

博士学位论文

论文题目： （三号黑体，居中）

\_

**作者姓名**  **（楷体三号，加粗，居中）**

**指导教师** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**学科专业** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**所在学院** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**提交日期** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**浙江工业大学硕士学位论文**

**论文题目（中文，小二号黑体）**

**作者姓名：三号楷体**

**指导教师：张三教授、第二导师**

**浙江工业大学计算机科学与技术学院**

**2010年 月**

**Dissertation Submitted to Zhejiang University of Technology**

**for the Degree of Master**

**Dissertation Title (Times New Roman, Bold, Centered, Multi-lines allowed)**

**Candidate:**

**Advisor:**

**College of Computer Science and Technology**

**Zhejiang University of Technology**

**Nov 2010**

**浙江工业大学**

**学位论文原创性声明**

本人郑重声明：所提交的学位论文是本人在导师的指导下，独立进行研究工作所取得的研究成果。除文中已经加以标注引用的内容外，本论文不包含其他个人或集体已经发表或撰写过的研究成果，也不含为获得浙江工业大学或其它教育机构的学位证书而使用过的材料。对本文的研究作出重要贡献的个人和集体，均已在文中以明确方式标明。本人承担本声明的法律责任。

作者签名： 日期： 年 月 日

**学位论文版权使用授权书**

本学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权浙江工业大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于

1、保密□，在一年解密后适用本授权书。

2、保密□，在三年解密后适用本授权书。

3、不保密□。

（请在以上相应方框内打“√”）

作者签名： 日期： 年 月 日

导师签名： 日期： 年 月 日

**基于先验和统计信息的前景背景分割方法研究**

**（小二号黑体，1或2行居中，小四号黑体上空3行）**

# 摘 要 (三号黑体，小四号黑体上空一行)

摘要的第一段主要阐述选题的意义。

第二段简单介绍研究动机、方法、目的。从存在的问题出发，阐述你采用什么方法，依据什么理论来做什么事，从而解决了前面指出的那些问题及取得的成果。本文的主要工作和成果如下：

1．针对....采用...，提出了,,,，效果如何；

2．500字左右；摘要应表达学位论文的中心内容，简短明了，摘取论文中的基本信息，体现科研工作的核心思想。内容包括本文的目的和意义、研究方法、研究成果、结论。主要介绍自己的工作和取得的成果。

5．（可能不止5点）最后，对全文进行总结，并对进一步的研究提出一些展望。

**关键词**：关键词数量为4－6个，每一关键词之间用逗号分开，最后一个关键词不用标点符号

**DISSERTATION TITLE IN ENGLISH （英文标题，小二号，Times New Romans字体，居中，全部用大写字母， 可分成1-3行）**

**ABSTRACT**

The abstract in English goes here. Abstract in English and that in Chinese presented on the previous page should agree. (英文摘要内容, 用小四号Times New Romans字体，每段开头留4个字符空格，英文摘要的内容应与中文摘要基本相对应)

**Key Words**: Key words in English go here and should be separated by a comma, but no comma is allowed after the last key word （关键词用小四号Times New Romans字体, 全部小写，每一关键词之间逗号分开，最后一个关键词后不打标点符号）

**目 录**（小二号黑体居中）

[摘 要 (三号黑体，小四号黑体上空一行) i](#_Toc536187934)

[第1章 绪 论（小二号黑体居中） - 1 -](#_Toc536187935)

[1.1 课题背景和意义 - 1 -](#_Toc536187936)

[1.1.1 小节标题（小四号黑体左起顶格打印，小节标号后空两格） - 1 -](#_Toc536187937)

[1.1.2 小节标题 - 1 -](#_Toc536187938)

[1.2 国内外研究现状 - 2 -](#_Toc536187939)

[1.2.1 显著性分割方法研究现状 - 2 -](#_Toc536187940)

[1.2.2 前景背景分割方法研究现状 - 2 -](#_Toc536187941)

[1.2.3 基于主题模型的分割方法研究现状 - 2 -](#_Toc536187942)

[第2章 图像分割图 表（小二号黑体居中） - 6 -](#_Toc536187943)

[2.1 节标题 - 6 -](#_Toc536187944)

[2.1.1 小节标题 - 6 -](#_Toc536187945)

[第3章 基于前景背景先验的室内场景多模态区域一致性显著性检测研究 - 8 -](#_Toc536187946)

[3.1引言 - 8 -](#_Toc536187947)

[3.2 相关工作 - 9 -](#_Toc536187948)

[3.3 算法介绍 - 11 -](#_Toc536187949)

[3.3.1多模态数据的过分割 - 12 -](#_Toc536187950)

[3.3.2 前景先验 - 13 -](#_Toc536187951)

[3.3.3 背景先验 - 14 -](#_Toc536187952)

[3.3.4 多模态区域显著性 - 15 -](#_Toc536187953)

[3.4 实验结果与分析 - 15 -](#_Toc536187954)

[3.4.1 评估多模态过分割 - 16 -](#_Toc536187955)

[3.4.2 多模态区域显著性的评估 - 19 -](#_Toc536187956)

[3.4.3 在真实机器人系统上评估实验 - 22 -](#_Toc536187957)

[第？章 结论与展望 - 24 -](#_Toc536187958)

[？.1 结论 - 24 -](#_Toc536187959)

[？.2 展望 - 24 -](#_Toc536187960)

[参 考 文 献（小二号黑体居中，字间空两格） - 25 -](#_Toc536187961)

[附 录（小二号黑体居中） - 28 -](#_Toc536187962)

[致 谢（小二号黑体居中） - 29 -](#_Toc536187963)

[攻读学位期间参加的科研项目和成果 - 30 -](#_Toc536187964)

注意事项：在具体章的题目中可能有空格，但在目录中无需空格；

可以点击鼠标右键并选择“更新域”（只更新页码）进行自动更新页码，在页码两边可能会出现两短划线，需要删除；

注意二级、三级标题的对齐位置；

章的标题为粗体。

# 

# 第1章 绪 论（小二号黑体居中）

### 课题背景和意义

随着精密仪器科学技术的发展，3D感知技术取得了长足的进步。3D感知技术依靠被动式手段（获取环境可见光等）或主动式手段（发射激光、结构光等），通过非接触式测量这些光在物体上的反射进行实时获取环境的景深数据，从而获得类似3D化的图像，通常由2D图像和深度图像表示。3D感知设备在人们的生活、工作和娱乐等方面广泛的应用。如图

### 1.1.1 小节标题（小四号黑体左起顶格打印，小节标号后空两格）

论文正文是主体，一般由标题、正文、图、表格和公式五个部分构成。写作内容可因科研项目的性质不同而变化，一般可包括理论分析、计算方法、实验装置和测验方法，经过整理加工的实验结果分析和讨论，与理论计算结果的比较以及本研究方法与已有研究法的比较等。

绪论部分主要阐述论文选题的意义，说明研究的必要性、学术价值和实际意义，有具体项目背景的还说明项目来源。

进而，结合提出的选题，通过文献综述说明和选题相关的国内外研究现状，指出尚存在的问题，从而提出本文的研究思路，即采用什么方法、什么理论来解决刚才提出的一些问题，期望达到什么样的结论。最后阐述全文的主要结构。

