

# 基于SVM与颜色/纹理组合特征的景物识别算法

雷宝权 杨丽华 程咏梅 赵春晖 吴燕茹

(西北工业大学自动化学院 西安 710072)

**摘要** 受多种因素影响,室外场景变换复杂,因此利用单个特征(通常多使用颜色或纹理)完成室外场景的识别,不能达到满意的效果。首先基于RGB空间的色直方图进行颜色特征提取,然后基于Gabor滤波器进行纹理特征提取,最后将两种特征结合,提出了基于SVM与颜色/纹理组合特征的景物识别算法。基于美国加州理工学院的Pasadena Houses2000数据库建立了室外场景中天空、道路、房子、树木和草地等5类样本训练库,进一步完成了基于SVM的单一特征和颜色/纹理组合特征的景物识别。实验结果表明,该算法对仅用一种视觉特征无法区分景物的室外场景图像能取得较好的分类结果。

**关键词** 景物识别, Gabor滤波器, 颜色直方图, 颜色/纹理组合特征, SVM

中图分类号 TP391 文献标识码 A

## Natural Object Recognition Algorithm Based on SVM and Color/Texture Combination Features

LEI Bao-quan YANG Li-hua CHENG Yong-mei ZHAO Chun-hui WU Yan-ru

(College of Automation, Northwestern Polytechnical University, Xi'an 710072, China)

**Abstract** Affected by many factors, outdoor scenes vary greatly, so using a single feature (color or texture) to complete the outdoor scenes recognition can not achieve satisfactory results. A natural object recognition algorithm based on SVM and color/texture combination features was presented. Firstly, the color histogram based on the RGB color space was extracted as color feature. Then, the texture feature was extracted based on multi-channel Gabor filters. At last, the color/texture combination features were presented. An image database of training samples including sky, road, house, tree and grass was created, which is obtained from Pasadena Houses2000 database of California Institute of Technology. And the natural object recognition based on SVM using a single feature and color/texture combination features was completed respectively. Experimental results show that this algorithm has good recognition effect on the images in which each natural object varies greatly from each other.

**Keywords** Natural object recognition, Gabor filter, Color histogram, Color/Texture combination features, SVM

## 1 引言

景物识别是对场景图像的各类景物进行识别分类,在很多领域,如图像检索、精确制导、移动机器人视觉导航、道路图像理解系统以及地图绘制等都得到了广泛的应用。提取与选择能充分反映每类景物属性的特征,是场景图像景物识别的关键问题之一,其有效性对后续处理至关重要。

用于图像识别的特征有很多,而图像的颜色、纹理特征是景物识别中常用的两种视觉特征<sup>[1,2]</sup>。由于室外场景的复杂及可变性,仅使用一种特征往往不能充分描述图像中的各类景物。因此提出将图像的颜色特征和纹理特征进行组合,从而更加准确、多方面地表达图像的内容信息。

SVM在解决小样本、非线性及高维模式识别问题中表现出许多特有的优势<sup>[3,4]</sup>,如泛化性能等,因此近年来受到各个研究领域,如人脸检测<sup>[5]</sup>、手写体数字识别<sup>[6]</sup>、文本自动分类<sup>[7]</sup>、语音识别等越来越多的重视。由于使用的室外场景图

像的样本数量相对偏少,同时场景中的景物分类问题也是一种非线性的模式识别问题,把分类性能卓越的SVM作为景物类别的分类器是一种合理的选择。

针对仅用一种视觉特征无法区分景物的室外场景图像,提出了基于SVM与颜色/纹理组合特征的景物识别算法,完成了基于SVM的颜色/纹理组合特征与单一特征的分类识别,并进行了对比分析。

## 2 颜色特征提取

颜色是图像非常重要的视觉特征,因而利用颜色特征进行图像分类受到重视,并最早得到应用。颜色特征的提取首先要选取合适的颜色空间。在景物识别中,常用的颜色空间有RGB空间<sup>[8]</sup>、HSV空间<sup>[9]</sup>以及CIE的L\*a\*b\*颜色空间<sup>[10]</sup>。RGB模型是目前应用最为广泛的彩色模型,其优点是描述简单,适合硬件实现。因此本文选取RGB颜色空间,根据R,G,B 3个分量提取彩色图像的颜色特征。

到稿日期:2008-12-11 返修日期:2009-03-02 本文受国家自然科学基金重点资助项目(60634030),高等学校博士学科点专项科研基金(20060699032),航空科学基金(2007ZC53037)资助。

雷宝权 博士研究生,主要研究方向为机载信号及信息处理、图像目标检测与识别、惯性导航技术, E-mail: lbqymn@yeah.net.

