

典型室内场景显著性稀疏识别

严 晗¹, 刘 佶², 龚建荣²

(1. 南京邮电大学通信与信息工程学院, 江苏 南京 210003)

(2. 南京邮电大学教育部工程研究中心, 江苏 南京 210003)

[摘要] 随着信息技术和智能机器人的发展与普及, 场景识别作为重要的研究内容, 已成为计算机视觉和模式识别领域的重要研究问题. 解决室内场景分类精度低的问题, 将有助于室内场景分类在场景图片检索、视频检索及机器人等领域中的应用. 针对常规场景识别方法在室内环境中性能显著下降的问题, 提出一种基于显著性检测的稀疏表示室内场景识别方法. 该方法利用显著性区域检测算法提取出场景图像中人眼感兴趣的区域, 并与稀疏表示结合进行场景识别. 实验结果表明, 将本方法应用在典型家庭室内场景(如卧室、厨房、衣帽间等), 在识别正确率方面有一定的优势.

[关键词] 场景识别, 室内场景分类, 显著性区域检测, 稀疏表示

[中图分类号] TP391 **[文献标志码]** A **[文章编号]** 1001-4616(2017)01-0079-07

Significant Spare Representation of Typical Indoor Scene Recognition

Yan Han¹, Liu Jixin², Gong Jianrong²

(1. College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

(2. Engineering Research Center of Education, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: With the development and popularization of information technology and intelligent robots, scene recognition as an important research content has become an important research in the field of computer vision and pattern recognition problem. Solving the problem of the low classification accuracy for indoor scene will help the indoor scene classification in some areas of application: the image retrieval, video retrieval of the scene and the robot. Conventional scene recognition methods have poor performance in indoor situations. For this reason, a sparse representation indoor scene recognition method is presented, which based on significant detection. This method is using significant recognition detection to extract the scene in the image area which we are interested in, and combined with sparse representation to scene classification recognition. Experimental results show that this method can be applied to a typical family indoor scenarios (e. g., bedroom, kitchen, closet, etc.) and have certain advantages in terms of recognition accuracy.

Key words: scene recognition, indoor scene classification, salient region detection, sparse representation

由于多媒体技术和互联网技术的快速发展, 每天都会产生大量的数字图像. 如何利用计算机自动将图像按照人类理解的方式分类到不同的类别, 从而快速有效地获取、管理和分类数量巨大的图像成为一个重要问题, 场景分类就由此产生. 针对场景分类难的问题, 不同阶段的研究提出不同的方法和模型. 早期, 场景图像分类一般都是基于整体模型^[1-2]开展, 利用颜色、纹理、形状等特征进行识别, 该类方法在训练集之外的泛化能力较差. 近年来, 诸如 SIFT^[3,19] (Scale-invariant Feature Transform, 尺度不变特征变换)、SURF^[4] (Speeded-Up Robust Features, 加速鲁棒特征)、HOG^[5] (Histogram of Oriented Gradient, 梯度方向直方图)、SDA^[6] (Subclass Discriminant Analysis, 子类判别分析)等算子有着广泛应用. 比较流行的分类方法是利用各种算子提取环境特征, 应用最广的是视觉词袋模型 (Bag of visual Words, BoW^[7]), 这一思路在图像分类的应用中取得了令人鼓舞的结果, 受到了研究者的极大关注. 但是, 由于忽略了局部图像块的位置

收稿日期: 2016-08-20.

基金项目: 国家自然科学基金青年基金 (61401220)、江苏省自然科学基金青年基金 (BK20140884)、江苏省高校自然科学研究面上项目 (14KJB510022).

通讯联系人: 严晗, 硕士, 研究方向: 稀疏表示、图像处理. E-mail: 729838449@qq.com

信息,该方法属于一种无序的特征表示,即缺少位置信息的全局特征表示.为解决这个问题,Lazebnik 等人提出以一种基于空间金字塔匹配(Spatial Pyramid Matching,SPM^[8,18])的方法来改进传统的视觉词袋模型.但是 SPM 策略存在着较大的量化误差,进而导致比较严重的信息损失.为了解决这个问题,Yang J^[9]等人首次提出使用稀疏编码(Sparse Coding)策略来学习视觉词典,然后用稀疏编码方法对整幅图像的关键点进行编码,最后用基于最大池(Max Pooling)的方法表示图像特征,他们称这种方法为稀疏的空间金字塔匹配(Sparse coding Spatial Pyramid Matching,ScSPM).此后,又有一系列工作对 ScSPM 中的不足做了改进,如 Wang^[10]等人在稀疏编码中加入了位置信息的约束,这使得编码效率和性能得到了改善;Boureau^[11]等人在视觉学习中引入了类别信息,提出了监督的系数词典学习方法.