### 1.1.2 小节标题

可以根据需要，分多个小节，小节前空一行。

正文：采用小四号宋体打印，里面的英文单词采用小四号Times New Roman，变量用斜体，不是变量的英文单词、字母用正体。变量应该用公式编辑器输入。

基于先验与统计信息的前景背景分割方法研究

### 1.2 国内外研究现状

### 1.2.1 显著性分割方法研究现状

2D视觉显著性检测方法研究现状

3D视觉显著性检测方法研究现状

### 1.2.2 前景背景分割方法研究现状

### 1.2.3 基于主题模型的分割方法研究现状

国内外研究现状：

显著性：

前景背景分割：

**自然语言处理中的主题模型：**

**视觉领域的主题模型：**

<Topic modeling and improvement of image representation for large-scale image retrieval> 2016

Nguyen[4]提出一种具有更鲁棒检索能力的新检索框架以及一个有效的re-ranking方法用于度量查询图像与候选数据库图像之间的相似度。每幅图像被建模为具有两阶段过程（图像编码，主题建模）的局部特征的集合。

图像编码Image Encoding：改进了VLAD（Vector of Locally Aggregated Descriptors，一种类似BOW的图像描述方法），提出新的编码模式soft-VLAD。使用基于软分配的两种聚合方法将每个局部描述符映射到多个可视单词并将它们聚合成单个向量。第一种：基于距离的分配，每个分配被指定到一个跟视觉单词到局部描述子之间的距离成正比的权重。第二种：基于稀疏编码的分配，采用稀疏编码算法将局部描述子映射到学习的码书上，然后计算每个分配的权重。这个方法受稀疏编码在图像分类中成功应用的启发。通过使用这些方法，克服了硬分配的局限性，把每幅图像编码为高辨别的向量，方便索引和检索候选的初始列表。

基于软分配局部描述子的聚合

把图像定义为局部区域的集合，这为对象姿势的变化以及局部形变提供了一定程度的鲁棒性。聚合了给定码书相关联的描述子去产生向量表示。

每个局部描述符

主题建模Topic Modeling：利用人对物体对象外观和背景的先验信息，提出一种产生式隐含主题模型，Spatial Latent Topic Model with Background distribution（SLTMB），

提取出图像的背景区域和主题区域。每个主题对应一个对象或者频繁出现在图像中的对象的某个部位。具体的，一幅图像包含带有某个空间约束的对象实例，每个对象实例也能被相关视觉单词的外观特征所表示。提出的SLTMB模型旨在利用针对不同主题的视觉单词上的概率分布。经常在同一幅图像中以某种特别空间分布同时出现的视觉单词往往属于同一个对象或者主题区域。自然地，相同的对象实例有相似的概率分布，因此SLTMB能使我们推测出图像中出现的未知对象是什么以及对象的位置在哪。结合到主题模型中的视觉单词的空间位置有效的用来补偿编码图像的产生的约束限制，加强了图像表示。

重排序Re-ranking：提取出候选图像的主题区域后，利用主题信息精炼搜索结果。如果查询图像跟数据库中的图像包含相似的主题区域，那么它们是相似图像。提出一种re-ranking算法，是一种针对大尺度图像匹配的快速有效几何打分模式。（1）建立查询图像跟候选图像之间的匹配特征对。（2）基于匹配特征对的几何信息生成共同主题区域之间的相似性得分。（3）每个图像对的新得分是通过计算主题相似性得分的和得到的，然后用这个新得分去执行re-ranking。这个策略大大加速了re-ranking并且执行在中等大小的数据集上。此外，采用抽取出来的主题区域，文章的方法能处理每幅图像中对象外观的局部变化。

总结：提出一种soft-VLAD编码方案生成整幅图像的向量描述，其中通过采用软分配聚合的方式扩展了VLAD方法，基于人对图像结构的先验信息构建的主题模型，限定不同主题和背景区域的外观和位置

# 

# 第2章 图像分割图 表（小二号黑体居中）

### 2.1 节标题

### 2.1.1 小节标题

图：图的前后各空一行，图题（反映图中内容的文字描述）用中文5号字体。引用图应在图题右上角标出文献的来源。

图号按章顺序编排，如图3-2为第三章第二幅图。如果图中含有几个不同部分，应标注分图号，如，图3-2a，并在各分图中给出分图的图题。

绘图必须工整、清晰、规范。其中机械零件图按机械制图规格要求；示意图应能清楚反映图示内容。照片应在右下角给出放大标尺。

表格：每张表格的前后空一行，表格按章顺序编号，如表5-4为第五章第四张表。每张表应有表题，表题用5号字体打印，放在表的上方。表中的内容用5号字体。

例如：图2-1：

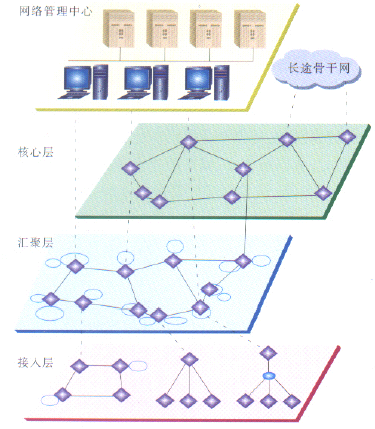


图2-1 之前空两格，5号字体

表格2-1：

表2-1 之前空两格，居中

|  |  |  |
| --- | --- | --- |
| 表格内内容选单倍行距，在“格式－段落”的段前段后选3磅 | 总的原则是美观 | 模转换成模 |
| 在“表格－表格属性”中选居中 |  |  |
| 不同 |  |  |

公式应在文中另起一行。公式居中编排，后应注明序号，按章顺序编排，且编号右对齐。对公式特别多的论文也可以按节编排，但全文要统一，例如，

 （2-1）

 （2-2）

其中，T是转置符号，d是微分符号，常数e在公式中均需采用正体。

# 第3章 基于前景背景先验的室内场景多模态区域一致性显著性检测研究

视觉显著性是复杂场景中对象检测的一种非常重要的特征。然而，基于图像显著性受到室内场景中杂乱背景和相似对象的影响，并且基于像素的显著性不能给整个对象提供一致的显著性。因此，在本文中，我们提出一种基于室内场景获得的多模态数据去计算视觉显著性图并且同时保持区域一致性的方法。来自一个场景的多模态数据首先是通过RGBD相机获取得到的。然后用自适应方法对该场景进行过分割，用来组合场景的颜色图和深度图。基于这些过分割结果，我们提出两种先验作为提高最终显著性图的领域知识，包括从彩色图像获得的焦点区域，以及从点云数据获得的平面背景结构。因此，通过融合场景中颜色数据、深度数据和点云数据产生新的显著性图。实验中，我们将所提出的方法与最前沿的方法进行了深入的比较，并且我们还将所提出的方法应用于真实的机器人系统以检测感兴趣的对象。实验结果表明，本文提出的方法在精度和召回率方面优于其他方法。

### 3.1引言

视觉显著性，作为一种重要的视觉特征，最近几年已经得到越来越多的关注。[1-4, 26, 27]视觉显著性表明子区域是否与周围环境有明显差异兵器能快速引起观察用户的注意。自从Itti[7]提出并应用计算机视觉的视觉注意力理论以来，所有后来提出的方法都是遵循文献[3]中总结的人类视觉注意四个基本原则中的一个或者几个，这些原则考虑了局部底层因素，全局因素，视觉组织规则以及高级因素。视觉显著性可以被用到许多计算机视觉任务[28]或者图像处理[29]中。许多机器人系统也利用视觉显著性去进行对象识别与检测[5, 6]。