颜色直方图是被广泛采用的颜色特征<sup>[11]</sup>,具有良好的尺度和旋转不变性、特征鲁棒性好等特点,特别适合描述那些难以进行自动分割的图像。对于用 RGB 颜色空间表示的图像, R, G, B 任何一个分量都可以构成自己的直方图,此直方图依然描述了图像颜色的统计特征。但是一幅图像的 R, G, B 值的范围在 [0, 255] 之间,因此颜色直方图矢量的维数也会非常多。如果直接进行特征计算,运算量将会很大,很难保证快速性。如果对 RGB 空间进行适当量化后再计算直方图,则计算量要小得多。同时,由于人的眼睛对颜色的分辨能力有限,也不需要将这些量全部细分计算,因此在计算特征量之前,应对 RGB 空间进行量化,以减少特征矢量的维数。

由于自然场景图片中,每种景物的颜色差异较大,如天空和树木、道路等,因此采用等间隔量化,对 R, G, B 依据各自的范围均匀分为 4 等份。这样可得到下面量化的 R, G, B 计算式:

$$\begin{aligned}
 R &= \begin{cases} 0 & \text{if } r \in [0, 63] \\ 1 & \text{if } r \in [64, 127] \\ 2 & \text{if } r \in [128, 191] \\ 3 & \text{if } r \in [192, 255] \end{cases} \\
 G &= \begin{cases} 0 & \text{if } g \in [0, 63] \\ 1 & \text{if } g \in [64, 127] \\ 2 & \text{if } g \in [128, 191] \\ 3 & \text{if } g \in [192, 255] \end{cases} \\
 B &= \begin{cases} 0 & \text{if } b \in [0, 63] \\ 1 & \text{if } b \in [64, 127] \\ 2 & \text{if } b \in [128, 191] \\ 3 & \text{if } b \in [192, 255] \end{cases}
 \end{aligned} \quad (1)$$

但是,颜色直方图也存在缺陷。首先,颜色直方图只包含了图像中某一颜色值出现的频数,丢失了像素点所在的位置信息。而且,虽然对任意一幅图像都能唯一地给出与它对应的直方图,但不同的图像却有可能具有相同的直方图特征,亦即直方图与图像之间并非一一对应。为了解决这个问题,本文对图像的每个  $W \times H$  区域提取颜色直方图,得到区域的颜色特征向量:

$$\vec{C} = (R_0, \dots, R_3, G_0, \dots, G_3, B_0, \dots, B_3) \quad (2)$$

其中,  $W=4, H=4, R_i, G_i, B_i$  分别表示区域  $W \times H$  中  $R(w, h)=i, G(w, h)=i, B(w, h)=i$  出现的像素的频数,  $w=1, \dots, W, h=1, \dots, H$ 。

### 3 纹理特征提取

纹理是指人们所观察到的图像像元(或子区域)的灰度变化规律,它是除颜色特征之外另一种应用广泛的图像视觉特征。纹理特征的提取主要有 4 种方法:统计法、结构法、基于模型的方法以及信号处理的方法<sup>[12]</sup>。目前信号处理方法中使用最为广泛的方法是基于空间/频域联合分析法,主要包括 Gabor 变换和小波变换<sup>[13]</sup>,且 Gabor 滤波器提取纹理特征最为有效<sup>[14, 15]</sup>。

用 Gabor 函数形成的二维 Gabor 滤波器不仅有优良的滤波性能,还有与生物视觉系统相近的特点,它的方向和径向频率带宽以及中心频率均可调,在空间域和频域同时达到了最佳的分辨率<sup>[16]</sup>。

Gabor 函数通过定义不同的 Gabor 函数核,就可以得到

一组 Gabor 滤波器。二维 Gabor 滤波器基函数的一般形式<sup>[17]</sup>是:

$$h(x, y, \theta_j, \lambda, \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\pi\left[\left(\frac{x\theta_j}{\sigma_x}\right)^2 + \left(\frac{y\theta_j}{\sigma_y}\right)^2\right]\right\} \cdot \exp\left(\frac{2\pi i x \theta_j}{\lambda}\right) \quad (3)$$

