

# D-S 证据理论在图像情感标识中的应用

高彦宇, 尹怡欣

(北京科技大学 信息工程学院, 北京 100083)

**摘要:** 图像情感标识就是为图像标注形容词性关键词, 以反映用户对该图像的情感或印象. 图像的视觉特征以及语义内容是决定用户对该图像情感理解的 2 项关键因素, 而图像内容识别具有较高的不确定性, 人类的情感理解也具有很强的主观性, 因此采用 Dempster-Shafer 证据理论能较好实现图像视觉特征及语义内容到图像情感标识的不确定性推理. 考虑到图像内容识别的不确定性, 研究中按一定比例扩大了图像语义内容对各情感因子的不确定性区间, 并构建了一个原型系统对自然风景图像进行自动标识. 实验表明 Dempster-Shafer 证据理论在处理情感标识上是很有效的, 而调整不确定性区间有助于进一步提高标识准确率.

**关键词:** D-S 证据理论; 图像情感标识; 情感因子; 分级语义

**中图分类号:** TP391.4 **文献标志码:** A **文章编号:** 1673-4785(2010)06-0534-06

## Application of the Dempster-Shafer theory to affective image annotation

GAO Yan-yu, YIN Yi-xin

(School of Information Engineering, University of Science & Technology Beijing, Beijing 100083, China)

**Abstract:** Affective image annotation involves labeling an image with adjectives, so that those labels reflect the user's emotional understanding of the image. The low-level visual features and the image semantic content are two decisive factors in the user's emotional understanding of an image, while image content recognition is highly uncertain and affective understanding is strongly subjective. In the following study, the Dempster-Shafer theory was applied to represent the visual image characteristics and to model the uncertainty reasoning from those decisive factors to affective understanding. In response to the semantic recognition error, the uncertainty range of image contents to each affective factor was enlarged and a prototype affective annotation system was built to automatically label natural scenic images. Experimental results show that the Dempster-Shafer theory is promising for ambiguous annotation, and enlarging the uncertainty range is helpful for improving annotation precision.

**Keywords:** Dempster-Shafer theory; affective image annotation; affective factor; hierarchical semantics

图像蕴藏着丰富的情感内涵——夕阳晚景图会让人感到忧伤寂寞, 而塞外雪景图则令人感到平静安逸. 挖掘图像的情感内涵, 实现图像情感自动标识不仅能够帮助杂志编辑、专栏作者或广告策划者在数以万计的图像库中为他们的作品找到合适的图片或插图, 而且能帮助工薪族找到合适的图片, 以调整情绪、减轻压力, 甚至能帮助心理障碍者缓解心灵上的痛苦. 在工作日趋繁忙、生活倍感压力的今天, 图像情感自动标识有广阔的应用前景.

## 1 研究背景

除了个人经历、教育背景、宗教信仰和生活环境之外, 图像视觉特征及语义内容对用户从情感上认识图像起着至关重要的作用. 称个人经历、教育背景为主观因子, 而宗教信仰和生活环境为客观因子. 因为主观因子涉及面广, 且与人类情感认知之间的关系复杂, 目前的研究对这些因素大都采取回避的态度, 而将重点放在建立客观因子 (尤其是图像视觉特征) 与情感理解的关系上. 例如, Kobayashi 等<sup>[1]</sup>通过实验指出颜色和方向多分辨率的对比对于人的主观感知非常重要. 毛峡等<sup>[2]</sup>通过二维波动数学模型及心理实验证实了符合  $1/f^0$  波动规律的图像给人

