

基于区域语义子概念分布的场景分类

王灵江 阮佳彬 杨育彬

(南京大学计算机软件新技术国家重点实验室 南京 210093)

摘要 大量的需求使得图像检索成为当前研究的热点。提出了对图像进行均匀分割、识别区域子概念并在此基础上对概念分布向量进行融合的场景分类方法。实验结果表明,这种方法提高了分类结果的 P/R 值,获得了很好的分类结果。

关键词 区域子概念,概念分布向量,特征融合,场景分类

Scene Classification Method Based on the Feature Fusion of Regional Semantic Sub-concept Occurrence

WANG Ling-jiang RUAN Jia-bin YANG Yu-bin

(State Key Laboratory for Novel Software and Technology, Nanjing Univ., Nanjing 210093, China)

Abstract A large quantity of practical demand has made image retrieval a research focus. We proposed a scene classification method based on the fusion of concept occurrence vectors, which are gained by segmenting images uniformly and recognizing corresponding regional sub concepts. Experiments show that our method increases the P/R values of scene classification, bringing a better result.

Keywords Regional sub concepts, Concept occurrence vector, Feature fusion, Scene classification

1 引言

用户对图像的检索有多种模式, Datta 等人^[1]将用户的检索模式分为基于关键词、自由文本、示例图像、草图以及综合方式这 5 种。基于关键词的检索是一般检索系统(例如谷歌、百度和雅虎等)最常用的方式,这种方式只涉及少数几个词的匹配问题,相比而言,自由文本的方式需要采用自然语言理解的技术。示例图像的方式是 CBIR 技术最常采用的方式,这种方式本质上是用示例图像的视觉特征和数据库中其它图像的视觉特征进行比较分析,从而得到检索结果。草图的方式需要用户对目标图像提供大致的描述,对用户要求比较高,在 QBIC^[2]等系统中得到了应用。每一种方式都各有优缺点,所以综合各种方式来权衡利弊是当前的一种趋势,而具体采用何种方式与用户的检索行为习惯有关。

Smeulders 等人^[3]将图像域分为狭域和广域两种,在狭域图像的检索行为中,用户的检索目标是很明确的,通常其视觉特征都能得到准确表达, CBIR 系统能够满足这种检索需求;而在广域图像的检索行为中,用户通常从自己的主观概念出发,对目标所具有的视觉特征没有明确的规定,检索的结果也很难明确说明其好坏,这种行为需要系统具有较强的语义处理能力,一般的 CBIR 系统都表现得不好。Smeulders 认为“语义鸿沟”就存在于这两者之间。另外, Datta 等人^[1]则直接将检索行为分为联想检索(search by association)、明确的检索(aimed search)和类别检索(category search)。在联想检索

的方式中,用户的初始目标不明确,但随着检索和对结果浏览的迭代过程不断清晰;而明确的检索,顾名思义,与联想检索截然相反,是对一幅或者某几幅具体图像的查找;类别检索则接近于前两者的折中,是对一个语义类别图像的查找,例如山峰、高楼等。从其它与用户行为相关的研究^[4,5]和检索方式的特点,可看到一般情况下用户更倾向使用简单的基于关键词检索的方式,因此广域图像的检索行为占了绝大部分。而在这种简单的检索模式下,基于语义的系统将使检索结果更能准确地满足用户的需求,基于语义的检索更有市场。因此,使图像的检索更能处理语义的问题成了当前的热点,而获取图像的语义对实现这个目标是不可或缺的。

图像场景分类是一个获取场景语义的方法,即对图像的个体赋予场景语义的过程。一方面,图像高层语义包括了场景语义,因此场景分类可以为进一步理解图像提供基础(例如提取事件语义和情感语义),从而更好地检索图像;另一方面,场景分类也要以图像的其它特征为基础。

场景分类的应用有很多, Boutell 等人^[6]将其概括为图像检索、数码冲印和自动的图像对准等,涉及众多技术,获取场景的语义对图像检索这个应用而言是十分重要的。

场景语义主要是将对象和对象关系语义联系起来,经过思考而得出的一种主观认识。场景分类方法主要有两种模式:一是基于图像的全局特征,一是基于局部特征。前者提取了图像的全局特征,然后利用机器学习的方法来对图像进行分类;后者首先对图像进行分割,提取区域特征,然后利用区

