

一种图像敏感对象的识别方法

张永梅¹, 吴攀¹, 李炜²

(1. 北方工业大学 信息工程学院, 北京 100144; 2. 北京航空航天大学 计算机学院, 北京 100191)

摘要: 针对现有的图像识别方法识别某一指定对象, 提出了一种图像敏感对象的识别方法, 该方法改进了现有目标识别的局限性; 首先通过计算得到图像中对象的敏感度值, 然后对图像对象进行纹理、色彩、形状和显著特征提取, 最后结合敏感度值识别出图像中敏感度大的对象, 并与经典的敏感度检测方法比较; 实验结果表明该方法敏感信息丢失少, 且能精确得到图像对象敏感度值, 通过敏感度的不同, 识别出人眼感兴趣的对象信息。

关键词: 敏感度计算; 特征提取; 图像识别; 感兴趣对象

A Recognition Method of Image Sensitive Objects

Zhang Yongmei¹, Wu Pan¹, Li Wei²

(1. School of Information Engineering, North China University of Technology, Beijing 100144, China

2. School of Computer, Beihang University, Beijing 100191, China)

Abstract: For the conventional image recognition methods, this paper proposes a recognition method of image sensitive objects which improves the limitations of present approaches. This method gives each object in the image a precise sensitivity value. Then it extracts the texture, color, shape and salient features of image object. Finally, according to the sensitivity values, it identifies the objects with large sensitivity value in the image and compares with the classical methods of sensitivity detection. Experimental results show the proposed method loses less sensitive information and can obtain more accurate image object sensitivity value and identify the objects of interest of the human eyes by different sensitivity values.

Keywords: sensitivity value calculating; feature extraction; image recognition; object of interest

0 引言

目前图像识别方法大致可概括为 3 种: 统计(或决策理论)法, 结构(或句法)方法和神经网络法。统计法是一种分类误差最小的方法, 它以数学上的决策理论为基础, 建立统计学识别模型, 但是物体的结构特征为主要特征时, 便会很难识别。句法方法则是用符号来描述图像特征, 采用分层的方法将复杂图像分解为单层或多层的简单子图像。神经网络方法是指用神经网络的算法对图像进行识别, 该方法侧重于模拟和实现人的认知过程。

但是在观察一副图像时, 人们的视觉注意力往往是分散在多个感兴趣对象上, 目前的图像识别方法只针对某一指定对象, 为此, 本文提出了一种敏感度对象的识别方法, 通过这种显著性特征甄选出人们感兴趣的图像对象。该方法为图像中的每一个对象计算敏感度值, 通过敏感度值的大小反映图像对象的显著性信息。

1 图像对象的敏感度计算方法

图像对象的敏感度计算方法的主要思想是: 首先结合颜色特征对图像进行分割, 大致确定图像各对象的区域^[1]。图像对象区域确定后, 根据图像场景中各对象之间的上下文关系来计算对象在当前图像的敏感度; 最后结合预先设置的初始图像对

象敏感度对图像对象进行甄选和预判断, 这里的初始敏感度由图像的应用场景和个人的喜好来设置。

图像中对象的敏感度取决于图像的应用场景、使用人的个人喜好以及图像场景中各对象之间的上下文关系 3 个因素, 其中, 前 2 个因素相对固定, 因此, 可以根据人们浏览图像的一般常识, 采用预先定义的方式给出各对象的初始敏感度。表 1 给出了本文图像对象初始敏感度的定义。