由式(3)可以看出, Gabor 函数是一个被复正弦函数调制的 Gaussian 函数,是一个  $R^2 \rightarrow C$  上的复函数。其中  $\lambda$  和  $\theta_j$  分别为正弦波的波长和方向。 $\theta_j$  的定义如下:

$$\theta_j = \frac{\pi}{n_0}(j-1), j=1, 2, \dots, n_0 \quad (4)$$

其中,  $n_0$  决定了滤波器方向的个数;  $\sigma_x$  和  $\sigma_y$  分别为高斯包络在  $x$  和  $y$  方向上的标准差,它们决定了高斯包络的空间扩展。 $x_{\theta_j}$  和  $y_{\theta_j}$  为:

$$\begin{cases} x_{\theta_j} = +x\cos(\theta_j) + y\sin(\theta_j) \\ y_{\theta_j} = -x\sin(\theta_j) + y\cos(\theta_j) \end{cases} \quad (5)$$

实际上,式(5)是坐标旋转公式,可以获得不同方向的滤波器。式(3)中的参数  $\lambda, \sigma_x$  和  $\sigma_y$ , 反映了 Gabor 滤波器的多尺度特性。通过设置这些参数就可以得到一组不同方向、不同尺度的滤波器组。

Gabor 滤波器可以定义为<sup>[18]</sup>:

$$m(x, y) = O_h(i(x, y)) = |i(x, y) * h(x, y)| \quad (6)$$

其中,  $i(x, y)$  代表图像,  $m(x, y)$  是其滤波输出,  $*$  表示卷积。

基于 Gabor 滤波器对图像库中的图像进行了纹理特征提取。步骤如下:

1) 选取 Gabor 滤波器组的尺度数和方向数,  $n_F=6, n_0=4$ 。Gabor 滤波器组中共有尺度和方向各不相同的滤波器  $6 \times 4=24$  个;

2) 得到上述不同方向和尺度的 24 个 Gabor 滤波器以后,对图像库中的图像经过 Gabor 滤波器组进行滤波。给定一个大小为  $X \times Y$  的图像,图像经每个滤波器的 Gabor 变换之后表示为  $T_{ij}(x, y)$ 。取变换后图像幅值的均值和标准差作为纹理特征:

$$\mu_{ij} = \frac{1}{XY} \sum_{x=1}^X \sum_{y=1}^Y |T_{ij}(x, y)| \quad (7)$$

$$\sigma_{ij} = \frac{1}{XY} \sum_{x=1}^X \sum_{y=1}^Y \sqrt{(|T_{ij}(x, y)| - \mu_{ij})^2} \quad (8)$$

则可以得到区域的纹理特征向量:

$$\vec{T} = (\mu_{11}, \sigma_{11}, \dots, \mu_{1n_0}, \sigma_{1n_0}, \dots, \mu_{n_F 1}, \sigma_{n_F 1}, \dots, \mu_{n_F n_0}, \sigma_{n_F n_0}) \quad (9)$$

### 4 颜色、纹理特征组合

自然界中的图像千差万别,有时图像灰度起伏变化剧烈,呈现出很强的纹理特性。但是颜色却比较相近,例如草地和树木;有时图像由若干个纹理相似但颜色有差异的区域构成,例如天空和道路。颜色和纹理都只是描述了图像的一部分特征。有时一幅图像用颜色特征就能区别其中不同的景物,而用纹理特征却区别不出来,相反的情况也是经常发生的。单纯地只用一种特征来表征图像的信息,并不能全面地描述图像所包含的内容。目前已提出的方法主要侧重于单个特征的描述。即使采用了多特征,也只是把几个特征向量简单地放在一起,因此对景物的描述还是不全面的。所以,本文将颜色和纹理特征相结合,这样不仅能够表达更多的图像信息,而且

可以从不同方面更加详细地对图像进行描述,以得到更好的分类结果。

把颜色和纹理特征组合得到一个 60 维的特征向量如下:

$$\vec{f} = [R_0, \dots, R_3, G_0, \dots, G_3, B_0, \dots, B_3, \mu_{11}, \sigma_{11}, \dots, \mu_{1n_0}, \sigma_{1n_0}, \dots, \mu_{n_F1}, \sigma_{n_F1}, \mu_{n_Fn_0}, \sigma_{n_Fn_0}] \quad (10)$$

