

# 基于LBP和小波纹理特征的 室内室外场景分类算法

李锦锋 许勇

(华南理工大学计算机科学与工程学院, 广州 510006)

**摘要** 室内室外场景识别是图像处理的基本问题之一。对此问题提出了一种不需要图像分割和目标识别的方法。首先,利用小波变换对原图像进行五层小波分解,然后用旋转不变的LBP(local binary pattern)算法对第二层至第五层的LL子图提取纹理特征,再计算第二层和第一层LH,HL,HH子图的能量均值和方差,最后连接这两个过程形成的低维特征向量进行室内室外图像场景分类。实验结果表明此算法分类效果比较好,且识别室内室外场景并不需要目标形状等局部详细信息,只需要图像模糊化的全局纹理信息。

**关键词** 室内室外场景识别 旋转不变 LBP 小波变换 纹理特征

中图法分类号: TN919.8 文献标志码: A 文章编号: 1006-8961(2010)05-742-07

## Indoor-outdoor Scene Classification Algorithm Based on the Texture Feature of LBP and Wavelet

LI Jinfeng, XU Yong

(School of Computer Science & Technology, South China University of Technology, Guangzhou 510006)

**Abstract** The scene recognition between indoor and outdoor is one of the basic problems in image processing. A solution to this problem which does not require image segmentation and object recognition is proposed in this paper. First of all, a 5-level wavelet transform was implemented on the image, then the rotation invariant LBP algorithm was used to extract the texture feature from 2-level to 5-level LL sub-images. Afterwards, the mean and variance of the LH, HL, HH sub-images from 2-level to 1-level were computed. Finally, these two low-dimensional eigenvectors were concatenated as the image feature to classify the indoor and outdoor scene images. The experiment results indicate that our algorithm performs well on indoor/outdoor classification, and the overall texture information of the fuzzy image is only needed in the scene recognition between indoor and outdoor, rather than the local detailed information such as object shape.

**Keywords** indoor-outdoor scene recognition, rotation invariant, LBP, wavelet transform, texture feature

## 0 引言

场景识别是图像处理的基本问题之一。一张图片的内容是室内还是室外的,是自然的还是人造的,这些最基本的问题对理解景物图片起了很重要的作用,为进一步识别具体的目标提供了依据。比如高

速公路上的实体汽车跟办公桌上的汽车模型,如果没有场景识别作为前提条件,仅仅根据局部区域特征,计算机将很难准确区分它们。

Oliva指出,图像的特征描述包含下层、基层和上层3个层次,这3个层次分别对应着图像不同的抽象描述以及语义信息<sup>[1]</sup>。

下层:这个层次的特征描述需要分析图像的局

基金项目:国家自然科学基金项目(60603022);国家863项目(2006AA01Z196)

收稿日期:2008-07-15;改回日期:2009-02-04

第一作者简介:李锦锋(1983—),男。2009年于华南理工大学获计算机软件与理论专业硕士学位。研究方向为数字图像处理。E-mail: li.jinfeng@mail.scut.edu.cn

通讯作者:许勇。E-mail: yxu@scut.edu.cn

部结构。典型的代表是Lazebnik提出的金字塔模型<sup>[2]</sup>,但这种方法因局部信息计算复杂度比较高,不太适用于场景识别。

基层:这个层次的特征描述跟类别有很大关系,比如森林、山和街道。Lipson使用图像预先确定区域的空间和光度关系来建立全局的图像特征<sup>[3]</sup>,Oliva和Torralba则提出用图像的空间信封模型来建立一个全局特征<sup>[1]</sup>。

上层:这个层次具有很高的抽象意义,因此,它只具有很少的类别:室内和室外或者城市风景和自然风景。Szummer和Picard描述了一个室内室外的分类方法<sup>[4]</sup>。

综上所述,室内室外图像分类实际是属于最高层次的场景识别,Oliva和Torralba指出,场景识别应该是基于基层特征和上层特征的,是不用图像的局部信息的<sup>[1]</sup>,他们提出的空间信封模型,提出场景由纹理性、开放性、粗糙性、自然性、扩展性和倾斜性组成,比较适用于场景分类,但对于这些性质的特征向量,不仅难以定义其训练样本(比如开放性,哪些图像具有开放性,更大程度上依赖于个人的认定),而且特征向量的提取需要用到主分量分析方法,复杂度比较大。