表 1 图像对象敏感程度定义表

种类	敏感度(范围:0~1.0)	种类	敏感度(范围:0~1.0)
汽车	1.0	建筑物	0.8
天空	0.3	草地	0.5

根据一般的经验, 我们知道: 图像对象的敏感度不仅与该对象的种类有关, 同时还受图像场景中对象之间的上下文关系影响, 例如: (1) 对象的面积, 通常在图像中占据较大面积的对象为图像的背景区域, 具有较低的敏感程度。(2) 距离图像中心的距离, 距图像中心越远, 则敏感程度越低。本文在实验和分析的基础上找出影响对象敏感度的上下文因素, 并据此给出图像上下文敏感度的计算公式, 其基本形式可用公式(1)表示:

$$SD_T = W_1 * A_T + W_2 * L_T$$

$$W_1 + W_2 = 1 \quad (1)$$

式中, SD_T 为某个图像对象的上下文敏感度, W_1 、 W_2 为影响因子的权重, A_T 表示该对象的面积与所有感兴趣对象面积和之比, L_T 表示该对象与图像中心的距离所处的等级。

2 场景分类器的构建

为了能较好地识别出图像中感兴趣对象, 本文采用了构建

收稿日期: 2014-03-10; 修回日期: 2014-04-11。

基金项目: 国家自然科学基金项目(61371143); 国家科技支撑计划项目(2012BAH04F00); 北京市自然科学基金项目(4132026)。

作者简介: 张永梅(1967-), 女, 山西太原人, 教授, 博士, 主要从事图像处理、智能识别方向的研究。

场景分类器来识别图像对象。具体方案是：首先获得大量图像对象的样本作为训练图像，提取对象的视觉特征，进行多特征融合构造出图像对象的分类特征，再采用机器学习的方法，学习出准确率高的图像对象分类器。

2.1 特征提取

本文考虑的图像对象主要是汽车、建筑物、天空、草地等。不同对象的图像灰度值变化不一样，有的较剧烈，纹理能量较大，有的比较平缓，纹理能量就相对较小，建筑物的纹理相对比较杂乱，熵值较大；而纹理能量值容易突出大片均匀的天空或者草地区域^[2]。颜色是彩色图像最重要的内容之一，可以借助颜色特征完成部分对象识别和分类，比如：大面积的蓝色区域可能是天空或者水域，大面积的绿色区域可能是草地等。汽车的车牌特征，建筑物的平行线特征是区别于其他对象的显著特征。综上所述，为了正确的识别这些对象，本文分别提取了纹理特征、颜色特征、形状特征和各对象的显著特征。

本文选取大量的图像对象训练样本，计算其纹理能量 E 和其它 3 个纹理描述符 W_E, W_C, W_{CO} ，具体定义如下所示：

$$E = \sum_{i=1}^k \sum_{j=1}^k (G(i, j))^2 \tag{2}$$

$$W_E = \sum_{i=1}^k \sum_{j=1}^k G(i, j) \log G(i, j) \tag{3}$$

$$W_C = \sum_{n=0}^{k-1} n^2 \sum_{|i-j|=n} G(i, j) \tag{4}$$

$$W_{CO} = \sum_{i=1}^k \sum_{j=1}^k \frac{(i, j)G(i, j) - u_i u_j}{s_i s_j} \tag{5}$$

式中， $G(i, j)$ 表示灰度共生矩阵，

$$u_i = \sum_{i=1}^k \sum_{j=1}^k iG(i, j),$$

$$u_j = \sum_{i=1}^k \sum_{j=1}^k jG(i, j),$$

$$S_i^2 = \sum_{i=1}^k \sum_{j=1}^k G(i, j)(i - u_i)^2.$$

能量反映了图像灰度分布均匀程度和纹理粗细程度，如果共生矩阵的所有值均相等，则 E 有较大值； W_E 是图像所具有的信息量的度量，若灰度共生矩阵值分布均匀，也即图像近于随机或噪声很大，熵会有较大值； W_C 反映图像的清晰度和纹理沟纹深浅的程度，灰度共生矩阵中远离对角线的元素越大，对比度越大； W_{CO} 反映了图像纹理的一致性，如果图像中有水平方向纹理，则水平方向矩阵的自相关值大于其余矩阵的自相关值。