到稿日期:2010-11-21 返修日期:2011-03-12 本文受国家自然科学基金项目(60875011,60723003,60721002),江苏省自然科学基金重点项目(BK2010054)资助。

王灵江(1985-),男,硕士生,主要研究方向为海量数据挖掘和机器学习, E-mail: yangyubin@nju.edu.cn; 阮佳彬(1983-),男,硕士,主要研究方向为海量数据挖掘和机器学习; 杨育彬(1977-),男,博士,副教授,主要研究方向为海量数据挖掘和机器学习。

域对象的特征以及对象的空间位置关系等来对图像进行分类。相关的研究有很多,Payne 和 Singh^[7]对图像的边缘信息进行分析,提了一种对户内、户外场景的分类技术;Oliva 和 Torralba^[8]提出了 6 个图像的描述子(naturalness, openness, roughness, expansion, ruggedness)来描述图像的整体特征,对不同类别的图像进行建模,他们的实验取得了良好的效果。Rasiwasia 和 Vasconcelos^[9]基于概率模型提出了一种半监督的场景分类方法。Vogel 等人^[10,11]对图像进行均匀分割,并据此提出了概念分布向量,对图像场景进行建模,也取得了良好的分类效果。在很多情况下,不同场景的界限并不是十分明显,场景语义相对第二层次的图像语义有更强的模糊性,因此有很多文献针对场景的模糊性提出了相应的方法。如 Boutell 等人^[12,13]提出了多标记的场景分类技术,其对于一个图像的场景区标不只一种,而是根据图像的内容而定;Zhou 和 Zhang^[14]将其提出的多示例多标记学习构架应用到场景分类。

本文为了改变 LabelMe^[15]非均匀的多标注图像底层视觉特征处理复杂的状况,对其进行了均匀化处理,从而使其底层视觉特征的处理相对简单,进而对区域子概念做识别;并对不同的视觉特征及其融合的分类效果做了对比。本文在识别区域子概念的基础上,提取概念分布向量^[10,11](Concept Occurrence Vector,简称 COV)作为特征向量,进而通过对训练集的图像类别建模,对测试集进行分类。实验证明,本文所述的方法取得了良好的场景分类效果。

本文第 2 节对非均匀的图像标注进行均匀化处理,进而对区域子概念做识别;第 3 节在区域子概念识别的基础上,对概念分布向量进行融合以用于场景分类;第 4 节描述实验过程,以及实验结果的分析;最后是对全文的简单总结。

2 区域子概念的识别

本文采用 LabelMe 的非均匀化图像数据,对其进行均匀分割,并采用机器学习的方法对区域子概念进行识别。步骤如下:

1. 图像的多边形标注;
2. 多边形标注层次的确定;
3. 图像的均匀分割;
4. 区域子概念的识别。

2.1 图像的多边形标注

LabelMe 的图像标注用多边形表示,多边形区域的子概念为相应的标注,然而多边形标注并不一定能将整个图像覆盖,所以本文将这些未覆盖区域统一用“unknown”表示。

2.2 多边形标注层次的确定

LabelMe 的多边形标注是可以重叠交叉的,例如,图 1 中的“rock”所标示的区域,既处于“rock”多边形当中,也处于“sea”多边形当中,因此无法直接利用多边形标注来判断该区域的类别。这是区域概念获取的一大难点,因为二维图像的区域特征应是处于最上层的物体的视觉特征,而多边形标注的层次的确定是需要通过人为语义判断才能准确得知的。根据文献^[15]给出的判断规则,总结判断过程如下:

for 重叠区域的物体 A 和 B

if A(B)总是处于最底层的(例如:没有边界的天空,不会覆盖其他物体的道路等)

then B(A)就处于上层

else if A(B)完全包含在 B(A)里面

then A(B)就处于上层

else if 多边形 A(B)在重叠区域拥有更多的顶点

then A(B)就处于上层

else 根据图像特征来判断哪个物体处于上层(例如:分别计算物体 A 和 B 的颜色直方图,再与重叠颜色直方图作比较,比较接近的那个就是处于上层的)