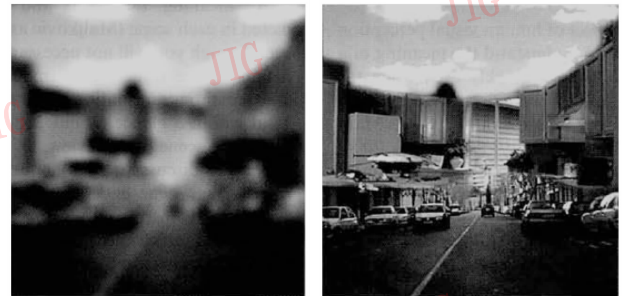
目前在研究室内室外场景分类的算法中,都是根据图像下层特征来分类的,如Szummer和Picard<sup>[4]</sup>的算法是对图像进行网格划分,对每个格子小图像计算其颜色特征和纹理特征,再综合这两个特征进行分类。Payne和Singh则指出室内图像因墙壁、桌椅等因素,其边缘比起室外图像更具有水平性和垂直性的特点,进而提出了一个利用图像的这些下层特征——边缘信息来进行室内室外场景分类的算法<sup>[5]</sup>。

为了克服Szummer<sup>[4]</sup>和Payne<sup>[5]</sup>仅利用局部特征的缺点,提出了一个基于图像全局特征的算法来进行室内室外分类。首先,利用小波变换对原图像进行5层的小波分解,然后应用旋转不变的LBP算法提取第2层到第5层LL子图的纹理特征,再计算第2层和第1层LH,HL,HH子图的能量均值和方差,最后连接这两个过程形成的低维特征向量形成图像特征,并据此进行室内室外图像场景分类。

## 1 室内室外场景识别的关键

Oliva和Torralba<sup>[1]</sup>认为人认识一个场景应该是

根据人眼闪过图片的一刹那对图像的“认知”,如图1所示,只需要浏览图1(a)的图像大概信息便能识别到这张图片是属于街道的,并不需要对图1(b)提取车辆、建筑等这些具体信息。



(a) Gabor 变换模糊图

(b) 原图

图1 图像 Gabor 变换

Fig. 1 The Gabor transform on image

Walker<sup>[6]</sup>也认为对图像简单的纹理分析能够寻找到场景分类的线索,而且基于这样简单的纹理分析能够使图像识别过程和计算简单化。也就是说,场景识别应该是对图像模糊预处理后的第2次处理,模糊预处理的效果就是类似图1(a)的形式。

模糊预处理的常用技术有Gabor变换和小波变换。Walker<sup>[6]</sup>和Oliva<sup>[1]</sup>利用了Gabor技术,但应用Gabor技术处理的缺点是特征向量维数比较多。本文使用小波技术,既可以模糊预处理图像又可对图像降维,效果比较好。

小波分析的基本思想是,通过小波变换将图像分解为不同的频域子图,计算频域的能量等参数作为纹理描述子。由于其具有良好的局域特性,所以频域仍然具有原图的某些视觉特性,如图2所示。

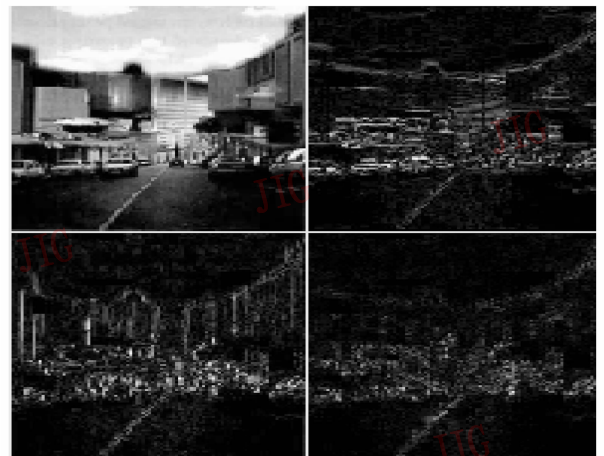
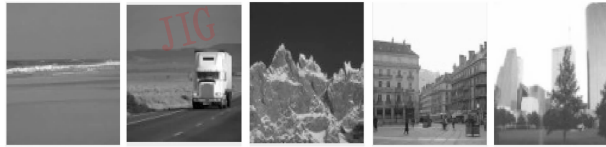