烦躁的感觉,符合  $1/f^2$  波动规律的图像给人单调乏味的感觉,而符合  $1/f$  波动规律的图像则给人和谐舒适的感觉. Cho 等<sup>[3]</sup> 利用离散小波变换提取图像特征,然后采用交互式遗传算法实现情感图像检索. 王上飞等<sup>[4]</sup> 从心理学的“维量”思想出发,采用语义量化技术和因子分析的方法建立情感空间,并应用径向基函数将图像从颜色和形状特征空间映射到情感空间. 王伟凝等<sup>[5]</sup> 通过定义描述色彩语义的自然语言词汇,设计了基于模糊隶属函数的图像色彩语义描述方法,并在此基础上定义一种查询语言,将人类对图像内容理解的经验与图像色彩描述联系起来,实现图像高层情感语义的查询. Bae 等<sup>[6]</sup> 对图像—情感之间的关系做了较深入的研究. 他们首先通过问卷调查及人工筛选确定了 52 种图像模式(包括 20 种颜色、12 种形状、10 种纹理、10 种图案)和与这些图像模式相对应的 55 种情感因子. 然后他们建立情感因子空间模型来定义图像视觉信息与情感之间的相对关系,并根据该模型进行情感度量. Shin 等<sup>[7]</sup> 建立一个情感预测系统,通过给定的纹理图案预测人的情感取向. 他们的实验表明多层感知器在情感预测上要优于 K 均值聚类算法和朴素贝叶斯模型,能达到 92% 的预测准确率.

上述研究在图像视觉特征与情感理解方面进行了多方位探索,使图像情感标识与检索在近 10 年获得了长足进步. 然而这些研究均未考虑图像语义内容对情感理解的影响. 针对自然风景图像,提出一种应用 Dempster-Shafer (D-S) 证据理论融合图像视觉特征及语义内容信息,推理出图像情感标识的方法. 系统框架如图 1 所示.

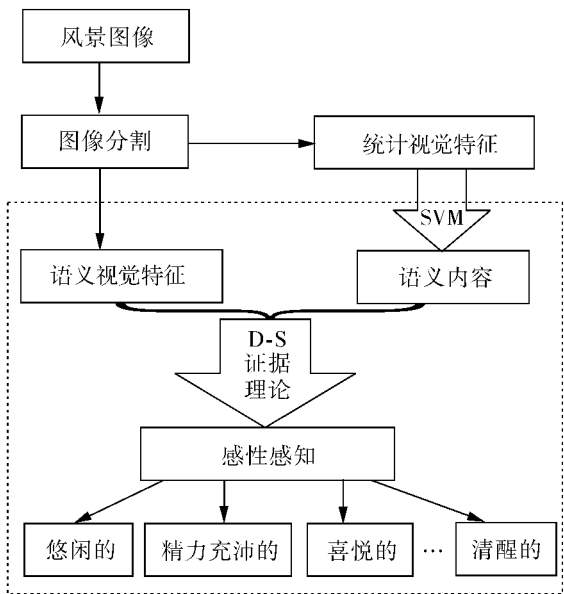


图 1 图像情感标识系统的框架图

Fig. 1 Schematic illustration of the proposed affective image annotation system

## 2 视觉特征与语义内容

准确获得图像视觉特征及语义内容是成功实现图像情感标识的关键环节. 本节重点解决图像分割、区域视觉特征提取及区域内容识别等问题.

### 2.1 图像分割

图像分割就是根据颜色分布、区域边缘以及空间布局将图像像素组织成一个个区域,并尽量使各分割区域与实际物体有密切的相关性. 本文采用基于空间限制混合模型的像素聚类法<sup>[8]</sup> 实现图像分割. 该方法由于考虑了像素位置信息,所以具有较高的分割精度. 通过实验比较,将初始迭代循环限制为 20 次,这样不仅不会对分割精度有很大影响,而且可以大大减少图像分割时间,有利于实际应用.

### 2.2 视觉特征抽取

提取每个分割区域的代表颜色作为其语义视觉特征. 此外,计算各分割区域的颜色矩和 Gabor 小波纹理特征,并应用这些统计特征向量来识别区域内容.

#### 2.2.1 代表颜色

基于对人类视觉系统的观察,发现一幅图像的颜色基调对该图像在人脑中引起的情感起着重要作用. 因此,提取各分割区域的代表颜色作为其语义视觉特征. 由于 HSV 颜色空间能更好地反应人类对色彩的认知,选择在 HSV 颜色空间建立颜色名称与颜色值的关系(如表 1 所示). 根据各种颜色的值域范围,为每一分割区域绘出颜色直方图,直方图上占比例最多的颜色即为该区域的主要颜色.