其中,前 12 维是 RGB 颜色直方图特征向量,后 48 维为 Gabor 纹理特征向量。

由于颜色特征和纹理特征的物理意义不同,不直接具有可比性,因此需对颜色和纹理的向量特征进行外部归一化<sup>[19]</sup>。归一化的公式如下<sup>[20]</sup>:

$$f'_i = \frac{f_i - f_{\min}}{f_{\max} - f_{\min}}, i = 1, 2, \dots, 60 \quad (11)$$

其中,  $f_i$  是特征向量  $\vec{f}$  的第  $i$  维特征值,  $f_{\max}$  和  $f_{\min}$  分别是特征向量  $\vec{f}$  中的最大和最小值,归一化后得到的  $f'_i$  值的取值范围是  $[0, 1]$ 。

## 5 实验结果及分析

为了对以上提出的几种特征提取方法进行评价,下面基于一对多的 SVM 分类方法<sup>[21]</sup>分别用基于单一的颜色特征、单一的纹理特征以及颜色与纹理的组合特征对景物进行识别,并对识别结果进行比较。其中训练及测试图像均取自美国加州理工学院的 Pasadena Houses2000 数据库,该库共有 157 幅典型的室外场景图像,其中 100 幅图像用来训练,另外 57 幅图像用来测试。手工将各类景物的样本从训练图像中分割出来,得到天空、道路、树木、草地、房子 5 类景物的样本图像。每类景物有 100 个样本,样本图像大小为  $64 \times 64$ 。并选择径向基核函数(惩罚因子  $C$  为 10、径向基核函数的宽度  $\sigma$  为 1)作为 SVM 模型的参数对样本进行训练<sup>[21]</sup>。采用不同的特征得到的识别结果如图 1 所示。

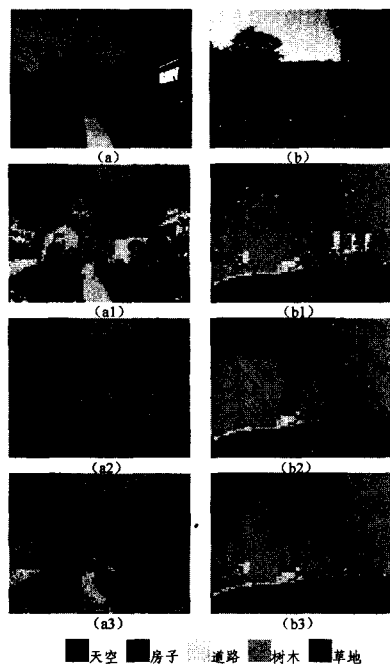


图 1 单特征与多特征的识别结果比较 ( $C=10, \sigma=1$ )

通过实验分析表明,对于每类目标颜色差异较大的场景

图像,仅使用颜色特征就可以得到很好的识别效果。但是颜色较相近的目标,例如树木和草地,仅使用颜色特征就无法区别。仅使用纹理特征也一样,像天空和道路的纹理特征较相近,就无法通过纹理特征进行分割识别。而使用两种特征的组合就可以明显提高识别的精度。

**结束语** 对于室外场景图像的景物识别问题,特征提取是一个非常重要的因素,而颜色与纹理特征是应用最广泛的两种特征。首先选择基于 RGB 空间的颜色直方图方法以及基于多通道 Gabor 变换的方法分别提取图像的颜色和纹理特征,然后为了更全面充分地描述图像中的景物,对颜色和纹理特征组成的特征向量通过外部归一化之后得到组合特征。再采用一对多的 SVM 分类方法对美国加州理工学院的 Pasadena Houses2000 数据库中的图像进行测试,比较了提出的几种特征对识别结果的影响。实验结果表明,颜色与纹理的组合特征在景物识别中更有效。