从图像中提取颜色特征时，很多算法都先要对图像进行量化处理，量化处理容易导致误检，并且产生的图像特征维数较高，不利于检索。因此本文提取颜色矩作为颜色特征。颜色矩是一种简单有效的颜色特征表示方法，本文采用一阶矩（均值，Mean）和二阶矩（方差，Variance）表达图像的颜色分布。以下是两个颜色矩的数学定义：

$$\mu_i = \frac{1}{N} \sum_{j=1}^N p_{i,j} \tag{6}$$

$$\delta_i = \left(\frac{1}{N} \sum_{j=1}^N (p_{i,j} - \mu_i)^2 \right)^{\frac{1}{2}} \tag{7}$$

式中， $p(i, j)$ 表示色彩图像第 i 个颜色通道分量中灰度为 j 的像素出现的概率， N 表示图像中的像素个数。图像的 3 个分量

N, U, V 图像的前两个颜色矩组成一个 6 维直方图向量，即图像的颜色特征表示如式（8）：

$$F_{color} = [\mu_Y, \delta_Y, \mu_U, \delta_U, \mu_V, \delta_V] \tag{8}$$

矩特征是图像识别的一种统计特征的提取方法，可以将一些基本的二维形状推导出来。本文利用了 Hu 不变矩来描述对象的形状特征，它们在连续图像条件下可以保持旋转、缩放和平移的不变性。

除了对训练对象分别提取了 3 种基础的公共特征之外，本文在构建图像对象分类器还针对人眼敏感的对象进行对象的显著特征提取。例如：汽车的车牌特征、建筑物的平行线特征等。

最后，将对象的纹理、颜色、形状和显著特征组合为一个特征向量，放入到特征向量库中。

2.2 对象自动识别

对所有图像对象的训练样本集进行特征提取，得到图像对象的特征库。待测图像输入后，对待测样本进行特征提取，与特征库的特征进行匹配，从而完成对象的自动分类和识别。常见的分类方法有：神经网络、遗传算法、K 最近邻算法（KNN）、支持向量机（SVM）等。神经网络分类需要大量的参数，如网络拓扑结构、权值和阈值的初始值；遗传算法的编程实现比较复杂，首先需要对问题进行编码，找到最优解之后还需要对问题进行解码；SVM 对缺失数据敏感，对非线性问题没有通用解决方案；相比较而言，KNN 算法简单有效，重新训练的代价较小，由于本论文需要处理的样本容量较大，KNN 算法适用于样本容量比较大的类域自动分类，故本文采用 KNN 分类法。

KNN 分类的基本思路是：给出一个待识别样本集，系统在训练集中找到最近的 K 个近邻，观察这 K 个近邻多数属于哪一类，就将待识别的样本归为此类。本文在实验中采用夹角余弦来表示两个相似样本之间的距离，即两个样本之间的相似性，计算如式（9）所示：

$$C_{ij} = \frac{\sum_{k=1}^n X_{ik} * X_{jk}}{\sqrt{(\sum_{k=1}^n X_{ik}^2)(\sum_{k=1}^n X_{jk}^2)}} \tag{9}$$

式中， $X_{ik}, X_{jk}, (k = 1, 2, \dots, n)$ 为两个样本向量。两个样本之间的夹角，如果夹角越小，说明两个样本形似性越大。

为了检测整幅图像，正确识别对象，本文做了如下的处理：在图像中移动搜索窗口，检测每一个位置来确定可能的目标；为了搜索不同大小的目标物体，分类器被设计为可以改变尺寸。图像识别的总体流程如图 1 所示。

3 实验结果与分析

用 100 幅图像对本文提出的图像敏感对象的识别方法进行测试。这些测试图像包括 5 类对象：汽车、树、建筑物、草地等。鉴于篇幅有限，下面只是汽车敏感对象的识别为例以示说明。

3.1 敏感度的计算

经过图像预处理和分割，结合颜色特征可以得到图像对象的上下文关系。为了计算方便，本文将图像分割后得到的连通域近似认为是在图像中各对象的面积。因此，区域的中心即为对象的质心。表 2 给出了图像和对象的大小、中心信息。