表 1 HSV 值与颜色名的一一对应关系

Table 1 The relationship between HSV and color names

颜色名	饱和度	亮度	色度
黑色	-	<20%	-
灰色	<8%	≥20% & <95%	-
白色	<8%	≥95%	-
红色	≥8%	≥20%	[0,8) ∪ [345,360]
橙色	≥8%	≥20%	[8,36)
黄色	≥8%	≥20%	[36,63)
绿色	≥8%	≥20%	[63,160)
青绿色	≥8%	≥20%	[160,218)
蓝色	≥8%	≥20%	[218,282)
紫色	≥8%	≥20%	[282,315)
洋红色	≥8%	≥20%	[315,345)

#### 2.2.2 颜色矩

计算了每个分割区域颜色分布的一阶、二阶和三阶中心矩作为该区域的统计颜色特征,并采用最完善、感知上最一致的颜色空间 CIE  $L^*a^*b^*$  作为工作空间. 这 3 种矩特征的定义分别为

$$\begin{aligned}\mu_i &= \frac{1}{N} \sum_{j=1}^N f_{ij}, \\ \sigma_i &= \sqrt{\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^2}, \\ s_i &= \sqrt[3]{\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^3}.\end{aligned}\quad (1)$$

式中:  $f_{ij}$  代表图像像素  $j$  在  $L^*a^*b^*$  的第  $i$  个颜色通道上的值,  $N$  为图像像素数. 一阶矩特征反映了图像像素分别在  $L^*a^*b^*$  3 个色彩通道上颜色值的平均值, 二阶矩和三阶矩表示图像像素在  $L^*a^*b^*$  3 个色彩通道上颜色值的方差和扭曲程度. 由于所有中心矩具有相同的单位, 所以易于实现度量和比较.

### 2.2.3 Gabor 小波纹理特征

Gabor 小波特征已被证明是一种非常有效的纹理特征. 它的基本思想是应用 Gabor 小波分解法在多尺度、多方向上抽取纹理特征<sup>[10]</sup>. 本文对每一个分割区域计算了 3 个尺度和 4 个方向上的 Gabor 小波变换系数的均值和方差, 它们组成一个 24 维的特征向量.

### 2.3 语义内容识别

本文研究的是对自然风景图像的情感标识问题. 通过对大量风景图片的观察以及参考 WordNet (<http://wordnet.princeton.edu/>), 列出 55 项常见的自然物名. 其中, 一些名称应用范围广, 适于作为景物的总称, 如石头; 另外一些名称应用范围窄, 只适用于形容特定物体, 如鹅卵石. 将这些名词分成 2 个等级: 广义关键词和狭义关键词 (如表 2 所示), 并用狭义关键词标记语义内容.

表 2 分级图像语义关键词

Table 2 Semantic keywords organized in 2 hierarchies

广义关键词	狭义关键词
水	海洋、河流、湖泊、瀑布
石头	小鹅卵石、大岩石、石碑、石墙、礁石
地面	土壤、沙滩、沙漠、雪地、草地、枯草地、泥
山	小山丘、冰山、雪山、悬崖、岛屿
天空	晴天、艳阳天、日升/日落、乌云密布的天空、云彩
植物	灌木、高树、芦苇、杂草
花	单支花、并蒂花、花簇、花床
路	柏油路、铁道、隧道、楼梯、小路
建筑物	砖制建筑、石制建筑、木制建筑、篱笆、雕塑

对每一个被分割的区域, 一方面根据其代表颜色得出其语义视觉特征, 该特征将被直接用于情感推理; 另一方面计算其统计视觉特征向量 (33 维, 包括 9 个颜色矩特征和 24 个纹理特征), 并根据该统计特征识别区域内容. 为了消除特征值范围的不一

致性, 将这些特征向量规范化, 使其具有 0 均值和单位方差.