室内场景分类是场景分类的一个研究领域,解决室内场景分类精度低的问题,将有助于室内场景分类在场景图片检索、视频检索及机器人等领域的应用.但是,现在所有的算法只是对室外场景处理较好,对于室内场景的识别还存在很多的不足.这是因为相较于户外场景,室内环境通常缺少显著的局部或全局视觉特征.本文针对常规场景识别方法在室内环境中性能显著下降的问题,提出一种基于显著性检测的稀疏表示室内场景识别方法.该方法利用显著性区域检测提取出场景图像中人眼感兴趣的区域,并与稀疏表示结合进行场景识别.实验结果表明,本方法能得到较高的识别正确率.

1 基于显著性检测的稀疏表示算法框架

场景识别技术的典型框架是特征表示加分类器,其中的特征表示算子都是人为设计的,需要有特征提取的预处理过程.

本文的算法框架如图 1 所示.从框架流程图中可以看出,该室内场景识别算法结合了显著性检测和稀疏表示算法,场景图片首先通过显著性检测得到图像的显著性图,通过该显著性图得到不同的分割图像,这些分割图像作为最终稀疏表示算法的训练输入,最终训练出类别字典进行场景识别.显著性检测不仅分割出图像中最主要的目标,减少了背景噪声的干扰,而且提高了稀疏表示算法的运算速度和可靠性.

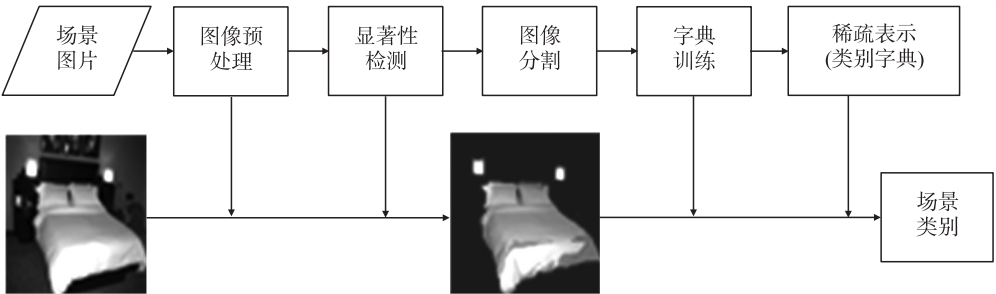


图 1 结合显著性检测和稀疏表示的室内场景识别框架

Fig. 1 Indoor scene recognition framework based on saliency detection and sparse representation

1.1 显著性检测算法

当浏览一个场景的时候,人类视觉具有倾向于忽略不重要的区域而快速地搜索到感兴趣目标的能力.这些区域通常被大家称为奇异点、视觉焦点或者是显著性区域^[12].图像显著性区域检测能够帮助大家采用不同的策略处理不同的区域.例如,通常采用精确的方法处理显著性区域,采用近似的方法处理非重要的区域.通过这种方式,避免了对整幅图像应用复杂的算法,从而提高了图像处理速度.

最早提出来的显著性算法是基于生物启发模型.这个方法模拟人的视觉神经,通过计算“center-surround difference”来获取到显著性对象的位置信息.但是由于该方法只计算局部特征的对比度,得到的显著度图只高亮了对象的边缘信息.因此,后来的显著性区域检测算法主要集中尝试利用各种策略避免结果中只高亮边缘信息.当前显著性检测的方法有很多,其中最为典型的方法有基于局部对比分析的算法、基于图论的算法、基于频谱特性的 SR(Spectrum Residual)算法等^[13],这些方法遵循的视觉显著性规律不同、使用的图像特征不同、对特征的处理方式也不同,所得到的显著图也都有自己的特点,但从所有方法中都可以发现设计显著性检测方法的基本思路.

本文使用的显著性目标检测方法是 Huaizu、Jiang 等人在 2013 年提出的不同区域特征融合

(Discriminative Regional Feature Integration, DRFI) 的显著性检测方法^[14].