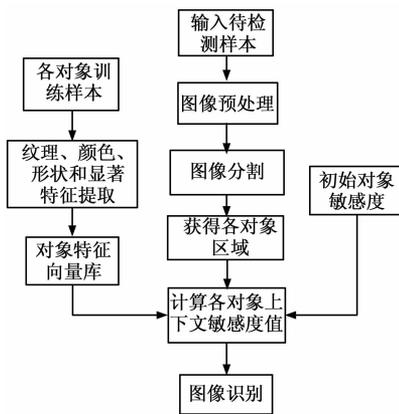


图 1 图像识别整体流程

表 2 图像和对象的信息表

	大小(像素个数)	中心坐标
图像	475 * 346	(287, 173)
汽车	230 * 225	(239, 183)

根据表 2 和上下文敏感度计算公式 (1) 可知, A_T 的取值为 1, 对象质心距离图像中心的距离为 49.03, 而图像中心距图像最边缘的距离为 335.11, 故对象与图像中心的距离所处的等级 L_T 为 0.8. 设置权重 W_1, W_2 都为 0.5, 最后可求得对象在该图像中的上下文敏感度为: $SD_T = 0.9$.

从实验结果可以看出: 本文提出的对象敏感度计算方法得到的结果与人眼实际观察的感兴趣对象显著程度很相似, 本文方法将人眼的观察的兴趣和重点量化的显示出来, 为后续图像识别工作提供了精确参考。

将本文的敏感度计算方法和其他经典的视觉显著度检测方法相比, 可以发现本文的方法能准确计算出各对象的敏感度值, 通过敏感度值来确定显著程度; 而其他方法只是粗略给出显著的对象, 对比结果如图 2 所示。

对于复杂背景或者多对象的图像而言, 敏感对象信息丢失度是显著度计算的另一个检测指标。显著度检测之后的结果与原图的重要信息匹配得越多, 认为显著度检测的结果越好。本文将原图作为比较标准, 统计 4 种方法得到的显著性结果与原图信息的差异, 表 3 从另一个角度验证了本文提出的方法的准确性。

表 3 显著性检测方法敏感信息丢失对比

显著性检测方法	敏感信息丢失度(%)
Itti-Koch 方法	40.4
光谱剩余假说方法	28.1
Hu-Rajan-Chia 方法	16.8
敏感度计算方法	9.5

Itti^[3]等人用高斯模糊连续采样对原始输入图像的亮度、颜色、方向 3 种特性进行连续采样, 根据高斯差计算得到各特征维的醒目图, 融合得到显著图。Itti 方法对一般的图像比较准确, 但不适用于显著物体较大、颜色比较均匀以及背景中有反常点的图像。Hou 和 Zhang^[4]等提取输入图像的 log 光谱图像在光谱领域的光谱剩余, 得到空间结构的显著性图, 但是它不适用于显著物体形状复杂的图像。Hu^[5]等人利用局部上下

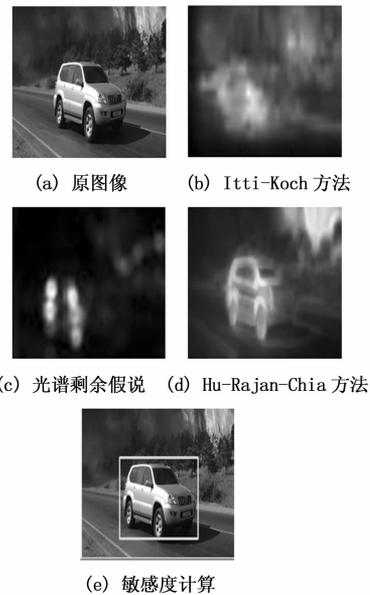


图 2 各显著性测试方法对比

文信息区分注意区域和非注意区域, 计算各个图像块的抑制因子和各个特征与其邻近区域的变化对比图, 结合亮度、颜色和抑制因子生成显著图。该方法使用多尺度纹理特征提取, 能很好的识别混乱背景中的小尺度对象, 但是对大尺度对象的识别能力不佳。而本文方法则是通过计算各对象的敏感度值, 通过敏感度值来确认显著程度, 因此它能更好的保留图像中的敏感对象信息。