接着从训练图像集合中手工分割出 200 个区域样本, 这些样本涉及狭义关键词指定的所有类, 而且每一类至少有 2 个训练样本. 接着计算这 200 个区域样本的统计视觉特征, 并用它们训练分类器. 最后用训练好的分类器对待识别图像区域进行分类, 以得到区域语义标识. 鉴于训练样本有限, 为了获得较高的识别率, 支持向量机 (SVM) 无疑是最合适的分类器. 它比传统的分类器 (如神经网络分类器) 有更好的泛化能力, 而且只需很少的训练样本就能获得较高的识别精度. 本文采用 RBF 函数作为 SVM 的核函数, 其主要参数为  $C$  及  $\gamma$ . 为了得到合适的核参数, 对训练图像应用三重交叉验证和网格搜索算法, 这里  $C$  及  $\gamma$  从指数增长序列  $C = 2^{-3}, 2^{-1}, \dots, 2^{11}$  及  $\gamma = 2^{-11}, 2^{-9}, \dots, 2^3$  中选择. SVM 本质上是一个二值分类器, 为了实现多类分类, 几种组合分类方案已被提出, 如一对一支持向量机 (1- $v$ -1 SVM), 一对多支持向量机 (1- $v$ - $r$  SVM), 有向无环图支持向量机 (DAGSVM), 基于不确定样本的多类支持向量机 (US\_MSVM)<sup>[11]</sup> 等. 本文采用比较成熟的 DAGSVM 实现语义分类.

## 3 基于 D-S 证据理论的不确定性推理

鉴于图像内容识别的不确定性以及情感认知的主观性, 建立图像视觉及语义描述与人类情感之间关系的问题可以看作是不确定性推理问题. 另一方面, 由于视觉特征和语义内容在人类对图像的情感认知方面都起决定性作用, 因此将这两方面联系起来推导人对图像的感知还可以看成是多传感器融合问题. D-S 证据理论是最适合解决这类多传感器融合及不确定性推理问题的方法之一.

### 3.1 D-S 证据理论

D-S 证据理论是由 Dempster 提出来的用概率上下限表示实际问题中不确定性的方法, 后来由 Shafer 进一步研究和推广, 成为系统化、理论化的不确定性推理理论. 其主要特点是通过证据的不断积累逐步缩小假设集, 从而逐步减小信息的不确定性. D-S 证据理论的具体推理过程由如下 4 步组成.

1) 定义辨别框架  $\Theta$ , 它是由  $n$  个完备的、互不相容的子命题  $a_1, a_2, \dots, a_n$  构成.  $\Theta$  中所有可能的子集构成一个超集  $2^\Theta$ , 即

$$2^\Theta = \{\emptyset, \{a_1\}, \dots, \{a_n\}, \{a_1, a_2\}, \dots, \{a_{n-1}, a_n\}, \dots, \Theta\}.$$

2) 定义一个函数  $m: 2^\Theta \rightarrow [0, 1]$ , 在满足

$$m(\emptyset) = 0, \sum_{A \subseteq \Theta} m(A) = 1 \quad (2)$$

的条件下,  $m(A)$  称为  $A$  的基本概率分配函数或基本

可信度,表示对命题 $A$ 的直接支持. 每一个 $H$ 的子集 $A$ ,如果有 $m(A) > 0$ ,则 $A$ 称为基本概率分配函数 $m$ 的一个焦元(focal element).

3)对给定的基本概率分配函数 $m$ ,计算置信函数 $\text{Bel}$ 和似真函数 $\text{Pl}$ :

$$\text{Bel}: 2^\theta \rightarrow [0, 1],$$

$$\text{Bel}(A) = \sum_{B \subset A} m(B) \quad (\forall A \subset 2^\theta),$$

$$\text{Pl}: 2^\theta \rightarrow [0, 1],$$

$$\text{Pl}(A) = \sum_{B \cap A \neq \emptyset} m(B) = 1 - \text{Bel}(\bar{A}) \quad (\forall A \subset 2^\theta).$$

式中: $\text{Bel}(A)$ 表示对命题 $A$ 为真的信任程度, $\text{Bel}(\bar{A})$ 表示对命题 $A$ 为假的信任程度, $\text{Pl}(A)$ 表示对 $A$ 为非假的信任程度,因此总有 $\text{Bel}(A) \leq \text{Pl}(A)$ 成立.由下限函数 $\text{Bel}(A)$ 和上限函数 $\text{Pl}(A)$ 构成的区间 $[\text{Bel}(A), \text{Pl}(A)]$ 表示命题 $A$ 的不确定性区间(如图2所示).