该方法有 3 个主要步骤组成,包括多尺度分割、区域显著性计算和多尺度显著性融合. 其原理如图 2 (详见文献[14])所示.

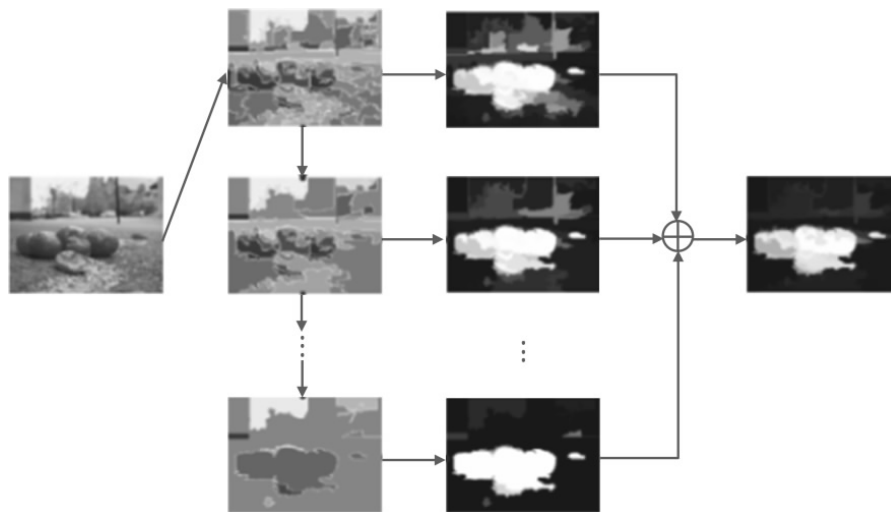


图 2 不同区域特征融合显著性检测方法原理示意图

Fig. 2 Schematic diagram of the principle of saliency detection method in different regional feature fusion

1.1.1 多尺度分割

多尺度分割中采用基于图的图像分割方法. 给一张原始图片 I , 将其进行 M 尺度的分割得到 $S = \{S_1, S_2, \dots, S_M\}$, 其中每个 S_m 都是原始图片 I 的分割, 包括 K_m 个区域. S_1 是最好的分割, 它包含了最大数量的分割区域, S_M 是最粗糙的分割, 它拥有最少的分割区域.

1.1.2 显著性分数计算

和大多数显著性检测方法不同的是, 该显著性算法是先设计出一些对比特征向量, 然后将这些有效的对比特征向量作为训练特征来训练一个随机森林回归器, 该回归器负责计算不同图片的显著性分数.

显著性特征向量分为 3 个主要部分: 区域对比描述、区域属性描述和区域背景特征描述. 其中区域对比描述主要包括每张分割图中相邻区域之间颜色和纹理之间的差异, 包括 RGB 均值、 $L * a * b$ 均值、LM 滤波器绝对响应、LM 滤波器的最大响应、 $L * a * b$ 直方图、灰度直方图、饱和度直方图和纹理直方图. 其中一般特征向量之间的差异以如下形式来计算:

$$d(x_1, x_2) = (|x_{11} - x_{21}|, \dots, |x_{1d} - x_{2d}|). \quad (1)$$

而直方图之间的差异计算方式为:

$$\chi^2(h_1, h_2) = \sum_{i=1}^b \frac{2(h_{1i} - h_{2i})^2}{h_{1i} + h_{2i}}. \quad (2)$$

区域属性描述又包括外表属性描述和集合属性描述, 外表特征视图描述图像区域的颜色和纹理特征, 它们可以作为区分显著性区域和背景的最基本的属性. 几何特征包括大小和位置关系, 这些对描述显著性和背景之间的空间关系有一定影响.

1.1.3 多尺度显著性融合

多尺度显著性融合的目的就是将多尺度分割后计算得到的显著性图融合成一张显著性图, 该方法使用了一个线性融合器 $A = \sum_{m=1}^M \omega_m A_m$ 来进行融合操作, 这个线性融合器通过使用最小化均方误差估来学习参数, 即最小化目标函数:

$$\left\| A - \sum_{m=1}^M \omega_m A_m \right\|_F^2. \quad (3)$$

1.2 稀疏表示分类算法

稀疏表示 (Sparse Representation) 理论是一种新兴的信号表示方法, 此方法使用超完备字典对信号进

行分解,对信号的误差与噪声比传统方法更稳健.在图像压缩领域中,更稀疏的字典能够得到更高的压缩比;在图像重建领域中,更稀疏意味着更高质量的图像重建.由此可见,稀疏性对于图像表示(Image Representation)是至关重要的.另外,基于稀疏分类框架^[15-17]的目标跟踪算法和图像分类算法具有独特的抗噪声与遮挡能力.