当采用视觉显著性进行对象检测时，由于人类总是把对象看作一个整体，因此很重要的一点是一个对象中的所有子区域具有相同或相似的显著性值。然而，大多数已有的方法使用基于像素的方法去计算显著性，这意味着显著性值能在像素间变化。此外，大部分已存在的计算显著性值的方法仅仅是基于二维图像的，其中显著对象和背景之间的相似颜色和纹理会造成显著性的提取正确率降低。

为了解决上述提到的问题，通过采用多模态数据，我们提出一种新方法来计算基于区域的显著性。提出的方法首先精确地分割对象，然后将唯一的显著性分配给属于一个对象的区域中的所有像素点。基于多模态数据，本文提出的方法利用不同模态之间的互补性，因此减少其他对象和背景造成的影响。特别地，本文首先根据一种自适应方法，通过使用场景的颜色图和深度图，对场景进行过分割。接下来，通过扩展文献[10]中的方法，对颜色图和点云数据中的每个过分割去计算局部区域对比的显著性。然后，考虑基于两种先验（cues=先验？）的显著性，如计算颜色图焦点区域的先验，以及计算点云的平面背景的先验。最终，我们结合这两种先验信息，得到在区域中数值一致的多模态显著性图。

在实验中，我们首先用过分割精确度去评估颜色图和深度图的互补效果。然后，将我们的方法和目前提出的最好的方法进行比较，并且将我们提出的方法应用到真实机器人系统上去检测感兴趣对象。

本文的剩余部分结构安排如下。首先在下一节回顾显著性的相关工作。在第三节详细描述我们提出的方法，接下来是实验结果与讨论。最后一小节，我们得出结论。

### 3.2 相关工作

显著性是根据人类视觉注意机制，人类借助显著性就能够迅速注意到跟周围环境的差异性最大的区域。人类的视觉注意力的四个基本原则[3]已经被大多数显著性方法所采用（读者可以阅读详细的综述文章[8]）。本文主要关注的是自下而上的显著性检测相关文献，包括在[3]中的三个基本原则。这些方法可以被简单的分成两种原子计算单元的组别：基于像素的显著性检测方法[2, 3, 12-17]，以及基于区域的显著性检测方法[9-11]。

在第一组中，基于像素级别计算得到的显著性，其中每个像素有唯一的显著性值。在文献[2]中，作者提出采用高维高斯滤波去计算完整的对比度和显著性，得到像素精确的显著性图。文献[3]中，考虑多尺度的局部对比度和全局对比度，以及像素的直接上下文，来计算像素级别显著性图。文献[12][13]中，作者提出类似过滤器的方法，根据颜色对比度和光照对比度来计算显著性图。文献[14]中，作者提出采用马尔科夫方法，通过定义两个点的不相似性为其中一个点与两个量的比率之间的距离，来计算像素级别的显著性图。文献[15][16][17]中，通过采用增量编码长度或者有损编码长度来提出基于光谱信息的方法，以测量每个特征的透视熵增益。所有这些方法都能获得可以突出感兴趣对象的比较好的结果，然而，基于像素的方法因为环境干扰也可以给背景过高的显著性。比如图1，由于第一行桌面上的镜面反射引起的两个高亮区域具有很高的显著性。并且在第二行中的两个垃圾桶跟背景颜色相似，这两个垃圾桶的显著性值是很低的。

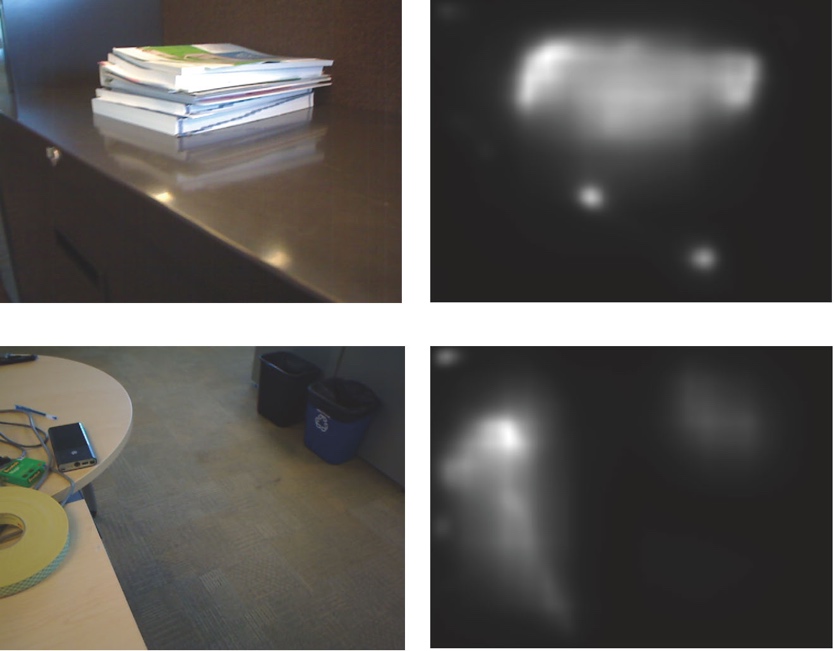


图3.1 被背景干扰的显著性图的例子

最近，越来越多的基于区域的显著性检测方法被研究人员提出来。这是因为当旨在解决显著对象检测问题时，区域一致显著性能获得更好的对象检测结果。文献[9]中，作者提出对图像进行过分割，然后通过图像的其他部分计算由一个或多个过分割区域组成的窗口的显著性值。Cheng等人[10]提出基于区域的显著性，其中一个过分割区域中的所有像素具有相同的显著性值，并且文献[30]在这个基础上结合颜色信息扩展了该方法。这些方法结合了超像素和显著性的优势，并且因此能检测具有高显著性值的显著对象。然而，这些方法只是针对二维图像，并且因此显著性检测的性能也会因为图像中存在两个相似的对象或者混乱的背景而降低。最近，Niu等人[11]扩展了文献[10]中的方法，其中提出结合二维图像到过分割图像中得到视差图，并且利用视差对比和领域知识去计算区域显著性值。然而，Niu等人提出的方法丢弃了图像信息并且，领域知识不适合深度图和点云数据。

所有上述提到的方法没有考虑背景信息。但是，背景也是显著性检测的一个重要的信息。最近，一些学者利用背景先验去计算显著性值并且获得了最好的性能[23-25]。在文献[23]中，作者提出一种鲁棒的软背景连通性去增强显著性的计算。文献[24]中，作者采用了边界先验和背景连通性先验。Yang等人也根据超像素的排序，利用背景信息去计算显著性值。这些方法给我们提供了显著性的考虑，然而，这些方法都是假设跟图像边界相连的区域属于背景，但这并不总是正确的，特别当我们在混乱场景中检测到显著的对象的时候。在本文的研究中，我们能从背景检测的多模态数据中获益。我们首先通过点云数据检测平面结构，然后将场景边缘相连的平面区域作为背景，这样操作很明显更加合理。

跟已有方法不同的是，我们提出的方法利用了多模态信息的优势，比如，从二维图像中提取前景先验，从三维点云数据中计算背景先验信息，以及利用二维图像和三维信息的互补效果产生更加精确的过分割区域。