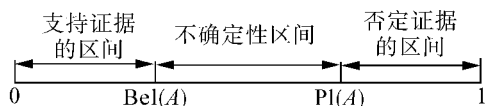


图2 图示置信函数与似真函数

Fig. 2 Illustration of belief function and plausibility measures

4)当获得了同一辨别框架下的一组基本概率分配函数 $m_1, m_2, \dots, m_t$ 后,通过Dempster合成法则,可计算出综合置信度:

$$m(D) = \frac{1}{1 - k} \left( \sum_{B_1 \cap B_2 \cap \dots \cap B_t = D} m_1(B_1) * m_2(B_2) * \dots * m_t(B_t) \right), \quad (1)$$

式中: $D = B_1 \cap B_2 \cap \dots \cap B_t$ ,  $B_i$ 代表 $m_i$ 的核心元素,且 $1 \leq i \leq t$ , $*$ 代表普通的乘法, $k$ 值称为冲突度,其值域为闭区间 $[0, 1]$ ,且定义为

$$k = 1 - \sum_{B_1 \cap B_2 \cap \dots \cap B_t \neq \emptyset} m_1(B_1) * m_2(B_2) * \dots * m_t(B_t).$$

### 3.2 基于D-S证据理论的情感标识

本文的目标是对自然风景图像进行情感标识,标识语为描述主观情感的形容词.以利用图像进行心理治疗为背景,选择4个描述积极情绪的形容词(悠闲的、平静的、清醒的、精力充沛的)作为情感标识语,称这些形容词为情感因子.下面以表3中的雪景图像及情感因子“悠闲的”为例,说明采用D-S证据理论实现图像内容信息与视觉特征融合的过程.

首先将图像分割成5个区域,提取各区域的视觉特征并采用DAGSVM法分别对各区域进行内容分类标识.这5个区域对应4个狭义语义类:晴天、云彩、雪地、小山丘和4种颜色特征:蓝色、白色、白色、黑色.对于情感因子“悠闲的”,D-S辨别框架 $\Theta$

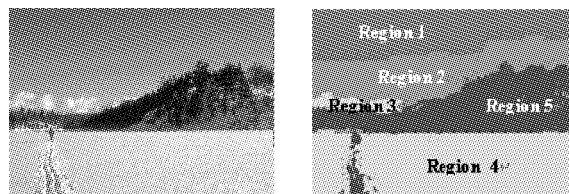
由{悠闲的,非悠闲的}这2个值组成,辨别框架中所有可能的子集构成一个超集 $2^\theta = \{\emptyset, \{\text{悠闲的}\}, \{\text{非悠闲的}\}, \{\text{悠闲的}, \text{非悠闲的}\}\}$ .根据各种视觉特征及语义内容对“悠闲的”的基本可信度及Dempster合成法则,可以计算得到各区域对“悠闲的”的综合置信度.由于只考虑颜色和语义因素,Dempster合成法则(1)可简化为

$$m(A) = m_1 \oplus m_2(A) = \frac{\sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1 - \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j)}.$$

式中: $A_i, B_j \subset 2^\theta$ .假设图像各区域彼此独立,最终的综合置信度可通过下式计算.

$$P = \sum_{i=1}^N m(r_i) * A(r_i).$$

式中: $A(r_i)$ 表示第 $i$ 个区域的面积百分比, $m(r_i)$ 表示第 $i$ 个区域的综合置信度.