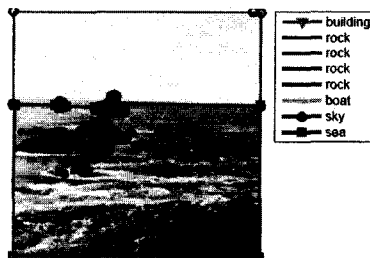


图 1 LabelMe 图像的有重叠的多边形标注(图像属于“coast”类别,包含的物体有:sky, sea, rock, boat, building)

本文按此规则将图像分割成无重叠的区域,例如,对于图 1 做无重叠分割处理,其结果如图 2 所示。



图 2 无重叠的多边形标注图像。

2.3 图像的均匀分割

在完成多边形标注的基础上,本文对图像进行均匀分割,一个均匀区域的概念为其中面积比例最大的无重叠区域的概念,例如对图 1 的区域做均匀化后,各区域的概念如图 3 所示,



图 3 均匀化的多边形标注图像

其中的“building”和“boat”区域因为过小而消失了,但是这对图像场景的语义的影响不大。采用均匀分割的理由如下:

1. 图像分割仍是当前研究的一个难点,常见的方法效果都不是很好。
2. 均匀分割比较简单,用方形的图像分块做特征提取等

处理比较方便。

3. 分割后的区域比较小,其中所含的物体比较单一,语义比较明确。

4. 相关的文献^[10,11]的实验结果表明,基于均匀分割的方法已能达到很好的效果。

本文对 LabelMe 的标注噪音进行过滤后,使其减少到 246 个概念,并且其中大部分都是平均分布值很小的概念,但这仍对概念的预测造成了很大的困扰。文献^[10,11]只选用了 9 个代表性的概念来做分类,却取得了良好的效果。因此,这里我们选择了平均分布值最大的前 11 个子概念,如表 1 所列。

表 1 区域均匀化分割后的主要子概念

概念名称	分布均值
sea	0.049894
sky	0.246013
building, edifice	0.198488
sand	0.010075
mountain, mount	0.111331
tree	0.108263
plant, flora, plant life	0.012944
rock, stone	0.009256
river	0.015088
road, route	0.087275
field	0.037725

2.4 区域子概念的识别

虽然大部分的均匀化区域包含的物体比较单一,但由于它不是对物体的准确分割,因此部分区域就可能含有多种物体。而前述的均匀化方法是将占最大比例的物体的类别作为该区域的概念,所以区域就可能不纯。如果不纯区域来做训练数据,就可能使测试结果不准确,故本章采用纯度为 100% 的区域(即只含有单一物体的区域)来做训练数据。

本文采用的视觉特征有 HSI 颜色直方图(其中色调量化为 36 份,饱和度量化为 32 份,而亮度量化为 16 份)、边缘方向直方图^[16-18](Edge Orientation Histogram,简称 EOH,它包括 4 个方向(0°,45°,90°,135°)的边缘和一个无方向边缘)以及基于灰度共生矩阵^[19,20](Grey Level Co-occurrence Matrix,简称 GLCM)的相关纹理描述子(包括对比度(Contrast)、能量(Energy 或者 Uniformity)、熵(Entropy)、同质性(Homogeneity 或者 Inverse difference)、逆差分矩(Inverse difference moment))。

子概念的识别采用机器学习的方法,其主要思想为:记图像区域的特征集为 X ,各区域所对应概念为 $g(X)$,其中 $x \in X, g(x) \in C, C$ 是所有概念的集合,利用 $\{X, g(X)\}$ 学习 g 的一个近似 g' ,然后根据 g' 来对未知的数据做识别。

3 COV 融合

3.1 COV 特征向量

在获取了图像区域的概念后,本文提取文献^[10,11]中所述的概念分布向量(Concept Occurrence Vector,简称 COV)作为特征向量,其主要思想如下:设图像库中所有的子概念为 $C = \{c_1, c_2, \dots, c_n\}$,则图像 $x_i \in X$ 的 COV 特征是一个 n_c 维的向量,其第 k 维的值为:

$$COV(x_i, k) = \sum_{l=1}^{n_x * n_y} \delta(g'(x_i, l), c_k) \quad (1)$$

式中, x_i 采用的是 $n_x * n_y$ 的格子做均匀分割的, $x_{i,l} \in x_i, l =$

$1, 2, \dots, n_x * n_y, g'$ 是根据 4.2 节方法所建立的子概念模型, δ 函数定义如下:

$$\delta(a, b) = \begin{cases} 1, & a \neq b \\ 0, & a = b \end{cases} \quad (2)$$