### 3.3 算法介绍

我们提出的多模态区域一致的显著性由三个步骤组成。首先，我们采用颜色图像和深度图对场景进行过分割，然后，我们计算颜色图中的焦点区域作为前景先验信息，并且检测出点云中的平面区域作为背景先验信息。最终，我们通过结合全局对比度和先验计算最终的显著性图。如图3.2所示，结合标记为红色框感兴趣区域的彩色图以及第一行中的深度图，得到多模态过分割结果（第二行左图），并且产生针对平面分割的点云（第二行右上图）。平面分割结果映射到背景先验的深度图。第三行是最终的显著图。每个平面被指定了不同的颜色。

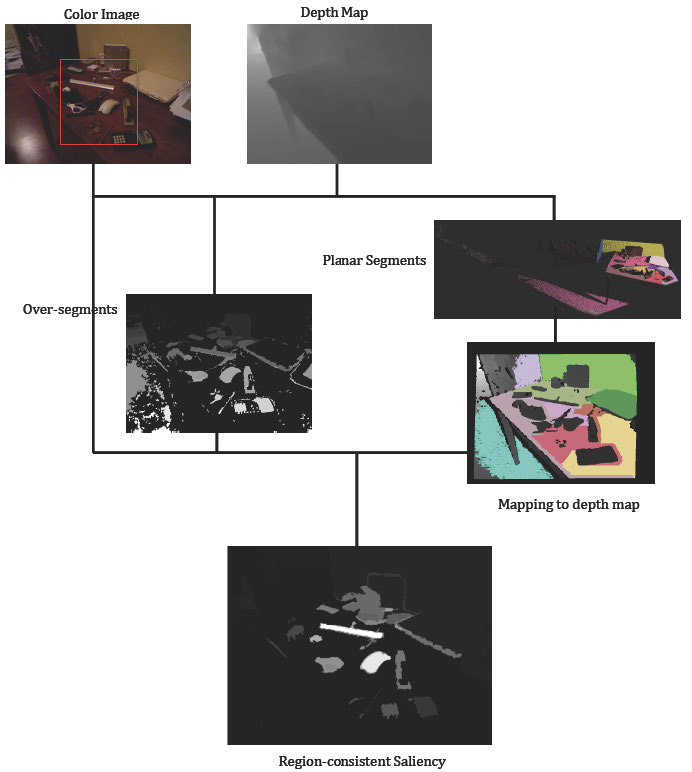


图3-2 提出的区域一致显著性方法

### 3.3.1多模态数据的过分割

在本文的研究中，我们采用文献[18]中的方法去获取图像以及对应深度图的过分割。这个算法根据图像区域特征之间的差异来分割图像。图像间特征的差异定义为：

其中，和分别是组件和的两个节点，E表示邻接关系，定义了边界权重函数。在文献[18]中，边缘权重函数是用来计算两个相邻像素之间的绝对强度差，，其中，和表示两个像素的强度。在本文的研究中，边缘权重函数同时度量颜色空间和深度空间之间的差异。边缘权重函数的正式定义如下：

其中参数控制颜色和深度通道的权重。

尽管这个公式只是简单的改进了文献[18]中的算法，然而，如何确定参数是获得满意的分割结果的关键，需要谨慎的去选择。此外，所有图像不可能设置相同的参数值，因为不同的场景存在多种图像差异和深度差异。并且因此，对于不同场景的过分割，图像和对应的深度图有不同贡献。如图3.3所示，不同的参数值将会导致不同的分割结果。当使用深度图的时候，只有少数分割部分是在深度图很大的地方得到的。因此结果的显著性图对于区别不同的对象是不精确的。当增加颜色通道的权值，分割将会变的越来越多。如果忽略深度通道并只采用颜色通道，就会有过多的分割。在这种情况下，太多的区域去计算显著性将占据过多的计算资源。并且将会得到过多的不同区域的显著性，这会减少显著性的精确度。

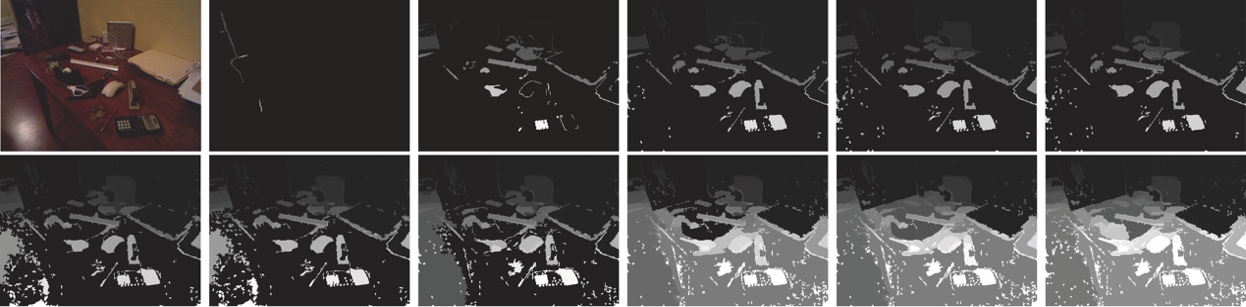


图3-3 给定不同权重参数的多模态过分割结果

我们提出一种自适应方法确定权重。首先采用文献[18]中描述的相同参数去分别分割颜色图和对应的深度图。因此可以获得两个过分割。然后通过下面的公式计算出：

其中和分别表示深度图和颜色图过分割的数目。

尽管上面的自适应方法是简单的，但是对于我们的实验是有效的。根据上述公式，我们得到深度图中的过分割数目越大，使用的值就越小。这是合理的，因为在深度图中更多的过分割意味着可以提供更多的细节信息，并且权重会变大去弥补颜色信息。此外，深度图的权重减小，颜色图的权重应该增大，来得到足够多的用于显著性计算的过分割。实验中，我们发现的值的范围是5到50，的值的范围是200到500。因此参数涌来调整颜色图和深度图之间的过分割的占的比重。实验结果上，我们发现当的值在5到15之间的时候，过分割结果是稳定的。在实验中，我们设置为10。另一个参数设为0.8确保生成过分割的时候颜色图至少占20%的贡献。因此，通过这个公式深度图和颜色图能够自动互补。

### 3.3.2 前景先验

我们观察到当一个人拍照的时候，最重要的物体将会放在照片的最焦点的位置。因此首先通过主体区域抽取算法（Subject Region Extraction, SRE）[19]提取出焦点区域(Focus Region, FR)。采用一系列不同尺寸的高斯核去卷积图像，并且计算所有卷积图像中每个像素的水平梯度和垂直梯度。然后，每幅卷积图像像素的分布通过直方图近似计算，通过下面的梯度对数似然函数来度量一个像素块区域如何较好的用一个k乘k大小的高斯核来表达：

其中和分别表示卷积图像的水平梯度分布和垂直梯度分布。N(x)指的是像素x的邻接窗口，表示邻接窗口内的一个像素。找到不同高斯核大小中最大的，该高斯核定义为，用于最佳的解释图像中一个窗口区域。高斯核的尺寸说明了该区域块的清晰程度。因此的值越小，中心为x的窗口区域块越清晰。对于每个像素，存在对应的高斯核，通过定义一个二值图来确定清楚区域还是模糊区域：

将二值图U投影到水平轴和垂直轴，统计二值图中水平方向和垂直方向上白色像素点的数目，得到两对坐标和，在这两个范围内白色像素的个数等于二值图中所有白色像素的。当我们去检测用专业摄影师拍摄的照片中的焦点区域时，其中焦点区域的清晰度有明显的差异，检测到的焦点区域是准确的。然而，我们实验的图像是通过移动机器人拍摄的，没有专业的技巧去拍着这类照片。因此，如果被设为0.9，检测到的区域太大以至于不能准确的确定出焦点区域，如图3.4所示。在我们的实验中，我们发现较小的值是适合我们的测试数据的。图3.4显示了实验中的一个例子。很明显，当是0.6时，检测到焦点区域时最准确的。