信号的稀疏表示是数据表示体系的重要组成部分之一.设字典 \mathbf{A} 由一组线性独立的基矢量 $[a_1, a_2, \dots, a_M]$ (原子)组成,这些基矢量能够张成整个矢量空间 $\mathbf{X} = [x_1, x_2, \dots, x_M] \in \mathbf{R}^M$,即空间中任意矢量 x 都可以通过这组基的线性组合进行重构,如式(4)所示:

$$x = \sum_{i=1}^M c_i \cdot a_i. \quad (4)$$

式中, $c_i = \langle a_i, x \rangle$ 是 x 在基矢量 a_i 上的展开系数.因为基是相互独立的,则这种展开的结果是唯一的.如果 $a_i \perp a_j$,则字典 \mathbf{A} 为空间 \mathbf{X} 的一组正交基.则式(4)可改写为式(5):

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{C}. \quad (5)$$

式中, $\mathbf{A} \in \mathbf{R}^{M \times N}$ 为字典矩阵,且每一行为 \mathbf{X} 的一个基矢量.若 $N < M$,则称字典 \mathbf{A} 为欠完备字典,即空间 \mathbf{X} 中有些矢量无法由字典有效表示;若 $N = M$,则称 \mathbf{A} 为完备字典;若 $N > M$,则称 \mathbf{A} 为超完备字典, \mathbf{C} 有无穷多解.

在过完备字典上求解稀疏表示问题可以用 ℓ_1 范数最小化方法来解决:

$$\hat{x}_1 = \operatorname{argmin} \|x\|_1 \quad s.t. \quad \mathbf{A}x = y. \quad (6)$$

在式(6)中, \mathbf{A} 代表训练字典, y 表示测试样本, x 表示稀疏系数.

在经典的稀疏表示算法中,稀疏表示拟解决的问题可以表示如式(7)所示:

$$\hat{\alpha} = \operatorname{argmin}_{\alpha} \|\alpha\|_1 \quad s.t. \quad \|y - \mathbf{X}\alpha\|_2 < \varepsilon. \quad (7)$$

式中, y 表示测试样本, \mathbf{X} 表示训练字典, α 则表示测试样本 y 在训练字典 \mathbf{X} 下的系数.通过计算每一类的残差 $e_i(y) = \|y - \mathbf{X}_i \hat{\alpha}_i\|_2$,并根据残差大小,从而可以判断测试样本所属的类别.

稀疏表示方法是基于每一类字典 \mathbf{X}_i 都是过完备的假设的.但本文针对的居家室内场景识别分类,这是一个小样本问题,所得到的训练字典 \mathbf{X} 也并不是完备的.如果依旧用测试样本 y 的类字典 \mathbf{X}_i 来重现样本,那么误差就会相当大,最终将会导致得到的残差 $e_i(y)$ 和 $\|\alpha\|_1$ 不精确,从而严重影响分类结果.

为解决上述问题,本文拟引用最小均方差准则下的协同表示分类来改进稀疏表示模型,从而大大提高算法的识别速度和分类效果.

利用类与类之间的相关性,即某些第 j 类的样本可能对第 i 类的测试样本的表示有着重要意义.因此,可利用字典中其他类的图像来扩充本类图像.在文献[15]中,就是利用这种方法来解决小样本问题的.

这样,实际上就是在 ℓ_1 范数的限制下,利用字典 $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n]$ 中所有的数据来协同表示测试样本 y .那么接近优化式(7)就可变为式(8):

$$\hat{\alpha} = \operatorname{argmin}_{\alpha} \|y - \mathbf{X}\alpha\|_2. \quad (8)$$

有关联表示 $\hat{y} = \sum_i \mathbf{X}_i \hat{\alpha}_i$ 为 y 在平面 \mathbf{X} 的投影.在 SRC(sparse representation-based classifier) 中,残差 $e_i(y) = \|y - \mathbf{X}_i \hat{\alpha}_i\|_2$ 用于分类,则可以导出:

$$e_i(y) = \|y - \mathbf{X}_i \hat{\alpha}_i\|_2 = \|y - \hat{y}\|_2 + \|\hat{y} - \mathbf{X}_i \hat{\alpha}_i\|_2. \quad (9)$$