3.2 COV 特征向量的融合

通常特征融合的方法有两类^[21,22]:早期特征融合(Early Feature Fusion)和晚期特征融合(Late Feature Fusion),这两种方法可用图 4 和图 5 进行概括,前者主要是在特征语义层上做处理,综合多种视觉特征产生效果更好的新特征,以供分类;后者主要在语义层上融合,将不同的分类器结果进行有机组合,得到综合的结果。



图 4 早期特征融合

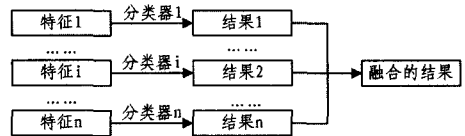


图 5 晚期特征融合

本文所采用的特征融合的方法近似于早期特征融合,区别在于这里融合的对象不是最早的视觉特征,而是不同特征 f_k 的概念分布向量,即 COV_k ,这是一种中间的特征。这里的 COV_k 都是根据式(1)计算得出的,是规范的结果,因此本文直接将各个 COV_k 并接,即 $COV = [COV_1, COV_2, \dots, COV_{nF}]$ 。

4 实验

本文的实验主要分为两个步骤:(1)通过均匀分割得到图像分块,进而用 SVM 算法识别区域子概念;(2)在此基础上,采用两种特征融合方法(对视觉特征进行融合和对 COV 特征进行融合)进行场景分类。

4.1 实验流程

在对区域子概念建立模型 g' 并获得图像各个区域的子概念后,本文提取图像 COV 特征,也记为 X ,相应图像类别为 $f(X)$,利用训练集的 $\{X, f(X)\}$ 学习得到 f 的一个近似 f' ,然后利用 f' 来预测未知图像类别。实验流程可用图 6 描述,主要包括以下步骤:

Step 1 对所有图像用 $10 * 10$ 的格子进行均匀分割得到图像分块。

Step 2 提取每个分块的视觉特征(即上文所述的 HSI 颜色直方图、EOH 及 GLCM 纹理特征)。对于训练集还需要对原有标注进行均匀化,得到相应的新标注。

Step 3 对训练集中的视觉特征及相应的分块标注进行学习,得到相应的概念模型。

Step 4 利用所得的子概念模型 g' ,识别所有图像分块的概念,并提取各图像分块的 COV 特征。

Step 5 根据训练集的 COV 特征及对应的图像类别,学习得到图像类别模型。

Step 6 根据图像类别模型 f' 和测试集的 COV 特征,对测试集进行类别预测。

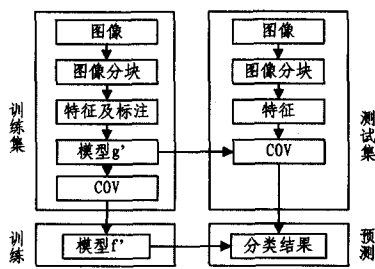


图6 基于子概念识别的COV模型流程

4.2 区域子概念的识别结果及分析

实验首先要对子概念进行识别,为了得到更好的识别效果,这里对视觉特征做了融合,即将规范化的各种特征进行拼接,这是常用的早期特征融合的方法。图像的底层视觉特征数据有以下4类:

1. HSI颜色直方图特征,记为 F_{HSI} 。
2. EOH特征,记为 F_{EOH} 。
3. GLCM纹理特征,记为 F_{GLCM} 。
4. 融合以上各组特征,记为 F_{COM} 。

本文随机选取了 $200 \times 11 = 2200$ 个概念纯度为 100% 的区域(每个主要子概念选取 200 个)作为训练数据,用 10 倍交叉验证的方法,利用 SVM 算法来实验,所得的结果如表 2 所列。