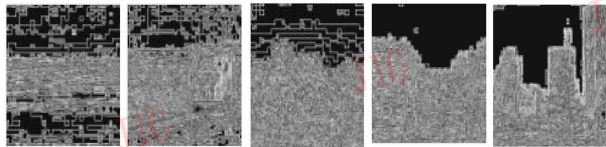
图2 图1(b)的小波变换图

Fig. 2 The wavelet transform images of Fig. 1(b)

而对于室内室外图像,除了模糊预处理之外,室外图像一般还包含有天空、沙地和路等信息。比如图 3(a)室外图像和图 3(c)室内图像,可以看出它们之间明显的不同。因此,室内室外场景识别的算法除了应具有场景识别的模糊纹理特征外,还应具有能从上到下区分天空、沙地和路等信息的能力,图 3(b)和图 3(d) RILBP(rotation invariant LBP)纹理图是用本文算法提取出来的,从中可以看出本文算法能够较好地提取这些信息。



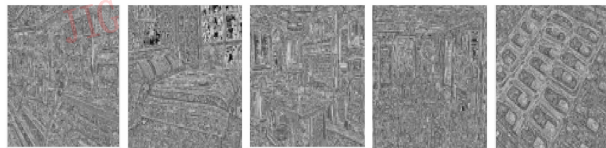
(a) 室外图像



(b) 室外图像 RILBP 纹理图



(c) 室内图像



(d) 室内图像 RILBP 纹理图

图 3 RILBP 效果

Fig. 3 The effects of RILBP

## 2 局部二值模式 (LBP)

Ojala 介绍了局部二值模式 (LBP) 纹理描述子<sup>[7]</sup>,对于图像中的一个  $3 \times 3$  邻域 (如图 4),9 个区域像素的灰度值记为

$$G = \{g_c, g_0, g_1, \dots, g_7\} \quad (1)$$

式中,  $g_c$  为区域中心像素的灰度值,  $g_n (n = 0, 1, \dots, 7)$  为中心像素周边像素的灰度值。

再定义纹理元  $T$  (texture) 为

$$T = \{p_0, p_1, \dots, p_7\} \quad (2)$$

$g_0$	$g_1$	$g_2$
$g_7$	$g_c$	$g_3$
$g_6$	$g_5$	$g_4$

(a) 图像  $3 \times 3$  区域的灰度表示

$p_0$	$p_1$	$p_2$
$p_7$		$p_3$
$p_6$	$p_5$	$p_4$

(b) 图像  $3 \times 3$  区域的纹理元表示

图 4 图像  $3 \times 3$  区域及其纹理元

Fig. 4  $3 \times 3$  image region and its texture

式中

$$p_n = \begin{cases} 1 & g_n \geq g_c \\ 0 & g_n < g_c \end{cases} \quad n = 0, 1, \dots, 7 \quad (3)$$

纹理元的每个元素均有两种取值可能,所以纹理元共有  $2^8$  种取值组合,定义  $lbp$  如式(4)所示:

$$lbp = \sum_{n=0}^7 2^n p_n \quad (4)$$

计算例子如图 5,最后其  $lbp$  值为 241,代入原来图像区域的中间像素。这样,遍历图像的每一个像素,形成了一个图像的纹理值图,如图 6 所示。统计整幅图像中纹理单元的分布频率即可得到 256 维的纹理谱,这可使用直方图方法实现。

example	thresholded	weights
6 5 2	1 0 0	1 2 4
7 6 1	1 1 0	128 8
9 8 7	1 1 1	64 32 16

$$lbp = 1 + 16 + 32 + 64 + 128 = 241$$

图 5  $lbp$  计算例子

Fig. 5 The calculation example of  $lbp$



(a) 原图

(b) 纹理值图

图 6 图像及其纹理值图

Fig. 6 Original image and its texture image

需要说明的是:

1) Heikkila<sup>[8]</sup>指出上面的 LBP 算法在图像的平

坦区域并不具有鲁棒性。例如对于天空区域的灰度值,它的  $3 \times 3$  邻域像素的灰度值很接近于中心点的灰度值,假设其中心值是 30,周边值是 29,那么按照式(4)的计算方法,其  $lbp$  值是 0。但如果中心值不变,周边值是 30,其  $lbp$  却变为 256。为了解决这个问题,Heikkila<sup>[8]</sup> 提出式(3)中的阈值应变为中间值减去一个常值  $a$ ,在实验中, $a$  取为 3。