在式(9)中, $\|y - \hat{y}\|_2$ 对于所有的类来说都是连续的,那么很显然真正起作用的其实是:

$$e_i^* = \|\hat{y} - \mathbf{X}_i \hat{\alpha}_i\|_2. \quad (10)$$

如图 3 所示, $\chi_i = \mathbf{X}_i \hat{\alpha}_i$ 和 $\bar{\chi}_i = \sum_{j \neq i} \mathbf{X}_j \hat{\alpha}_j$.图 3 中显示了测试样本 y 在平面 \mathbf{X} 的投影,可看出 $\bar{\chi}_i$ 与 $y - \mathbf{X}_i \hat{\alpha}_i$ 平行,则:

$$\frac{\|\hat{y}\|_2}{\sin(\chi_i, \bar{\chi}_i)} = \frac{\|\hat{y} - \mathbf{X}_i \hat{\alpha}_i\|_2}{\sin(\hat{y}, \chi_i)}. \quad (11)$$

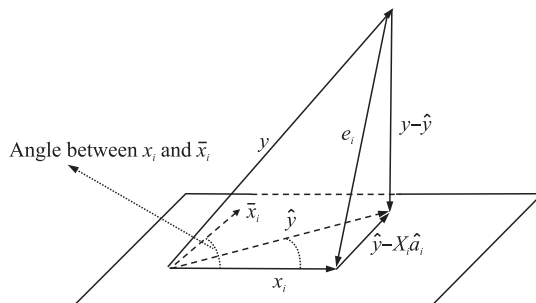


图 3 协同表示原理示意图

Fig. 3 Schematic diagram of the principle of collaborative representation

式中, $(\chi_i, \bar{\chi}_i)$ 是 χ_i 和 $\bar{\chi}_i$ 之间的夹角, (\hat{y}, χ_i) 是 \hat{y} 和 χ_i 之间的夹角, 根据式(11)可得出:

$$e_i^* = \frac{\sin^2(\hat{y}, \chi_i) \|\hat{y}\|_2^2}{\sin^2(\chi_i, \bar{\chi}_i)}. \quad (12)$$

从式(12)中, 可以看出, 在判断测试样本 y 是否属于第 i 类时, 不仅需要考虑 \hat{y} 和 χ_i 之间的夹角大小, 还需要兼顾到 χ_i 和 $\bar{\chi}_i$ 之间的夹角大小, 正是这种双重标准使得分类变得更加高效和鲁棒.

2 实验及分析

本文通过 MATLAB 平台进行仿真实验, 验证本文算法的场景分类效果. 本文的实验分为四个部分: 第一个部分为显著性检测算法使用不同阈值时室内场景分类结果的对比实验; 第二个部分为使用单个特征与使用多个特征融合分类结果的对比实验; 第三部分为前两部分实验最优结果的结合, 即本文最终算法的实验结果; 最后一部分是检测本文算法复杂度的实验, 实验通过处理不同像素大小的图片所需要的 cpu 时间来检验算法的复杂度.

前三部分的实验中, 所采用的室内场景均分为卫生间、卧室、衣帽间、厨房、客厅这 5 个场景类别. 从每个场景类别中随机选择 50 张图片作为字典训练的样本, 并随机选择 10 张图片作为测试样本.

而最后一个实验中, 分别选择卧室、厨房、衣帽间等 5 个室内场景图片中像素大小为 256×256 、 640×480 、 1080×810 的图片各 60 张, 其中 50 张用作训练字典, 10 张用作训练样本.

本文对实验中的室内场景图片有一定的要求, 即室内场景图片要相对简洁, 且图片中一定有每类场景的代表性主目标, 例如卧室图片一定包含床、客厅图片一定包含沙发或椅子等.

2.1 固定阈值的显著性检测算法实验

本实验部分针对的是显著性检测算法中阈值的选取问题. 为了得到最佳参数, 将显著性检测算法中的阈值分别设为 0 (即不设置阈值)、0.2、0.4、0.6 及 0.8, 其中, 阈值的大小即表示对图像的分割程度, 阈值越大, 则分割后剩余的图像特征越少, 不设阈值则表示不对图像进行分割.