图3-4 焦点区域检测的例子

检测到焦点区域后，我们定义第一个显著性先验信息。离焦点区域越近的区域，该区域的显著性值越高。因此我们正式得出下面的公式：

其中表示过分割块和焦点区域中心点之间的欧氏距离。表示所有过分割和焦点区域之间的最大距离。

一些其他的算法[2,3,14]强调了图像中心位置的显著性值，因为这些算法明确地或隐含地使用了图像中心店作为前景先验。但是，在一些情况下，关键的物体可能不在中心区域并且中心区域的显著性值也可能是不准确的。当采用焦点区域作为前景先验的时候，由于焦点区域不会被限定在图像的中心位置，上面这个缺陷可以被解决。

### 3.3.3 背景先验

在文献[23,24]中，已经采用背景作为一种鲁棒性先验去构建精确的显著性图，然而，这些方法考虑将大部分连接到图像边缘的区域作为背景，由于有时一些物体对象刚好是处在图像的边缘处，对于这种情况上述方法就不适用了，如图3.5所示。



图3-5 物体对象出现在图像边缘时的例子

为了克服这些缺点，我们利用平面分割去检测背景。在室内场景中，存在很多平面结构，比如桌子，地面，墙壁。如果这些平面结构到达了场景边缘，他们一般会被当作背景。由于我们的数据是从类似Kinect的设备上获取的，其同时提供了对应的颜色图和深度图。因此，我们能很轻易的产生点云并且采用文献[31]的方法检测出平面区域，然后计算每个连接场景边缘的平面区域的边界长度。我们通过下面的定义作为一个平面分割块的背景连通度：

其中表示与场景边缘相邻的边界长度， 是分割平面的面积。

基于计算得到的背景连通度，我们简单用下面的公式来定义过分割区域的背景权重：

其中P表示所有检测到的平面区域。

因此，过分割区域和所有分割平面的重合区域越多，背景权重的值越大，也说明过分割区域属于背景的可能性越高并且有较小的显著性值。很明显，

值的范围是0到1，因此我们基于背景权重去定义DK显著性值如下：

### 3.3.4 多模态区域显著性

最终，我们的多模态区域一致显著性是结合前景背景先验采用全局区域对比度（GRC）计算得到的。全局区域对比度扩展了文献[10]中提出的算法，定义如下：

其中和分别表示过分割区域到之间的颜色距离和空间距离，确定空间权值的长度，根据文献[10]中的建议设为0.4，并且表示区域的权重，通过计算区域内的像素数目得到。与文献[10]中提到的不同点是，我们扩展了空间距离为内的欧氏距离，计算对应于和中心的点云中两点之间的距离。然后最终的显著性图通过如下公式获得：

### 3.4 实验结果与分析

我们从多个方面评估了我们提出的这种方法。首先我们评估从多模态数据中计算的过分割的正确率。然后通过对比目前最好的方法，基于两种不同的策略，计算分割出感兴趣对象的准确率和召回率。最终，我们在真实机器人系统上实现提出的方法并且，显示出检测感兴趣对象的结果。

前两个实验是基于两个公开的RGBD数据集[20,21]。第三个实验，机器人系统检测实验室中感兴趣的对象。

### 3.4.1 评估多模态过分割

我们评估多模态过分割区域和颜色图像过分割区域的对象边缘对齐的准确性。从两个数据集中选择1000幅彩色图像和对应的深度图作为测试数据。因为在这两个数据集中只有表示感兴趣对象的边界框，我们给所有物体增加了边缘轮廓，如图6中第一列显示的例子。我们利用过分割区域和数据集提供的Ground Truth之间的重合度来评估准确性。准确度正式的定义如下：