(a)原始图像

(b)分割图像

图3 雪景图像及其分割图像

Fig. 3 Snow scene image its partition image

表3 情感因子“悠闲的”的置信度计算过程

Table 3 The calculation process of the belief of “relaxed”

区域	内 容
	颜色=蓝色 $m_1(\text{Yes})=0.8, m_1(\text{No})=0.1, [\text{Bel}, \text{Pl}]= [0.8, 0.9]$ . 内容=晴天
区域1	$m_2(\text{Yes})=0.6, m_2(\text{No})=0.3, [\text{Bel}, \text{Pl}]= [0.6, 0.7]$ . $\Rightarrow$ 冲突度 $k=m_1(\text{Yes})m_2(\text{No})+m_2(\text{Yes})m_1(\text{No})=0.3$ $\Rightarrow$ 融合置信度 $m(\text{Yes})=m_1+m_2(\text{Yes})/(1-k)=0.62/(1-0.3)=0.8857, m(\text{No})=0.1$ .
区域2	颜色=蓝色, 内容=晴天 $m(\text{Yes})=0.8857, m(\text{No})=0.1$ .
区域3	颜色=白色, 内容=云彩 $m(\text{Yes})=0.8478, m(\text{No})=0.125$ .
区域4	颜色=白色, 内容=雪地 $m(\text{Yes})=0.8056, m(\text{No})=0.1752$ .
区域5	颜色=黑色, 内容=小山丘 $m(\text{Yes})=0.2632, m(\text{No})=0.6316$ .
总结	$P(\text{Yes})=0.727265, P(\text{No})=0.23657, P(\text{Yes}) > P(\text{No})$ 决策: 是的, 这幅图像令我感觉悠闲.

上述计算过程同样适用于其他情感因子.计算求得整幅图像对各情感因子的综合置信度后,采用如下规则确定图像的最终情感类别:1)情感类别应具有最大的综合可信度,且其综合可信度与其他情感因子的综合可信度之差必须大于某一域值(根据实验经验本文规定此阈值为0.05);2)情感类别的不确定性概率 $P(\text{Yes}) - P(\text{No})$ 必须小于某一域值,

且小于该情感因子的综合置信度值,根据实验经验本文规定该阈值为 0.75.

#### 4 实验结果与分析

以 Matlab 为工作平台,开发了一个图像情感标识原型系统,系统分为训练和测试 2 个阶段.在训练阶段,从 Corel 图片库中选择 120 幅自然风景图像,从中手工分割出 200 个图像区域作为训练样本,训练图像内容分类器 SVM.这 200 个图像区域涵盖了狭义关键词描述的 44 种语义内容,并且为每个狭义关键词提供了至少 2 个图像样本.通过 3 重交叉验证和网格搜索算法实验,发现  $C = 2^7$ ,  $\gamma = 2^{-11}$  能获得最小的 3 重交叉验证误差,因此采用这 2 个核参数值代表训练好的 SVM.

另一方面通过小规模问卷调查来获得各种颜色和各语义内容的 D-S 不确定性区间,即得出各种颜色和各个语义内容给观察者带来“悠闲的、宁静的、清醒的、精力充沛的”这 4 种情感认知的概率范围.由于颜色及语义内容的多样性,让参与者回答所有颜色和语义的情感认知工作量太大,进而会影响调查的准确性.本实验请 80 位在校大学生参与,同时将 20 幅颜色图片(包括黑色、白色各一幅,其他颜色按照亮度不同各 2 幅)和上述 200 个手工分割的图像区域组成问卷调查图片库.这些图片被随机分为 22 组,每组由 10 幅图像组成.每位学生只需选择一组图像完成问卷调查即可.根据卷面结果统计出各种颜色和各个语义内容图像引起 4 种情感认知的概率范围.限于篇幅只列出部分颜色及图像内容对“悠闲的”、“清醒的”这 2 种情感因子的 D-S 不确定性区间(如表 4 所示).