人工选取阈值并不能有效处理各种各样的图像场景,Murat 提出另外一种思路:对于  $3 \times 3$  邻域,阈值为邻域内 9 个像素的平均灰度值,大于阈值的取 1,小于或等于阈值的取 0<sup>[9]</sup>。这样在上面所举的例子中,两种情况下得到的  $lbp$  值都是 0,符合局部纹理的特点。本文将采取 Murat 提出的纹理元计算方法<sup>[9]</sup>,即式(3)变为

$$p_n = \begin{cases} 1 & g_n > A \\ 0 & g_n \leq A \end{cases} \quad (5)$$

式中, $A$  为  $3 \times 3$  邻域的均值。

2) 图 7 中 3 个例图具有视觉一致性,应为同一纹理模式,但是如果按 Ojala<sup>[7]</sup> 算法计算,可以得到图 7 从左到右的  $lbp$  值分别为 130,160 和 10。也就是说,对于在视觉上一致的这 3 个二值图,按 Ojala<sup>[7]</sup> 的方法计算,却得到不同纹理模式了。

为了改进这个缺点,需要考虑图像的旋转不变性。图 7(b)按顺时针旋转  $270^\circ$ ,或者按逆时针旋转  $90^\circ$ 可以旋转到图 7(a)一致的模式,同样的道理,图 7(c)按顺时针旋转  $90^\circ$ ,或者按逆时针旋转  $270^\circ$ ,也可以旋转到图 7(a)一致的模式。图 7 中黑色为 1,白色为 0。这种旋转不变性,在场景分析中是一种很重要的性质,可以用一个现实的例子来解释,假如屋子里的桌子上放着一个瓶子,无论摆放者怎样摆放,无论照相者从哪个角度拍照,都不影响这张照片是室内照片的事实。

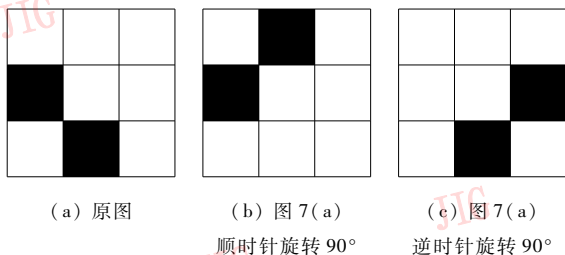


图 7 旋转不变例子

Fig. 7 Examples of rotational invariance

根据式(2)纹理元  $T$  的计算,图 7 中 3 个例图的纹理元分别为:0000 0101, 0100 0001, 0001 0100,

按照上面的分析,这 3 个纹理元应为同一纹理模式,即经过旋转三者的纹理元均为 0000 0101。用同样的方法处理 Ojala<sup>[7]</sup> 中的 256 种不同纹理元,最后可以归纳为下面 30 种纹理元:

- 1:0000 0000    11:0010 0101    21:0010 1111
- 2:0000 0001    12:0000 1111    22:0011 0111
- 3:0000 0011    13:0001 0111    23:0101 0111
- 4:0000 0101    14:0010 0111    24:0101 1011
- 5:0000 1001    15:0010 1011    25:0011 1111
- 6:0001 0001    16:0011 0011    26:0101 1111
- 7:0000 0111    17:0110 0011    27:0110 1111
- 8:0000 1011    18:0110 1011    28:0111 0111
- 9:0001 0011    19:0101 0101    29:0111 1111
- 10:0001 0101    20:0001 1111    30:1111 1111

在实验中,也使用直方图的方法来统计纹理元出现的频率,得到纹理谱,称这种纹理谱计算方法为 RILBP。由于 RILBP 有效体现了图像的局部特征和旋转不变性,它能很好地区分出室内室外图像,从图 3(c) 室外图像 RILBP 纹理图和图 3(d) 室内图像 RILBP 纹理图,从上、中、下 3 个层次可以明显看出视觉效果的区别。

### 3 室内室外场景分类算法

#### 3.1 室内室外场景纹理特征提取算法(ST 算法)

1) 对图片进行五层小波变换(在实验中,使用的图片都已经转换为  $256 \times 256$  像素),这样就得到了从  $256 \times 256$  到  $8 \times 8$  的分解子图, $4 \times 4$  以下对于 9 个像素的模板已经没有意义就不再分解(如图 8 所示);