其中N是对象个数，表示第j个对象的真实值，表示对感兴趣的所有过分割区域的并集。

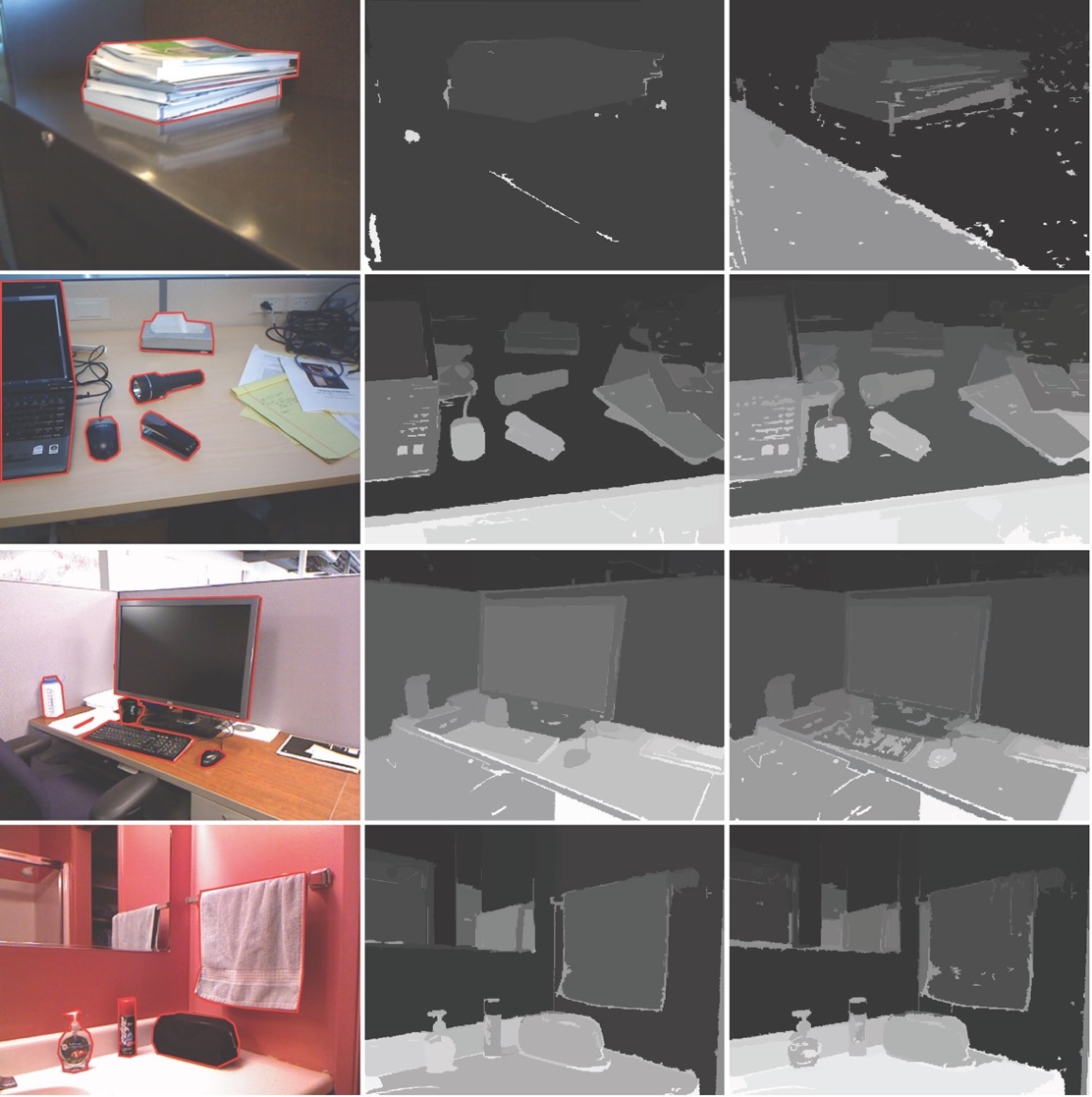


图3-6 一些过分割的例子。第一列示带有目标ground truth的原图，

第二列第三列的结果分别表示来自RGBD数据和颜色图像的过分割结果

表3-1 不同过分割方法下的对象边缘对齐的准确度

|  |  |  |
| --- | --- | --- |
|  | 多模态过分割 | 颜色图像过分割 |
| 准确度(%) | 78.5 | 72.2 |

准确度在表3-1中列出。我们提出的多模态过分割方法能够获得比基于颜色图像的过分割方法更好的准确性，这是因为多模态过分割方法能避免相似颜色和纹理的影响。对于背景噪声，比如高亮的反射光，多模态过分割方法也能避免它造成的影响并且获得更好的结果。一些过分割的例子如图3-6所示。第一行中，桌子上的两个高亮反射区域被分割成为跟第二列中桌子相同的区域，但是在第三列中，这两个区域又被分割成两个不同区域。在第三行第二幅图中，键盘和显示器之间的马克杯被正确分割为一个独立的区域，但第三幅图分割出现错误。在第二行和第四行中，我们的多模态过分割方法同样获得了更好的对齐到对象边缘。

在实验中我们研究了公式(3)中的参数和的影响。我们设置了十组不同的值，值的范围从0.1到1.0，设置0.1为间隔。那么对于每个，我们设置为值从5到15之间的不同整数值。对于每一组和，我们计算对齐获得的过分割区域的对象边缘准确性。结果如图3-7所示。当设为0.8和设为10的时候获得最好的准确性。当的值逐渐变大的时候，准确性逐渐提升。这是因为的值设置太大的时候，提供给过分割区域过量的深度信息，深度图中甚至没有足够的细节。改变的值的效果是相似的。当我们将的值从0.1增大到0.7，准确度也随之增加。然而，当大于0.8的时候，深度信息只能给过分割区域提供很少的贡献。因此，准确率不再提升。

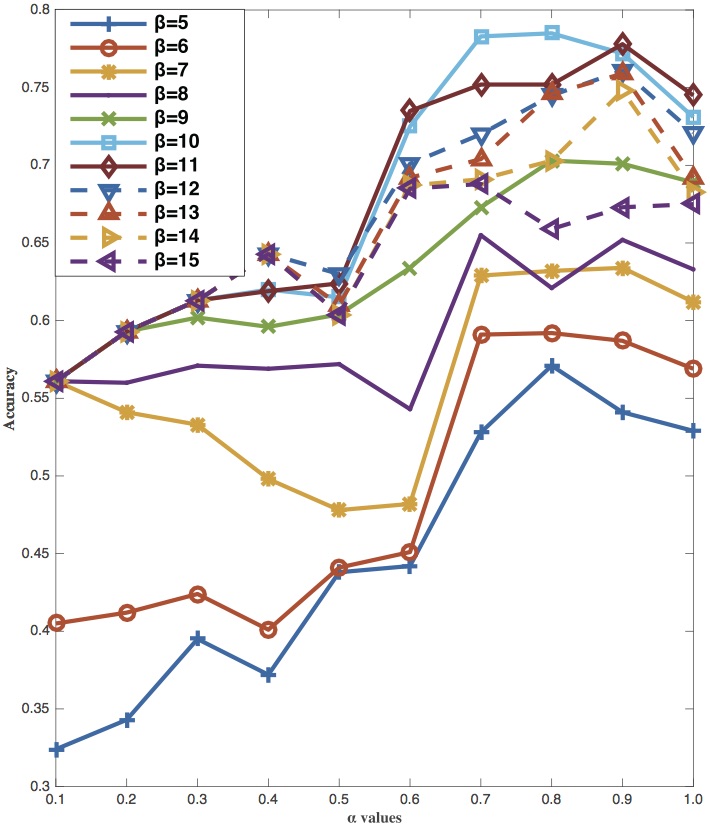


图3-7 不同的和值对于对象边缘对齐准确度的影响

### 3.4.2 多模态区域显著性的评估

这个实验中，首先评估提出的方法每一部分对于最终的显著性图的生成的贡献，然后通过比较其他经典的以及目前最好的方法，去评估多模态区域一致显著性方法的准确性。为了更全面的评估提出的方法，我们按照文献[10]中的两种策略（比如，简单的阈值分割以及迭代式图割方法）区分割显著性对象并且计算它们的准确度。对于这两种策略，分别计算精确率和召回率。

在我们的方法中，我们通过自适应的调整颜色图像和深度图像对生成最终过分割图像的贡献比例来提升过分割的效果。然后结合前景先验和背景先验信息进入GRC显著性。因此，我们采用四种不同的策略去计算显著性，如，采用改进的过分割方法加GRC的方法（IO+RC），采用改进的过分割加上前景先验信息再加上GRC的方法（IO+F+RC），采用改进的过分割加上背景先验信息再加上GRC的方法（IO+B+RC），最后一种是综合上面提到的所有方法。实验的准确率和召回率如图3-8所示。

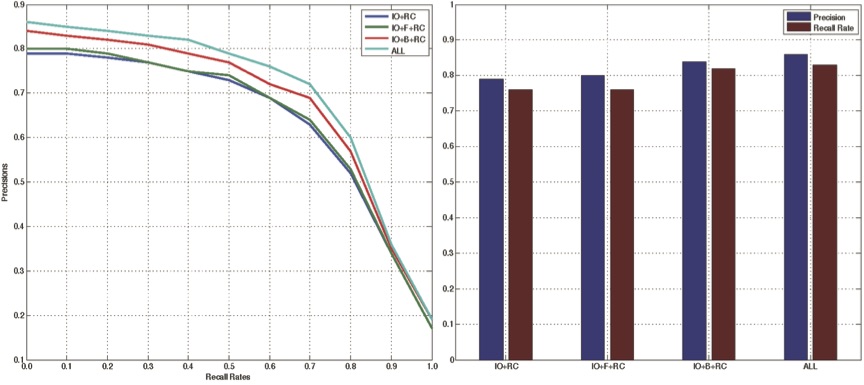


图3-8 采用不同策略获得的准确率和召回率

从图中我们发现，背景先验对于提升最终显著性图的生成有更多的贡献。如图3-8的左边所示，从“IO+RC”方法到“IO+B+RC”方法，准确率得到了很大的提升，并且如图3-8右边所示准确率和召回率的增幅分别达到5%和6%。这个说明在室内环境中，平面信息对于计算对象显著性非常重要。图3-9展示了相关的例子。如图中的第四列、第五列，我们观察到如果显著性值是通过结合背景先验信息计算的话，背景的显著性值能被极大的压抑住。

现在我们通过对比三种基于像素的方法来评估我们提出的方法：如Achanta提出的两种方法（A1）[12]和（A2）[13]，Goferman提出的方法（CA）[3]，以及对比了五种基于区域的方法：如Cheng提出的区域一致显著性方法（RC）[10]，显著性滤波器（SF）[2]，测地显著性（GS）[24]，基于图的流形排序方法（MR）[25]以及来自鲁棒性背景检测研究中的显著性优化（SO）方法。