3.2 敏感对象的特征提取

对敏感对象的训练图片提取纹理特征: 能量 E 和 W_E, W_C, W_{CO} ; 提取颜色一阶矩和二阶矩特征: $\mu_V, \delta_V, \mu_U, \delta_U, \mu_V, \delta_V$; 提取形状特征: 7 个 Hu 不变矩, 表 4 分别列出了汽车的纹理特征, 颜色特征和 Hu 不变特征。

(1) 消去水平轮廓。由于经过预处理后的图像拥有丰富的水平边缘信息, 而车牌区域拥有丰富的垂直边缘信息, 所以可以通过消去水平轮廓信息, 留下垂直轮廓信息^[6]。这样在车牌区域留下丰富的边缘信息, 而其他区域的信息会减少, 效果如图 3 所示。

表 4 汽车特征统计表

	汽车对象的特征
纹理	0.006087, 0.150090, 1.632622, 0.001988
颜色	0.5192, 0.2231, 0.3496, 0.2520, 0.1535, 0.1933
Hu 不变矩	2.44289, 6.151511, 9.503062, 9.904652, 19.732114, 19.789745

(2) 数学形态学处理。数学形态学主要采用了邻域运算形式, 它的基本思想是用具有一定形态的结构元素去度量和提取图像中的对应形状以达到对图像分析和识别的目的, 其基本运算有: 腐蚀 (Erosion)、膨胀 (Dilation)、开运算 (Open) 和闭运算 (Close)。以图 3 (b) 为例, 对其进行数学形态学处理后的结果如图 3 (c) 所示。

(3) 车牌粗定位: 经过上述形态学处理后, 从在图像中可以发现存在一些矩形区域, 这些矩形区域就是车牌的候选区域,

(下转第 2188 页)

应用水平，能够显著提升航空修理行业的生产效率和修理质量；同时该技术在测试测控及相关行业领域内的普遍推广与应用，也是具有实际意义的。

参考文献：

[1] Single-Channel, 1024-Position, 1% R-Tolerance Digital Potentiometer [EB]. Analog Devices Inc. 2011.

[2] PCI9052 Data Book [EB]. PLX Technology Inc. 2008.

[3] MAX II Device Handbook [EB]. Altera Corporation. 2009.

[4] PCI Local Bus Specification Revision 2. 3 [EB]. PCI Special Inter-

est Group. 2001.

[5] 宋宇峰, 郝文化. LabWindows/CVI 逐步深入与开发实例 [M]. 北京: 机械工业出版社, 2003.

[6] 钱能. C++ 程序设计教程 [M]. 北京: 清华大学出版社, 1994.

[7] 李贵山, 陈金鹏. PCI 局部总线及其应用 [M]. 西安: 西安电子科技大学出版社, 2003.

[8] 聂鑫, 田建生, 梁远灯. 基于 FPGA 的 PCI 总线仲裁器设计 [J]. 计算机测量与控制, 2005, 13 (8): 817-820.