## 参考文献

- [1] Kumar S, Loui A C, Hebert M. An Observation - constrained Generative Approach for Probabilistic Classification of Image Regions[J]. Image and Vision Computing, 2003, 21(1): 87-97
- [2] Bosch A, Munoz X, Freixenet J. Segmentation and Description of Natural Outdoor Scenes[J]. Image and Vision Computing, 2007, 25(5): 727-740
- [3] 边肇祺, 张学工. 模式识别[M]. 北京: 清华大学出版社, 2000
- [4] Vapnik V. Statistical Learning Theory[M]. New York: Wiley, 1998
- [5] Osuna E, Freund R, Girosi F. Training Support Vector Machines; An Application to Face Detection[A] // Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition[C]. New York: IEEE, 1997: 130-136
- [6] Joachims T. Text Categorization with Support Vector Machines; Learning with Many Relevant Features[A] // Proceedings of the 10th European Conference on Machine Learning [C]. Heidelberg: Springer Berlin, 1998: 137-142
- [7] Dumais S, Platt J, Heckerman D, et al. Inductive Learning Algorithms and Representations for Text Categorization[A] // Proceedings of the 7th International Conference on Information and Knowledge Management[C]. New York: ACM, 1998: 148-155
- [8] Campbell N W, Mackeown W, Thomas B T. Interpretation Image Databases by Region Classification[J]. Pattern Recognition, 1997, 30(4): 555-563
- [9] Tenenbaum J N. An Interactive Facility for Scene Analysis Research[R]. Report 87. Artificial Intelligence Center, Stanford Research Institute, 1974
- [10] Celenk M. A Color Clustering Technique for Image Segmentation[J]. Computer Vision, Graphics and Image Processing, 1990, 52(2): 145-170
- [11] 伯晓晨, 刘建平. 基于颜色直方图的图像检索[J]. 中国图象图形学报, 1999, 4(3): 33-37
- [12] Tuceyran M, Jain A K. Texture Analysis, in Handbook of Pattern Recognition and Computer Vision[M]. 2nd edition. Singapore: World Scientific Publishing Co., 1998: 207-248
- [13] 高燕. 基于小波多尺度分析的图像检索技术研究[D]. 扬州: 扬州大学, 2006

(下转第 283 页)

积极作用。图4的实验结果也验证了这一点。

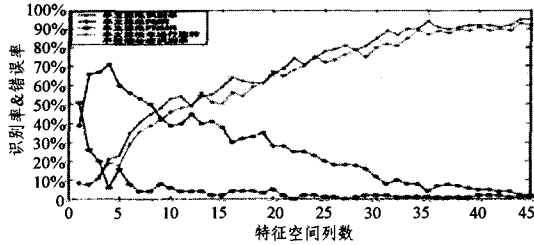


图4 特征空间大小与识别率、FMR(错误拒绝)、FNMR(错误接受)之间的对应图

图5是对 $r=45$ 时类内样本比较和类间样本比较的Hamming距离的分布直方图。从图中可以看出,类内对比结果数据比较分散,且Hamming距离较小;类间对比结果数据比较集中,且Hamming距离较大。说明本文算法可以有效地区分不同类别样本。实验验证,取 $k=5$ 和阈值为0.34,可以得到较快的识别速度和较好的识别效果,此时识别正确率可达到95%以上。

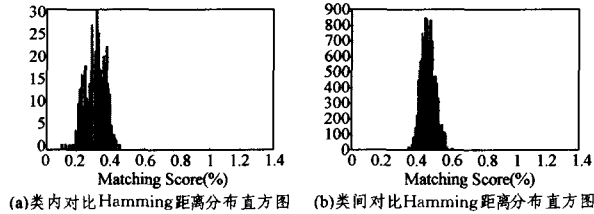


图5  $r=45$ 时样本比较 Hamming 距离分布直方图

此外,对本文提出的算法与2D-PCA和文献[3]中提出的基于小波系数直接量化获得特征向量的识别方法(简记Lim算法)等进行了对比,实验结果如表1和表2所列。

表1 本文算法和Lim、2D-PCA算法识别能力比较

算法	识别正确率(%)	错误率(%)
2D-PCA	89.6	11.5
Lim	93.4	6.5
Proposed	95.1	4.9

表2 本文算法和PCA、2D-PCA算法复杂度比较

算法	训练矩阵维数	协方差矩阵维数	输入空间个数
PCA	$48 \times 360$	$17280 \times 17280$	1
2D-PCA	$48 \times 360$	$360 \times 360$	1
Proposed	$6 \times 45$	$45 \times 45$	2

从表1和表2可以看出,本文提出的多频带2D-PCA算法,降低了2D-PCA输入空间的维数,同时使得2D-PCA的输入空间刻画识别样本特征的能力得到增强。而且由于选取不同子带的小波系数作为2D-PCA的输入空间,因而增强了算法的鲁棒性。实验表明,无论对比小波系数直接量化方法、PCA方法或者2D-PCA方法,该算法在识别效果和识别时间上均有一定程度的改进。

