification and retrieval[C]//Proc of ACM Multimedia, Vol 2. Orlando: ACM Press, 1999: 9-12.

[9] 王崇骏, 杨育彬, 陈世福. 基于高层语义的图像检索算法[J]. 软件学报, 2004, 15(10): 1461-1469.

[10] PASS G, ZABIN R, MILLER J. Comparing images using color coherence vectors[C]//Proc of CM International Conference on Multimedia. Boston: [s. n.], 1996: 65-73.

[11] HUANG Jing. Image indexing using color correlograms[C]//Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Juan: [s. n.], 1997: 762-768.

[12] 王涛, 胡事民, 孙家广. 基于颜色-空间特征的图像检索[J]. 软件学报, 2002, 13(10): 2031-2036.

[13] FAUQUEUR J, BOUJEMAA N. Region-based image retrieval: fast coarse segmentation and fine color description[J]. *Journal of Visual Languages and Computing (JVLC)*, Special issue on Visual Information Systems, 2004, 15(1): 69-95.

[14] ZHANG Y, MEHROTRA S, HUANG T S. A multimedia information retrieval model based on semantic and visual content[C]//Proc of the 5th International ICYCS Conference. Nanjing: [s. n.], 1999: 468-475.

[15] 王海晖, 彭嘉雄, 吴巍. 采用交互信息量评价遥感图像融合结果的方法[J]. 华中科技大学学报: 自然科学版, 2003, 31(12): 32-34.

[16] 吴薇. 基于最大模糊熵原理的多阈值图像分割新算法[J]. 系统工程与电子技术, 2005, 27(2): 357-360.

[17] WANG J, DU Y, CHANG C I, et al. Relative entropy-based methods for image thresholding[C]//Proc of IEEE International Symposium on Circuits and Systems (ISCAS). 2002: 265-268.

[18] JOHN Z. An information theoretic approach to content based image retrieval[D]. Louisiana: Louisiana State University and Agricultural and Mechanical College, 2000: 45-62.

[19] 孙君顶, 崔江涛, 刘卫光, 等. 基于熵的图像空间特征提取及检索方法[J]. 系统工程与电子技术, 2006, 28(6): 791-794.

[20] 李清勇, 胡宏, 施智平, 等. 基于纹理语义特征的图像检索研究[J]. 计算机学报, 2006, 29(1): 116-123.

[21] TAO W B, TIAN J W, LIU J. Image segmentation by three-level thresholding based on maximum fuzzy entropy and genetic algorithm[J]. *Pattern Recognition Letters*, 2003, 24: 3069-3078.