(上接第 2169 页)

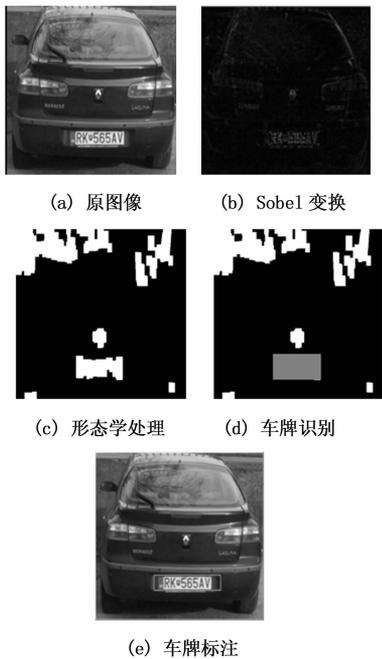


图 3 车牌识别结果

分别求得每个车牌候选区域的高、宽和坐标，依据车牌在原图像中的几何特征，定位车牌^[7]。考虑车牌的以下 4 种几何特征，从上述区域中筛选出包含车牌的区域。

- ①宽：车牌区域的宽度像素应该大于 90。
- ②高：车牌区域的高度像素应该大于 30。
- ③宽高比：车牌区域的宽高比例应大于 2。
- ④位置：车牌区域在车辆图像中出现的位置一般在某一范围内。

经过形态学处理后的图像中，车牌区域并不一定是完全是白色区域的矩形，本文规定只要白色面积大于矩形面积的 90%，就认为是车牌区域，实验结果如图 3 (d) 所示。

3.3 敏感对象识别

对所有的样本，分别提取纹理、颜色、形状和车牌特征，将得到的特征向量集放到后续识别的特征库中训练。训练完以后，就可以应用于输入图像中的敏感对象的检测识别。为了在图像中检测识别未知大小的敏感对象，本文的扫描程序采用不同比例大小的搜索窗口对图片进行多次扫描，最终优化检测结果可得到识别结果，图 4 给出了识别结果。

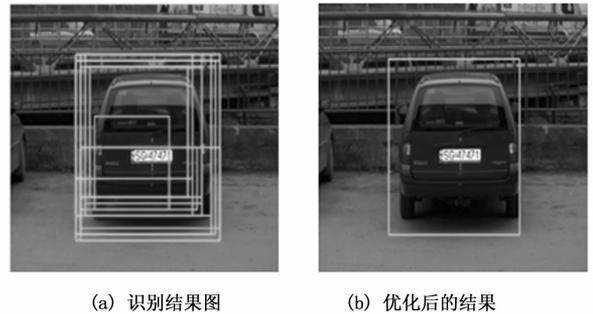


图 4 识别结果图

4 结论

一般方法只能识别指定的图像对象，而本文与其他方法不同，本文提出了一种图像敏感度对象的识别方法。该方法可以准确给出一幅图像中每个对象的敏感度值，结合敏感度值的大小，识别出图像中敏感度高的对象，这种方法更能接近于人类的视觉注意机制。

参考文献：

[1] Rafale C Gonzalez, Richard E Woods. 数字图像处理 [M]. 阮宇智, 等译. 北京: 电子工业出版社, 2004.

[2] 徐贵力, 毛罕平, 李萍萍. 彩色图像颜色和纹理特征提取的应用算法 [J]. 计算机工程, 2002, 28 (6): 25-27.

[3] Itti L, Koch C, Niebur E. A Model of Saliency Based Visual Attention for Rapid Scene Analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20 (11): 1254-1259.

[4] Hou X D, Zhang L Q. Saliency Detection: a Spectral Residual Approach [A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C], Minneapolis, USA: IEEE Computer Society, 2007.

[5] Hu Y, Rajan D, Chia L T. Adaptive Local Context Suppression of Multiple Cues for Salient Visual Attention Detection [A]. Proceedings of IEEE International Conference on Multimedia and Expo [C], Amsterdam, the Netherlands: IEEE Computer Society, 2005.

[6] 董燕, 朱永胜, 刘聪. 图像融合技术在车牌识别中的应用 [J]. 计算机测量与控制, 2013, 21 (3): 791-793.

[7] Clady X, Collange F, Jurie F. Cars detection and tracking with a vision sensor [A]. IEEE Proceedings of Intelligent Vehicles Symposium [C], 2003: 593-598.