表 4 部分颜色及图像内容对 2 种情感因子的不确定性区间  
Table 4 Uncertainty ranges of some colors and objects for “relaxed” and “clear-headed”

颜色名/内容	悠闲的	清醒的
黑色	[0.2, 0.4]	[0.5, 0.8]
灰色	[0.2, 0.7]	[0.2, 0.5]
红色	[0.4, 0.8]	[0.56, 0.86]
绿色	[0.8, 0.9]	[0.67, 0.9]
蓝色	[0.8, 0.9]	[0.83, 0.93]
白色	[0.6, 0.9]	[0.8, 0.95]
雪地	[0.67, 0.8]	[0.78, 0.98]
海洋	[0.4, 0.6]	[0.48, 0.6]
晴天	[0.6, 0.7]	[0.67, 0.97]
云彩	[0.72, 0.92]	[0.72, 0.82]
小山丘	[0.3, 0.7]	[0.25, 0.75]

在测试阶段,从 Corel 图片库中另选 200 幅自然风景图像,采用像素聚类法先将各图像分割成 3~5 个区域,然后用训练好的 DAGSVM 对每一区域的内容进行分类识别,并根据“识别率 = 正确识别区域

数/总分割区域数”统计出平均识别率为 78.2%.从识别结果可以看出,颜色、纹理特征越统一、越鲜明的图像区域,如雪山、蓝天、绿草地等,其识别效果越好,可达到 90% 左右;而对于建筑物和花等对图像形状特征要求较高的语义类,识别效果要差很多,只有约 50%.最后采用 D-S 证据理论计算每幅图像对各情感因子的综合置信度并判断最终情感类别.

为了验证情感标识的准确性,为系统增加了检索功能,检索关键词为情感因子.图 3 显示了 200 幅图像中关键词为“清醒的”的部分检索结果.

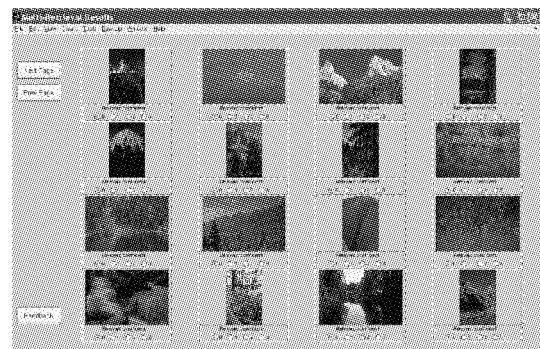


图 4 关键词为“清醒的”的部分检索结果

Fig. 4 A part of retrieval results of “clear-headed”

然后采用最常用的检索查准率 (precision,  $P_r$ ) 和查全率 (recall,  $R_e$ ) 进行检索效果度量.检索查准率和查全率的定义分别为  $P_r = N_c/N_r$ ,  $R_e = N_c/N_e$ , 其中  $N_c$  代表检索到符合用户此种情感认知的图像数,  $N_r$  代表被标识系统标上该关键词的图像数,  $N_e$  表示系统用户手工标注该关键词的图像数.随机抽选 8 位在校大学生参加检索测试,他们的检索查准率和查全率差距不大,限于篇幅,列出平均查准率最高和最低 2 位实验者的数据(如表 5 所示).

表 5 扩大语义内容的不确定性区间前后的情感检索结果  
Table 5 Precision and recall of affective retrieval before and after enlarging uncertainty range %

情感因子	扩大语义内容		扩大 30% 语义内容	
	不确定区间前		不确定区间后	
	查准率	查全率	查准率	查全率
悠闲的	77.78	84.00	82.40	84.0
平静的	83.80	77.50	92.30	90.0
清醒的	92.30	80.00	96.40	88.3
精力充沛	78.90	90.00	89.10	98.0
悠闲的	82.50	68.75	85.70	75.0
平静的	66.67	80.90	79.20	90.5
清醒的	83.80	89.60	91.60	94.8
精力充沛	85.10	76.90	90.00	86.5

考虑到图像内容识别会引入一定误差,按 30% 的比例扩大了图像语义内容对各情感因子的不确定性区间.例如,将“雪地”在“悠闲的”上的不确定性区间从 [0.67, 0.8] 扩大为 [0.650 5, 0.819 5].然

后根据新的语义内容可信度再次计算自动标识结果和上述8位在校大学生的检索查准率和查全率,结果如表5所示。

根据表5,发现对语义内容的不确定性区间进行调整后,对于相同的实验者检索查准率比调整之前要提高8%左右,最高可达96%,最低也在80%左右。虽然达不到文献[7]中的均值为92.37%和96.15%的查准率和查全率,但考虑到文献[7]仅限于对纺织品图像进行标识和检索,图像分析的难度远远小于自然风景图像,因此情感标识系统在图像检索、图像心理疗法等方面具有实用价值。