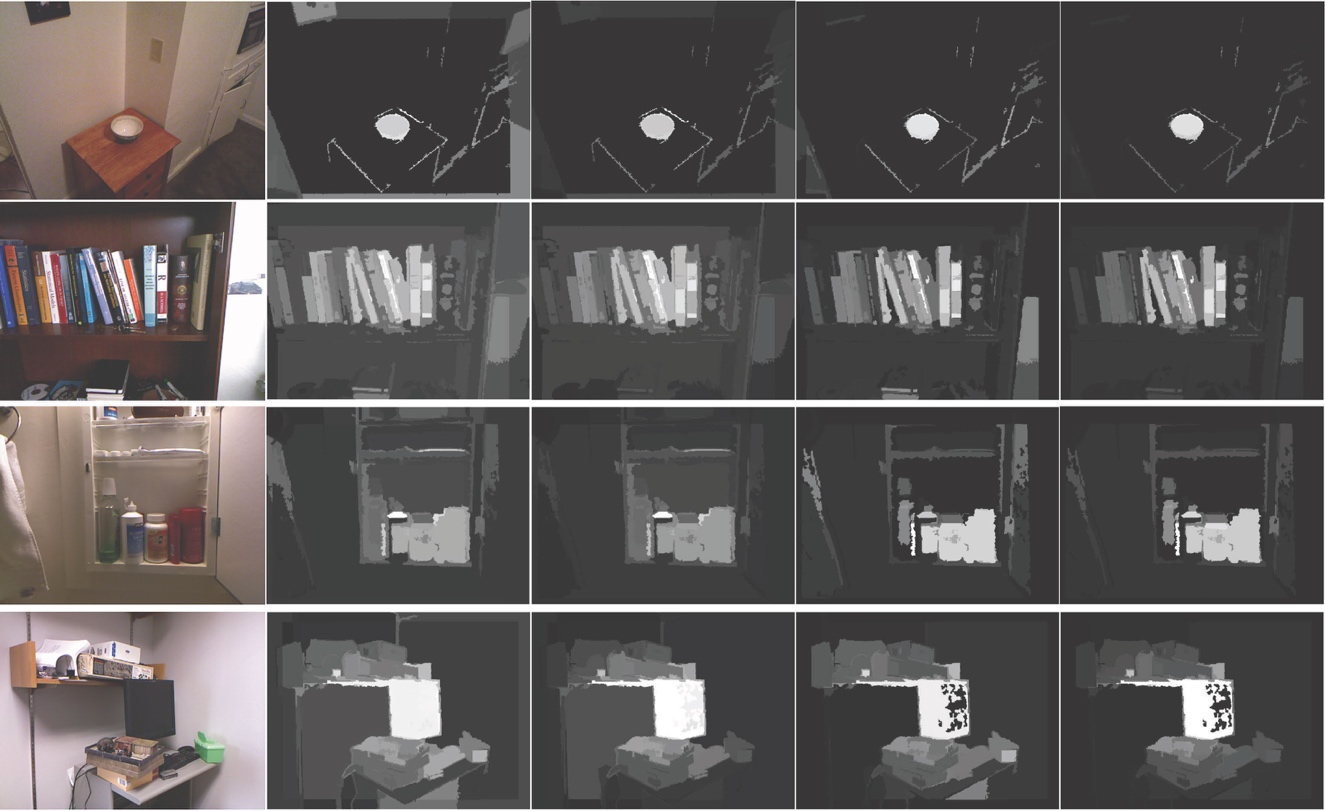


图3-9 采用不同策略计算的显著性图的例子。最左列是原图，从第二列到第五列分别是采用“IO+RC”，“IO+F+RC”，“IO+B+RC”以及综合所有的方法。

定量的对比结果如图3-10所示。在图3-10的左边，很明显地，我们提出的方法有更好的准确率和召回率。跟其他基于像素的方法对比（“A1”，“A2”，和“CA”），大多数基于区域的方法有更好的结果，因为基于区域的方法能够获取区域一致的显著性。由于我们着重设计了针对这种场景的背景先验，我们提出的多模态区域一致显著性方法在室内场景中比其他方法有更好的性能。相同的结果也能在图3-10的右边看到，我们提出的方法都得到了最好的性能。

值得注意的是所有方法在本文中的性能要比这些方法在所提出的论文中的性能要差。这是因为这些方法的性能是通过在经典数据集上的测试获得的。然而，我们的测试数据存在更复杂的背景，并且在大多数图像中，存在不只一个的感兴趣对象。因此，在真实环境中很难获得精确的显著性区域。

我们采用上述提到的方法和我们提出的方法去计算的显著性图，简单场景如图3-11所示，复杂场景如图3-12所示，其中我们提出的方法记为MRS。很明显地，我们提出的方法是唯一可以正确计算显著对象的显著性，比如在两个场景中的两只碗。对于三种基于像素的方法，采用显著性图去精确检测完整的对象是不可能完成的。对于其他基于区域的方法，采用简单的算法检测复杂对象也是很困难的，这是因为这些基于区域的方法没有利用三维信息，比如深度图以及点云数据。对于图3-12中的复杂场景，我们提出的方法很明显的避免了背景造成的影响（例如白色的墙壁），并且能获得清晰的显著性图。对于其他的方法，背景抑制了前景对象的显著性，例如采用SF，A1以及CA得到的结果，或者产生高的显著性值去干扰前景对象的显著性，例如通过RC，GS，MR，SO以及A2得到的结果。