**结束语** 为了获得更好的虹膜识别效果和更快的识别速度,本文提出了一种多频带2D-PCA的虹膜识别算法。该算法在保持识别效果和识别速度的前提下,通过设置多个归一化起始角度,使算法具有虹膜图像旋转不变的特性。同时采用小波的方向子带作为输入空间,提高了算法对形变和噪音的鲁棒性。实验结果验证了所提算法的有效性。

### 参考文献

- [1] Turk M, Pentland A. Eigenfaces for Recognition[J]. Cognitive Neuroscience J, 1991, 3(1): 71-86
- [2] Yang J, Zhang D, Frangi A F, et al. Tow - dimensional PCA : A New Approach to Appearance -based Face Representation and Recognition[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2004, 26(1): 131-137
- [3] Lim S, Lee K, Bveon O, et al. Efficient iris recognition through improvement of feature vector and classifier [J]. ETRI Journal, 2001, 23(2): 61-70
- [4] 明星,刘元宁,等. 基于平移不变预处理的小波变换的虹膜识别算法[J]. 计算机研究与发展, 2006, 43(7): 1186-1193
- [5] 李弼程,彭天强,等. 智能图像处理技术[M]. 北京: 电子工业出版社, 2004
- [6] Daugman J. High confidence visual recognition of persons by a test of statistical independence[J]. IEEE Trans, PAMI, 1993, 15(11): 1148-1161
- [7] 边肇祺,张学工,等. 模式识别[M]. 北京: 清华大学出版社, 2000
- [8] CASIA Iris Image Database[OL]. <http://www.sinobiometrics.com>, 2005
- [9] Mansfield A J, Wayan J L. Best practices in testing and reporting performance of biometric devices [R]. CMSC14 /02. Middlesex, U. K. : National Physical Laboratory, 2002
- [10] Mallat S G. A theory for multiresolution signal decomposition; The wavelet representation [J]. IEEE Trans, PAMI, 1998, 20(7): 674-693
- [11] ... (omitted for brevity) ...
- [12] ... (omitted for brevity) ...
- [13] ... (omitted for brevity) ...
- [14] Daugman J. High Confidence Visual Recognition of Persons by a Test of Statistical Independence[J]. Transaction on Pattern Analysis and Machine Intelligence, 1993, 15(11): 1148-1161
- [15] Jain A, Ratha N, Lakshmanan S. Object Detection Using Gabor Filters[J]. Pattern Recognition, 1997, 30(2): 295-309
- [16] Turner M R. Discrimination by Gabor Functions[J]. Biol. Cybern, 1986, 55(2): 71-82
- [17] Daugman J. Uncertainty Relation for Resolution in Space, Spatial frequency and Orientation Optimized by Two-dimensional Visual Cortical filters[J]. Journal of the Optical Society of America, 1985, A. 2: 1160-1169
- [18] Dunn D F, Higgins W E. Optimal Gabor Filters for Texture Segmentation[J]. IEEE Transaction on Image Processing, 1995, 4(7): 947-964
- [19] 刘中伟,章毓晋. 综合利用颜色和纹理特征的图像检索[J]. 通信学报, 1999, 20(5): 36-40
- [20] 张瑜慧. 基于SVM的语义图像检索技术的研究和实现[D]. 扬州: 扬州大学, 2007
- [21] 杨丽华. 复杂室外环境下景物识别算法研究[D]. 西安: 西北工业大学, 2008

(上接第276页)

- [14] Daugman J. High Confidence Visual Recognition of Persons by a Test of Statistical Independence[J]. Transaction on Pattern Analysis and Machine Intelligence, 1993, 15(11): 1148-1161
- [15] Jain A, Ratha N, Lakshmanan S. Object Detection Using Gabor Filters[J]. Pattern Recognition, 1997, 30(2): 295-309
- [16] Turner M R. Discrimination by Gabor Functions[J]. Biol. Cybern, 1986, 55(2): 71-82
- [17] Daugman J. Uncertainty Relation for Resolution in Space, Spatial frequency and Orientation Optimized by Two-dimensional Visual Cortical filters[J]. Journal of the Optical Society of A-