将不同阈值下得到的显著图与稀疏表示算法结合后进行室内场景的分类. 通过比较室内场景分类的正确率, 从而得到分类效果最好的显著性检测算法阈值. 图 4 为不同阈值情况下, 典型室内场景分类的平均正确率.

由图 4 可知, 不同的阈值设定使得室内场景的分类正确率不同, 且当显著性检测算法的阈值设为 0.4 的时候, 室内场景的分类效果最好, 正确率为 52%. 因此, 本文在最终算法中将把显著性检测算法的阈值设置为 0.4.

2.2 显著性区域特征与灰度图特征融合实验

本实验部分将各个场景的灰度图特征、显著图 (无阈值) 特征以及灰度图与显著图的融合特征分别作为稀疏表示算法的输入进行场景的分类, 并比较各个实验的分类正确率, 如图 5 所示.

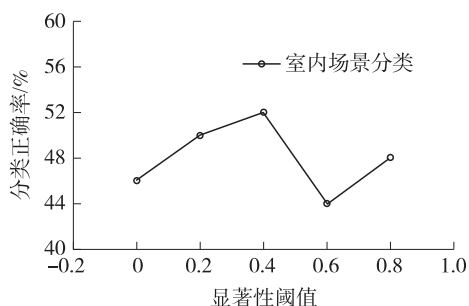


图 4 不同阈值下室内场景的分类正确率

Fig. 4 Classification accuracy of indoor scenes with different thresholds

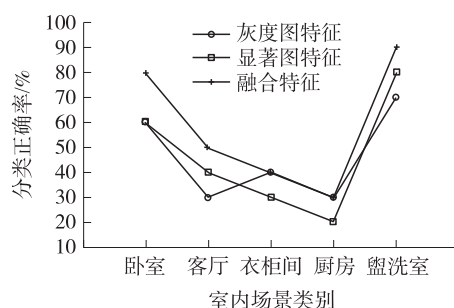


图 5 不同特征下室内场景的分类正确率

Fig. 5 Classification accuracy of indoor scenes with different features

由图 5 可知, 用单个特征即灰度图特征与显著图特征作为稀疏表示算法的输入时, 室内场景的分类正确率为 46%, 而用灰度图与显著图的融合特征作为算法的输入时, 场景的识别正确率为 56%, 这表明使用融合特征在一定程度上可以提高室内场景的分类正确率, 比使用单个特征的分类效果好. 因此, 在本文的

最终算法中,将采用特征融合的方式.

2.3 多特征融合算法实验

根据 2.1 章节与 2.2 章节的实验结果,本实验部分将显著性检测算法的阈值设为 0.4,并将融合特征作为稀疏表示算法的输入,即本文的最终算法为:将阈值为 0.4 的显著图与灰度图及显著图(未设阈值)特征融合,并作为稀疏表示算法的输入进行场景分类. 通过实验可知,本文算法的分类正确率为 62%,比 2.1 章节与 2.2 章节的最好正确率都有所提高. 可见本文算法能够提高室内场景的分类正确率,有一定的实用价值.

图 6 为部分场景及其灰度图与显著图.



图 6 部分场景及其灰度图与显著图
Fig. 6 Part of the scene and its grayscale and significant figure

2.4 算法复杂度分析

本实验部分,将像素大小为 256×256 、 640×480 、 1080×810 的图片分别进行实验,实验中,显著性检测的阈值设置为 0.4,并采用融合特征作为稀疏表示算法的输入. 通过比较不同像素大小的图片在实验时所需要的 cpu 时间来检验算法的复杂度,并通过柱状图直观展现. (其中,实验所用电脑配置为: Intel Core i5 3.20 GHz,4 GB 内存.)

如图 7 所示,可以看出,同一实验条件下,本文算法在处理不同像素大小的图片时所需的 cpu 时间基本不变,由此说明本文算法的时间复杂度很小.