表2 区域子概念识别的P/R值

数据	F_{HSI}	F_{EOH}	F_{GLCM}	F_{COM}
sea	38.3/36.0	25.7/33.5	41.7/50.0	45.7/45.0
sky	60.7/65.0	27.9/66.5	43.8/67.0	66.7/73.0
building	30.5/43.0	50.3/43.5	59.6/63.5	61.0/54.0
sand	50.0/43.5	8.6/5.0	22.7/13.5	50.3/47.5
mountain	18.7/13.0	20.0/15.5	21.8/18.5	30.7/27.5
tree	29.9/36.5	22.1/25.0	22.3/22.0	37.1/36.5
plant	44.2/55.0	24.3/29.5	29.7/38.5	42.3/45.5
rock	35.3/41.0	16.5/11.0	20.3/14.0	36.8/41.0
river	23.3/13.5	11.5/3.5	22.3/11.5	25.6/20.5
roand	48.9/44.5	21.1/25.5	25.8/28.0	49.2/58.0
field	43.9/39.5	14.0/8.0	23.4/30.0	38.9/40.5
AVG	38.5/39.1	22.0/24.2	30.3/32.2	44.0/44.5

从中可见子概念识别的 P/R 值都比较低,原因可能有以下几种:

1. 图像的分辨率不高(都是 256×256),区域特征不是很稳定。
2. 图像的内容比较丰富,而 LabelMe 的某些标注不是十分严格,可能有不少概念的视觉特征波动比较大。

表中的数据还表明,融合后的视觉特征 F_{COM} 的识别效果要比没有融合的好,这说明早期特征融合是有好处的,但在本文的实验中其对识别效果的提高还不够,原因除了数据本身的模糊性比较大之外,还可能是因为所用的特征不足以全面地表达图像的内容。

4.3 场景分类结果及分析

本文主要对比了两种特征融合的方法对场景进行分类:对视觉特征直接进行融合和对 COV 特征进行融合。首先是对视觉特征直接进行融合,对于 4.2 节所述的融合的视觉特征 F_{COM} ,本节的实验采用 SVM 算法对子概念进行识别,然后按式(1)提取 COV 特征向量,再分别利用 SVM 算法进行场景分类;其次是对 COV 特征进行融合,COV 特征的融合方法类似 3.2 节所述的 COV 特征融合,它是将底层特征学习得到

的 COV 特征进行拼接组合。这两种方法得到的效果差别比较大,这说明选择合理的特征融合方法是很重要的。为了对结果作对比分析,将图像的 COV 特征数据分为如下 5 类:

1. HSI 颜色直方图学习得到的 COV 特征,记为 D_{HSI} 。
2. EOH 学习得到的 COV 特征,记为 D_{EOH} 。
3. 利用 GLCM 纹理学习得到的 COV 特征,记为 D_{GLCM} 。
4. 融合以上各 COV 特征,记为 D_{COM1} 。
5. 融合各底层视觉特征,学习得到的 COV 特征,记为 D_{COM2} 。

对于以上 5 类数据,本文分别采用 SVM 进行分类。这里所用的 LabelMe 图像库“Spatial Envelope”有 8 个类别,每个类别随机选取 200 张作为训练集,其余用来做测试,得到的实验结果如表 3 所列。

表3 基于子概念识别的COV分类的P/R值

数据	D_{HSI}	D_{EOH}	D_{GLCM}	D_{COM1}	D_{COM2}
coast	39.9/36.9	72.7/60.0	56.0/49.4	76.0/73.1	71.4/68.8
forest	54.7/73.4	77.8/87.5	63.3/83.6	82.9/94.5	72.1/93.0
highway	41.3/51.7	32.1/70.0	35.7/68.3	49.0/80.0	43.2/68.3
insidecity	32.1/39.8	62.0/62.0	56.0/56.5	79.0/76.9	61.5/62.0
mountain	40.2/28.2	61.8/63.2	29.4/30.5	76.3/75.9	57.6/63.2
opencountry	64.3/51.4	65.9/58.1	45.7/40.0	73.9/66.2	69.3/59.0
street	38.1/48.9	46.5/51.5	43.9/39.1	79.5/71.7	62.9/60.9
tallbuilding	39.7/38.5	71.6/50.0	54.6/37.8	81.6/73.7	81.1/55.1
AVG	43.8/46.1	61.3/62.7	48.1/50.6	74.8/76.5	64.9/66.3