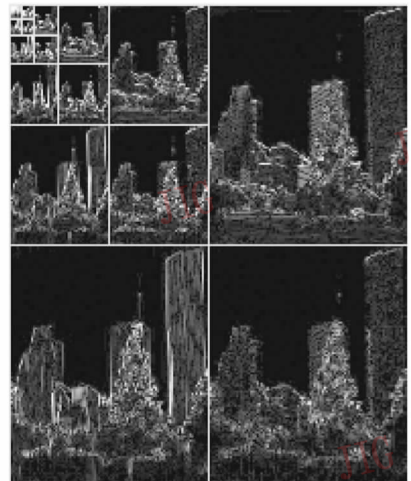


图 8 图 6(a) 的五层小波变换图

Fig. 8 5-level wavelet transform images of Fig. 6(a)

2) 小波逆变换提取图像的低频部分,在实验中,提取了从  $64 \times 64$  到  $8 \times 8$  低频部分,这样形成一个金字塔模型,如图 9 所示。

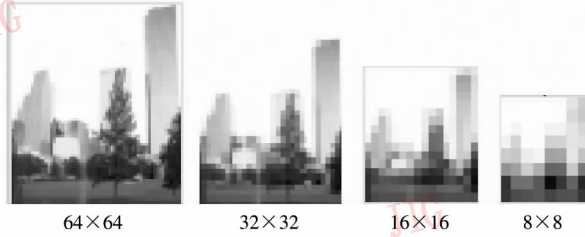


图 9 图 8 小波逆变换的 LL 部分

Fig. 9 The LL parts of the inverse wavelet transform in Fig. 8

对这 4 个低频部分,分别用第 2 节提到的 RILBP 算法计算其纹理特征,分别记为  $F_{64}$ ,  $F_{32}$ ,  $F_{16}$ ,  $F_8$ 。

最后对这 4 个特征向量进行如下的加权处理:

$$F_1 = 0.4 \times F_{64} + 0.3 \times F_{32} + 0.2 \times F_{16} + 0.1 \times F_8 \quad (6)$$

这样做的原因是:

(1) 金字塔模型既能体现局部的 RILBP 纹理特征,也能体现全局的 RILBP 纹理特征,从而使得特征向量  $F_1$  既包含图像的局部信息也包含图像的全局信息。

(2) 金字塔模型中,底层图像比顶层图像清晰,底层特征向量比顶层特征向量表示图像更多的特征信息,权重也应该比较大。另外,对于归一化处理,加权系数和应该为 1,实验结果表明:当  $F_{64}$ ,  $F_{32}$ ,  $F_{16}$ ,  $F_8$  这些特征向量加权系数分别取 0.4, 0.3, 0.2, 0.1 时,算法分类的结果最优。

3) 小波处理后的图像高频信息的纹理特征提取算法是:分别提取  $64 \times 64$  和  $128 \times 128$  的 LH, HL, HH 部分,对这 6 个部分的每部分计算其能量均值和方差,从而形成 12 维的特征向量  $F_2$ 。

4) 30 维的  $F_1$  特征向量值范围在  $0 \sim 1$  区间,12 维的  $F_2$  特征向量值则在  $0 \sim 100$  区间,为了使得特征向量归一化,把  $F_2$  特征向量值除于 100,再连在  $F_1$  特征向量后面,从而形成 42 维的特征向量  $F_0$ 。

5) Shu<sup>[10]</sup> 指出,为了保留不同局部区域的能量特征,应对原图的梯度图(如图 10)也做 RILBP 处理,因此在这里先计算图像的梯度图,再对其进行如上 1)~4) 步骤的处理,也得到一个 42 维的特征向量,称为  $F_g$ 。

6) 计算总的特征向量  $F$ ,基于数据归一化的考

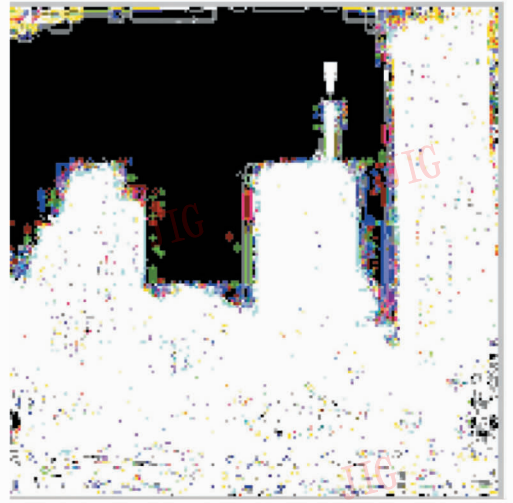


图 10 图 6(a) 的梯度图

Fig. 10 The gradient image of Fig. 6(a)