## 5 结束语

提出一种对自然风景图像应用D-S证据理论融合图像视觉特征及语义内容信息,得出图像情感标识的方案。实验证明D-S证据理论对这种模糊主观、不确定性强的情感推理问题能够取得较好的效果,且易于实现。本文采用的视觉特征只有区域主要颜色,而且颜色的分类只有11种,这在某种程度上限制了情感推理的精确性。因此下一步的研究方向有:1)根据HSV中的饱和度及亮度信息进一步细化颜色分类,例如将红色细分为暗红色、亮红色、粉红色、鲜红色,同时增加对比度、粗糙度、方向性等感性纹理特征作为图像视觉特征。这些新增颜色及感性纹理特征能帮助系统获得更高精度的标识结果。2)通过进一步邀请各行各业、各种教育背景的人员参与确定不确定性区间的问卷调查,扩大系统的适用范围。

## 参考文献:

- [1] KOBAYASHI Y, KATO P. Multi-contrast based texture model for understanding human subjectivity[C]//15th International Conference on Pattern Recognition. Barcelona, Spain, 2000, 3: 917-922.
- [2] 毛峡,丁玉宽,牟田一弥. 图像的情感特征分析及其和谐感评价[J]. 电子学报, 2001, 29(12A): 1923-1927.  
MAO Xia, DING Yukuan, MUTA I. Analysis of affective characteristics and evaluation on harmonious feeling of image[J]. Acta Electronica Sinica, 2001, 29(12A): 1923-1927.
- [3] CHO S B, LEE J Y. A human-oriented image retrieval system using interactive genetic algorithm[J]. IEEE Trans on Systems, Man, and Cybernetics, 2002, 32(3): 452-458.
- [4] WANG S F, WANG X F. A double-level emotion image retrieval model[J]. Journal of System Simulation, 2004, 16(9): 2074-2078.
- [5] 王伟凝, 贺前华. 基于色彩描述的图像情感语义查询[J]. 华南理工大学学报: 自然科学版, 2008, 36(1): 60-66.  
WANG Weining, HE Qianhua. Emotion-based image semantic query through color description[J]. Journal of South China University of Technology: Natural Science Edition, 2008, 36(1): 60-66.
- [6] BAEK S, HWANG M, CHUNG H, KIM P. Kansei factor space classified by information for Kansei image modeling[J]. Applied Mathematics and Computation, 2008, 205(2): 874-882.
- [7] SHIN Y, KIM Y, KIM E Y. Automatic textile image annotation by predicting emotional concepts from visual features[J]. Image and Vision Computing, 2010, 28(3): 526-537.
- [8] BLEKAS K, LIKAS A, GALATSANOS N P, LAGARIS I E. A spatially-constrained mixture model for image segmentation[J]. IEEE Transactions on Neural Networks, 2005, 16(2): 494-498.
- [9] MANJUNATH B S, MA W Y. Texture features for browsing and retrieval of image data[J]. IEEE Trans on PAMI, 1996, 18(8): 837-842.
- [10] 陈立伟, 王文姝, 王颖芳. 基于模糊理论的关键词识别算法[J]. 应用科技, 2010, 9: 5-8.  
CHEN Liwei, WANG Wenshu, WANG Yingfang. Keyword recognition algorithm based on fuzzy theory[J]. Applied Science and Technology, 2010, 9: 5-8.
- [11] WANG Y, HUANG S T. Reducing the number of sub-classifiers for pairwise multi-category support vector machines[J]. Pattern Recognition Letters, 2007, 28: 2088-2093.

### 作者简介:



高彦宇,女,1975年生,博士,讲师。主要研究方向为图像处理与模式识别、感性工学、人工智能。



尹怡欣,男,1957年生,教授,博士生导师,博士。主要研究方向为智能控制及自适应控制、人工生命、情感计算。