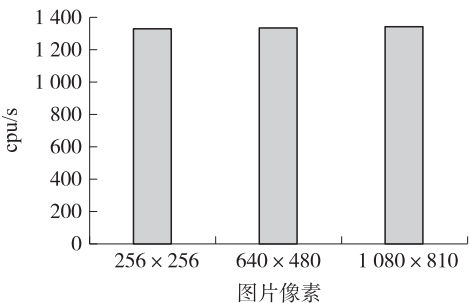


图 7 不同像素图片实验所需 CPU 时间
Fig. 7 Different pixel picture experiment required CPU time

3 结语

针对多数常规算法在室外环境表现良好、在室内环境性能下降的问题,本文提出一种基于显著性检测的稀疏表示室内场景识别方法,创新性地将稀疏表示方法应用在家庭室内场景的识别中. 该方法利用显著性检测算法找出场景图像中人眼感兴趣的区域,并通过设置阈值提取出主要目标区域,并将其与稀疏表示算法结合进行场景的分类. 在实验中,将显著性检测算法的阈值设为 0.4,并使用了多特征融合的方法. 通过这种方式,得到了各个场景的代表性目标,避免了对整幅图像应用复杂的算法,从而提高了图像处理速度与室内场景分类的正确率.

[参考文献]

[1] VAILAVA A,JAIN A,ZHANG H J. On image classification: city vs. landscape [C]//IEEE Workshop on Content-Based Access of Image and Video Libraries. Piscataway, USA:IEEE,1998:3-8.
[2] CHANG E,GOH K,SYCHAY G,et al. CBSA:content-based soft annotation for multimodal image retrieval using bayes point machines[J]. IEEE transactions on circuits and systems for video technology,2003,13(1):26-38.

- [3] 钱堃,马旭东,戴先中,等. 基于层次化 SLAM 的未知环境级联地图创建方法[J]. 机器人,2011,33(6):736-741.
- [4] 包加桐,宋爱国,郭晏,等. 基于 SURF 特征跟踪的动态手势识别算法[J]. 机器人,2011,33(4):482-489.
- [5] ZHANG H B, SU S Z, LI S Z, et al. Seeing actions through scene context[J]. IEEE visual communications and image processing, 2013, 8 575(VCIP):1-6.
- [6] BEKIOS-CALFA J, BUENAPASADA J M, BAUMELA L. Robust gender recognition by exploiting facial attributes dependencies[J]. Pattern recognition letters, 2014, 36:228-234.
- [7] LI F F, PERONA P. A Bayesian hierarchical model for learning natural scene categories[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition. USA:IEEE Computer Society, 2005:524-531.
- [8] LAZEBNIK S, SCHMID C, PONCE J. Beyond bags of features:spatial pyramid matching for recognizing natural scene categories[J]. IEEE conference on computer vision and pattern recognition(CVPR), 2006, 2:2169-2178.
- [9] YANG B J, YU K, GONG Y, et al. Linear spatial pyramid matching using sparse coding for image classification[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition. USA:IEEE, 2009:1794-1801.
- [10] WANG J, YANG J, YU K, et al. Locality-constrained linear coding for image classification[C]//IEEE Computer Society on Computer, Vision and Pattern Recognition. US:IEEE, 2010:3360-3367.
- [11] BOUREAU Y L, BACH F, LECUN Y, et al. Learning mid-level features for recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition. USA:IEEE, 2010:2559-2556.
- [12] GOPALAKRISHNAN V, HU Y, RAJAN D. Random walks on graphs to model saliency in images[C]//IEEE Conference on Computer Vision and Pattern Recognition. USA:IEEE, 2009:1698-1705.
- [13] AVIDAN S, SHAMIR A. Seam carving for content aware image resizing[J]. ACM Transactions on Graphics, 2007, 26(3):10-16.
- [14] JIANG H, WANG J, YUAN Z, et al. Salient object detection:a discriminative regional feature integration approach[C]//IEEE Conference on Computer Vision and Pattern Recognition. USA:IEEE, 2014:2083-2090.
- [15] WRIGHT J, YANG A Y, GANESH A, et al. Robust face recognition via sparse representation[J]. IEEE transactions on pattern analysis and machine intelligence, 2009, 31(2):210-227.
- [16] XUE M, HAIBIN L. Robust visual tracking and vehicle classification via Sparse representation[J]. IEEE transactions on software engineering, 2011, 33(11):2259-2272.
- [17] HAN A, JIAO J, ZHANG B, et al. Visual object tracking via sample-based adaptive sparse representation[J]. Pattern recognition, 2011, 44(9):2170-2183.
- [18] HUANG F X. Beyond bag of latent topics:spatial pyramid matching for scene category recognition[J]. 浙江大学学报(英文版), 2015, 16(10):817-828.
- [19] HAYAT M, KHAN S H, BENNAMOUN M, et al. A Spatial layout and scale invariant feature representation for indoor scene classification[J]. Computer science, 2015.

[责任编辑:黄 敏]