虑,取  $F_0$  和  $F_g$  的权重各为 0.5,即

$$F = 0.5 \times F_0 + 0.5 \times F_g \quad (7)$$

在本文中称此算法为 ST(scene texture)算法。

### 3.2 分类判别方法——SVM

为了实现对室内室外图像的分类,在实验中使用支持向量机(SVM)进行分类。

杨志民<sup>[11]</sup> 概括支持向量机的基本思想为:寻找一个最优分类超平面,使得训练样本中的两类样本点能被无错误地分开,并且要使两类样本点的分类间隔最大。而对线性不可分问题,通过核函数将低维输入空间的数据映射到高维空间,从而将原低维空间的线性不可分问题转化为高维空间上的线性可分问题,然后在这个新空间中求取最优分类面。

这里,采用径向基(rbf)函数作为核函数(如式(8)所示),输入是 ST 算法处理后的 42 维特征向量  $F$ 。

$$k(x, x_i) = e^{-\frac{|x-x_i|^2}{\delta^2}} \quad (8)$$

## 4 实验

### 4.1 数据库

目前室内室外场景并没有标准的数据库,为了进行相关实验,整理了两个数据库:数据库 1 是从 UIUC 的计算机视觉研究小组<sup>[12]</sup> 提供的 15 类图片数据库上整理的,其中室内图像有 1 245 张,室外图像有 1 400 张,在实验中称此图库为 uiuc 数据库,部分例图如图 11 所示。



(a) 室内图像



(b) 室外图像

图 11 Uiuc 图像数据库

Fig. 11 Uiuc image database

数据库 2 是从 google 图片库整理而来的,其中室内图像 437 张,室外图像 438 张,在实验中称此图库为 google 数据库,例图如图 12 所示。



(a) 室内图像



(b) 室外图像

图 12 Google 图像数据库

Fig. 12 Google image database

### 4.2 实验步骤

1) 按第 3.1 节介绍的 ST 算法分别对 uiuc (google) 数据库中 2 645 (875) 张图片计算 42 维特征向量;

2) 对 uiuc 数据库 (google 数据库), 随机采样一些样本放到 SVM 中训练判别函数, 其中 SVM 使用 Anton<sup>[13]</sup> 提供的代码, 核函数使用第 3.2 节说明的径向基核函数, 其参数  $\sigma$  取为 0.01, 约束优化条件取为 100;

3) 取图库中非训练的图片进行测试;

4) 为了减少误差, 对步骤 2) ~ 3) 重复进行 50 次, 取 50 次均值作为最终的分类结果。

### 4.3 实验结果比较

为了与 ST 算法做比较, 修改 Oliva 和 Torralba<sup>[1]</sup> 提供的八类室外场景分类方法为室内室外分类算法, 并称此修改的算法为 gist 算法。

用第 4.2 节提出的实验步骤对两个数据库进行分类实验, 其中 uiuc 数据库样本数分别取值 50, 100, 150, 200, 400, google 数据库则取值 10, 20, 40, 80, 160。结果如图 13 和图 14 所示。对 uiuc 数据库, 当样本数为 400 时, gist 算法的分类成功率为 96.035 8, ST 算法为 96.639 5; 对 google 数据库, 当样本数为 160 时, gist 算法的分类成功率为 94.221, ST 算法为 94.545 7。