将表 3 中各类数据产生分类的平均 P/R 值做对比,结果如图 7 所示。

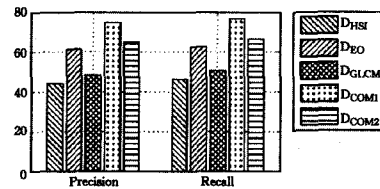


图7 表3中平均P/R值的对比

从图 7 可以发现,只要是做过特征整合的数据,其分类效果总会比用单一特征的分类效果好,这是因为特征整合综合了多种信息,更能描述图像的内容。而 COV 融合的方法比直接对视觉特征整合的方法产生的分类效果要好很多,这说明本文提出的 COV 融合的方法是比较有效的。融合的 COV 特征虽然使得分类结果更加准确,但所得的效果 ($P/R = 74.8\% / 76.5\%$) 并不算特别高,这可能是由于以下两方面的原因造成的:

1. 子概念识别的准确率比较低。
2. 所选用的主要概念数不是很多。

第 1 点可以从表 2 的结果看出,如果子概念识别准确率比较高,则分类结果的准确率可能会高得多。第 2 点与算法的特性有关,本文融合后的 COV 特征只有 33 维,数据量较少。而经过后的 LabelMe 数据库中原有标注的数量有 246 个,本文在子概念识别时,为使数据简单,只选取其中的 11 个,信息量也因此有所减少,分类准确率有所降低。

上述的第 5 类数据是按照 Vogel 等人^[10,11]的方法产生的,Vogel 等人只对底层视觉特征进行融合,他们所用的数据有 6 个场景类别,一共 700 张图像,所得的最好的平均准确率为 69.9%(采用 SVM 分类),而在本文的 8 类场景分类结果

(下转第 200 页)

class of convex objective functions [J]. Control and Cybernetics, 1997, 26(3): 375-390

- [9] 刘峰,刘贵忠,张苗生. 进化规划的 Markov 过程分析及其收敛性[J]. 电子学报, 1998, 26(8): 76-79
- [10] 高永超,李歧强. 退火进化规划算法及其收敛性[J]. 系统仿真学报, 2006, 18(3): 577-580
- [11] 杜海峰,公茂果,刘若辰,等. 自适应混沌克隆进化规划算法[J]. 中国科学 E 辑: 信息科学, 2005, 35(8): 817-829

(上接第 185 页)

中的平均 P/R 值为 64.9%/66.3%, COV 融合的方法使分类准确率得到了较大提高, 这说明了 COV 融合对场景分类是十分有效的。

结束语 本文首次将 LabelMe 非均匀的图像标注进行均匀化处理, 使其底层视觉特征的处理相对简单, 进而对区域子概念做识别。对不同的视觉特征及其融合的识别效果做了对比, 并利用所识别的子概念提取了图像的 COV 特征, 进而通过对训练集的图像类别建模, 对测试集进行分类。这个过程中, 对不同的子概念的识别方法产生的 COV 特征及融合的 COV 特征的分类效果做了横向对比分析, 同时也对 COV 特征融合及底层视觉特征融合的分类效果做了纵向的对比。实验结果表明, 对 COV 特征进行融合能取得良好的分类效果。