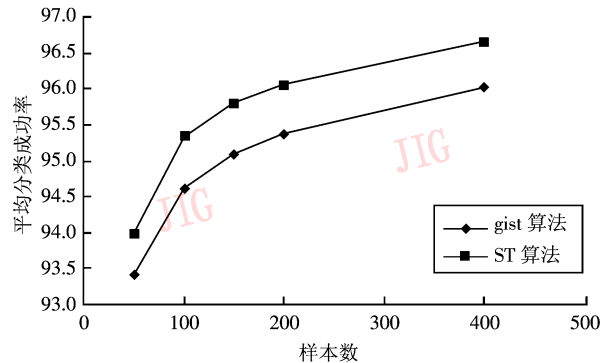


图 13 Uiuc 数据库分类结果

Fig. 13 The classification results on uiuc database

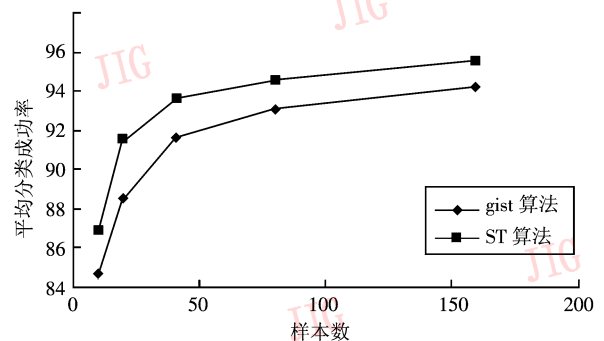


图 14 Google 数据库分类结果

Fig. 14 The classification results on google database

从两个图中, 可以看出 ST 算法无论在 uiuc 数据库上还是 google 数据库上, 其 50 次平均分类成功率都大于 gist 算法, 且 ST 算法的特征向量维数 42 远远少于 gist 算法中的 512 维。

实验表明, 提出的 ST 算法分类效果较好, 原因就在于 ST 算法不但可以用较少的维数包含模糊纹理信息, 还能像图 3 所示那样包含区分室外图像的上、中、下层的较全面的图像信息。而 gist 算法只是计算像图 1(a) 那样的模糊纹理特征信息。

## 5 结 论

分析了室内室外场景图像识别的要素, 并针对

这些要素提出了应用于室内室外场景图像识别的 ST 算法,得到了较好的实验效果。当然,对于场景识别的要素还很多,ST 算法还有很多需要改进的地方,为了使其应用于更广泛的场景识别邻域,还需要对此作进一步研究和优化。

### 参考文献 (References)

- [1] Oliva A, Torralba A. Modeling the shape of the scene: A holistic representation of the spatial envelope[J]. *International Journal of Computer Vision*, 2001, 42(3): 145-175.
- [2] Lazebnik S, Schmid C, Ponce J. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. New York, NY, USA: IEEE Computer Society, 2006, 2: 2169-2178.
- [3] Lipson P, Grimson E, Sinha P. Configuration based scene classification and image indexing[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Washington DC, USA: IEEE Computer Society, 1997: 1007-1013
- [4] Szummer M, Picard R W. Indoor-outdoor image classification[C]// *Proceedings of IEEE International Workshop on Content-Based Access of Image and Video Databases*. Bombay, India; [s. n.], 1998: 42-51.
- [5] Payne A, Singh S. Indoor vs outdoor scene classification in digital photographs[J]. *Pattern Recognition*, 2005, 38(10): 1533-1545.
- [6] Walker Renninger L, Malik J. When is scene identification just texture recognition? [J]. *Vision Research*, 2004, 44(19): 2301-2311.
- [7] Ojala T, Pietikainen M, Harwood D. A comparative study of texture measures with classification based on feature distributions[J]. *Pattern Recognition*, 1996, 1(29): 51-59.
- [8] Heikkila M, Pietikainen M. A texture-based method for modeling the background and detecting moving objects [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(4): 657-662.
- [9] Murat H, Liu Wei, Tong Qinye. Content-based image retrieval using texture spectrum histogram and latent semantic indexing[J]. *Bulletin of Science and Technology*, 2006, 22(3): 389-394. [木拉提·哈密提, 刘伟, 童勤业. 纹理谱直方图与潜在语义标引在图像检索中的应用[J]. *科技通报*, 2006, 22(3): 389-394.]
- [10] Liao Shu, Fan Wei, Chung Albert C S, et al. Facial expression recognition using advanced local binary patterns, Tsallis entropies and global appearance features [C]// *Proceedings of IEEE International Conference on Image Processing*. New York, NY, USA: IEEE Press, 2006: 665-668.
- [11] Yang Zhiming, Liu Guangli. *The Principle of the Uncertain Support Vector Machine and Its Application* [M]. Beijing: Science Press, 2007. [杨志民, 刘广利. 不确定性支持向量机原理及应用[M]. 北京: 科学出版社, 2007.]
- [12] Ponce J. The uiuc database from UIUC research group [DB/OL]. (2006-06-17) [2008-08-10]. [http://www-cvr. ai. uiuc. edu/ponce\\_grp/data/index. html](http://www-cvr. ai. uiuc. edu/ponce_grp/data/index. html).
- [13] Schwaighofer A. SVM code which is implemented by Anton Schwaighofer[CP/OL]. (2001-06-05) [2008-08-15]. <http://ida. first. fraunhofer. de/~anton/software. html>.