参 考 文 献

- [1] Datta R, Joshi D, Li J, et al. Image retrieval: Ideas, influences, and trends of the new age[J]. ACM Comput. Surv., 2008, 40(2): 1-60
- [2] Flicker M, Sawhney H, Niblack W, et al. Query by image and video content: The qbic system[J]. Computer, 1995, 28(9): 23-32
- [3] Smeulders A W M, Member S, Worring M, et al. Content-based image retrieval at the end of the early years[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22: 1349-1380
- [4] Choi Y, Rasmussen E M. Users' relevance criteria in image retrieval in American history[J]. Information Processing & Management, 2002, 38(5): 695-726. <http://www.sciencedirect.com/science/article/B6VC8-44M2C2F-3/2/bffb7decc20305f925f1f6-40cc31f27e>
- [5] Eakins J P, Briggs P, Burford B. Image retrieval interfaces: A user perspective[C]// CIVR. 2004: 628-637
- [6] Boutell M R, Brown C M, Luo J. Review of the state of the art in semantic scene classification[EB/OL]. <http://hdl.handle.net/1802/284>, 2002
- [7] Payne A, Singh S. Indoor vs. outdoor scene classification in digital photographs[J]. Pattern Recognition, 2005, 38(10): 1533-1545. <http://www.sciencedirect.com/science/article/B6V14-4FX23D4-4/2/3a59f47d1a21c4dba5021f3c02352a66>
- [8] Oliva A, Torralba A. Modeling the shape of the scene: A holistic representation of the spatial envelope[EB/OL]. Available: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=?doi=10.1.1.2,2001>

- [12] Huang H, Wu C G, Hao Z F. A Pheromone-rate-based Analysis on the Convergence Time of ACO Algorithm [J]. IEEE Transactions on Systems, Man, and Cybernetics-Part B, 2009, 39(4), 910-923
- [13] 黄翰,郝志峰,吴春国,等. 蚁群算法的收敛速度分析[J]. 计算机学报, 2007, 30(8): 1343-1353
- [14] 黄翰,郝志峰,秦勇. 进化规划算法的时间复杂性分析[J]. 计算机研究与发展, 2008, 45(11): 1850-1857

- [9] Rasiwasia N, Vasconcelos N. Scene classification with low-dimensional semantic spaces and weak supervision[C]//Proceedings of the IEEE Conference on Computer Rision and Pattern Recognition. June 2008: 1-6
- [10] Vogel J, Schiele B. Natural scene retrieval based on a semantic modeling step[C]// CIVR. Springer Verlag, 2004
- [11] Vogel J. Semantic scene modeling and retrieval[D]. Technische Wissenschaften ETH Zürich, Zürich, 2004
- [12] Boutell M R, Luo J, Shen X, et al. Multi-label semantic scene classification[R]. 2003
- [13] Boutell M R, Luo Jie-bo, Shen Xi-peng, et al. Learning multi-label scene classification[J]. Pattern Recognition, 2004, 37(9): 1757-1771. <http://www.sciencedirect.com/science/article/B6V14-4CF14JX-1/2/a17089f241a1d23f218e55d2c8d9f763>
- [14] Zhou Zhua, Zhang Ming. Multi-instance multilabel learning with application to scene classification[EB/OL]. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.67.4619>, 2007
- [15] Russell B C, Torralba A, Murphy K P, et al. Labelme: A database and webbased tool for image annotation[J]. Int. J. Comput. Vision, 2008, 77(1-3): 157-173
- [16] Park D K, Jeon Y S, Won C S. Efficient use of local edge histogram descriptor[C]// MULTIMEDIA '00: Proceedings of the 2000 ACM workshops on Multimedia. New York, NY, USA: ACM, 2000: 51-54
- [17] Wu P, Ro Y M, Won C S, et al. Texture descriptors in mpeg-7 [C]// CAIP '01: Proceedings of the 9th International Conference on Computer Analysis of Images and Patterns. London, UK: Springer-Verlag, 2001: 21-28
- [18] Won C S, Park D K, Park S-J. Efficient use of mpeg-7 edge histogram descriptor[J]. ETRI Journal, 2002, 24(1): 23-30
- [19] Haralick R M, Shanmugam K, Dinstein I. Textural features for image classification[J]. IEEE Transactions on Systems, Man and Cybernetics, 1973, 3(6): 610-621. <http://ci.nii.ac.jp/naid/10019290755/en/>
- [20] Clausi D A. An analysis of co-occurrence texture statistics as a function of grey level quantization[J]. Can. J. Remote Sensing, 2002, 28(1): 45-62
- [21] Tseng B, Lin C-Y, Naphade M, et al. Normalized classifier fusion for semantic visual concept detection[C]// Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference. vol. 2, sept. 2003
- [22] 伍卉. 基于综合特征的图像语义分类技术的研究[D]. 长沙: 湖南大学, 2009