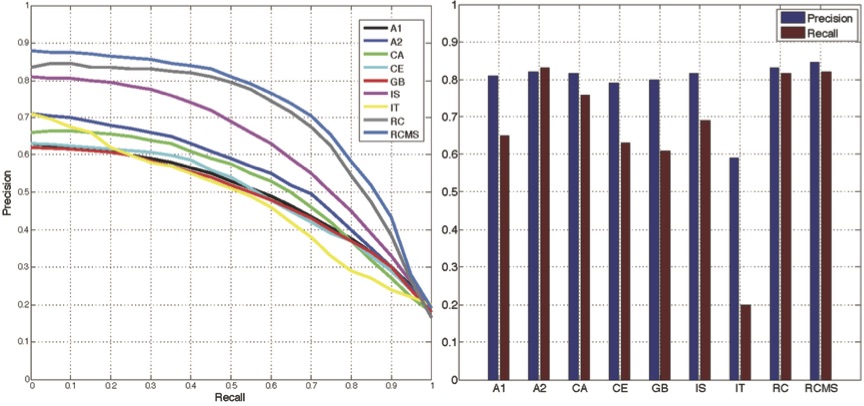


图3-10 提出的方法跟目前已有最好方法之间准确率和召回率的对比。

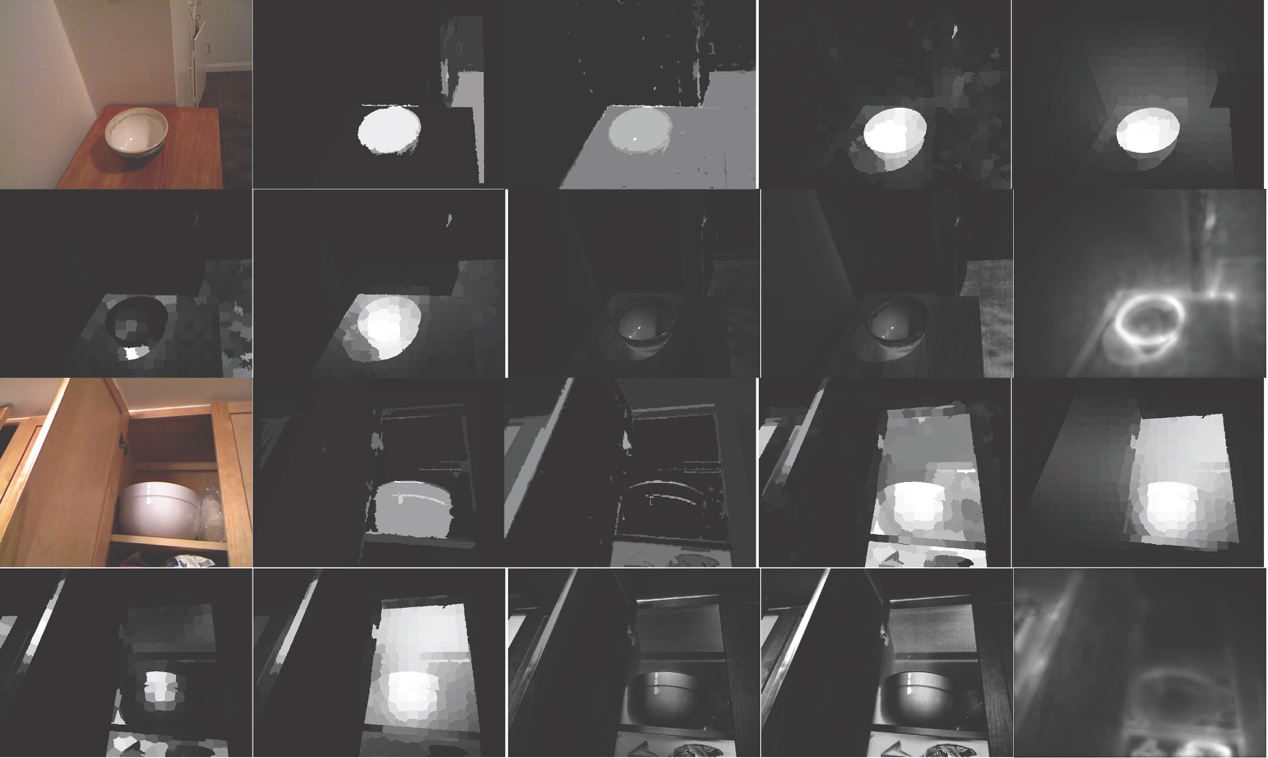


图3-11 从左到右分别是简单场景的原图以及采用MRS，RC，GS，MR，SF，SO，A1，A2

以及CA得到显著性图

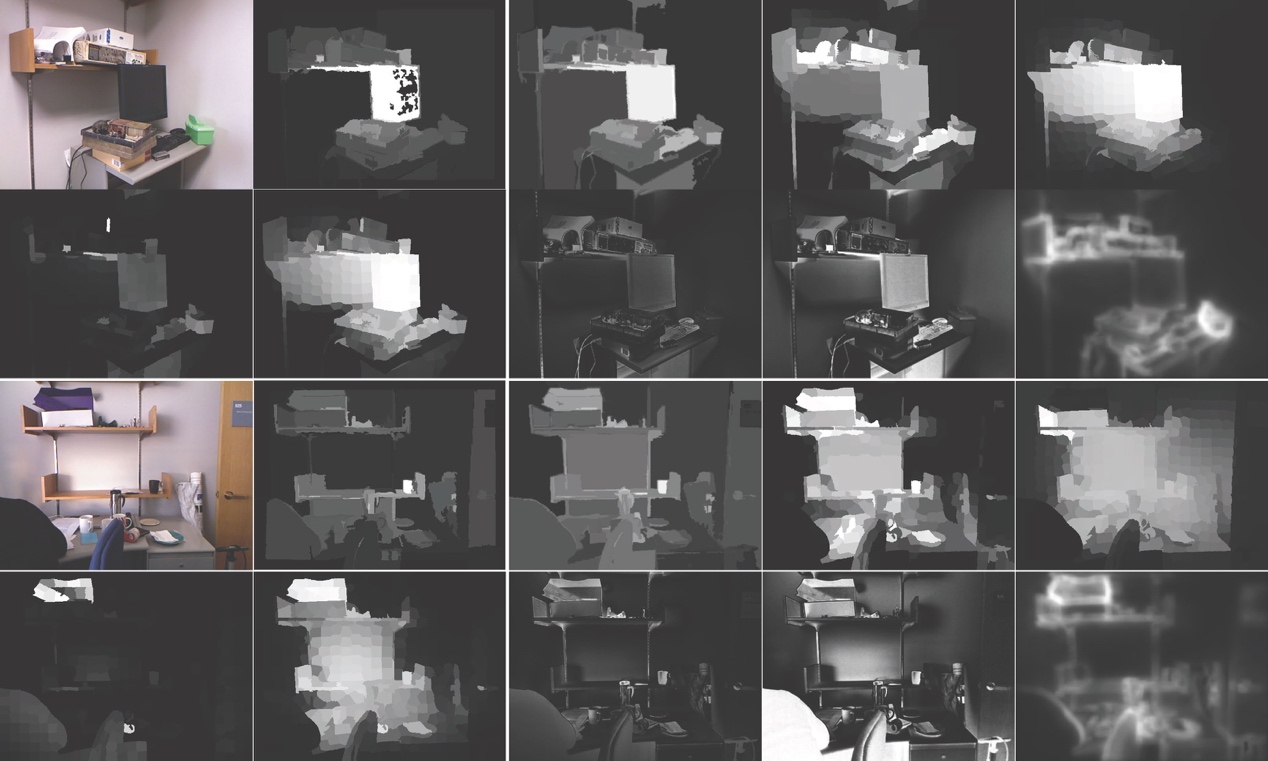


图3-12 从左到右分别是复杂场景的原图以及采用MRS，RC，GS，MR，SF，SO，A1，A2

以及CA得到显著性图

### 3.4.3 在真实机器人系统上评估实验

我们在真实机器人系统上对比我们提出的方法和其他方法（RC，GS，MR以及SO方法）。这些方法是跟本文方法最相关的。均是基于区域的方法。我们的方法采用跟RC方法相似的方式区计算区域对比度。GS，MR以及SO方法利用背景先验信息。

在实验中，我们将提出的方法以及其他对比的方法部署到我们实验室中的AS-R机器人上，采用文献[10]描述的显著性割方法去检测场景中的显著性对象。我们采用一个Kinect相机去获取彩色图像和深度图。对于RC，GS，MR以及SO方法，只采用了颜色图像。对于我们提出的多模态区域一致显著性方法MRS，同时用到了颜色图像和深度图像。由于我们的方法采用了延伸到场景的平面结构作为背景先验信息，它可以抵消背景的影响，并且因此获得更精确的对象检测结果。这些方法的准确度列在表3-2中以及实验结果如图3-13所示。很明显，我们的方法可以更高效的检测到感兴趣的对象。

表3-2 在真实机器人系统上检测到感兴趣对象的准确度

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | MRS | RC | GS | MR | SO |
| 准确度(%) | 75.2 | 69.7 | 68.5 | 71.2 | 71.4 |

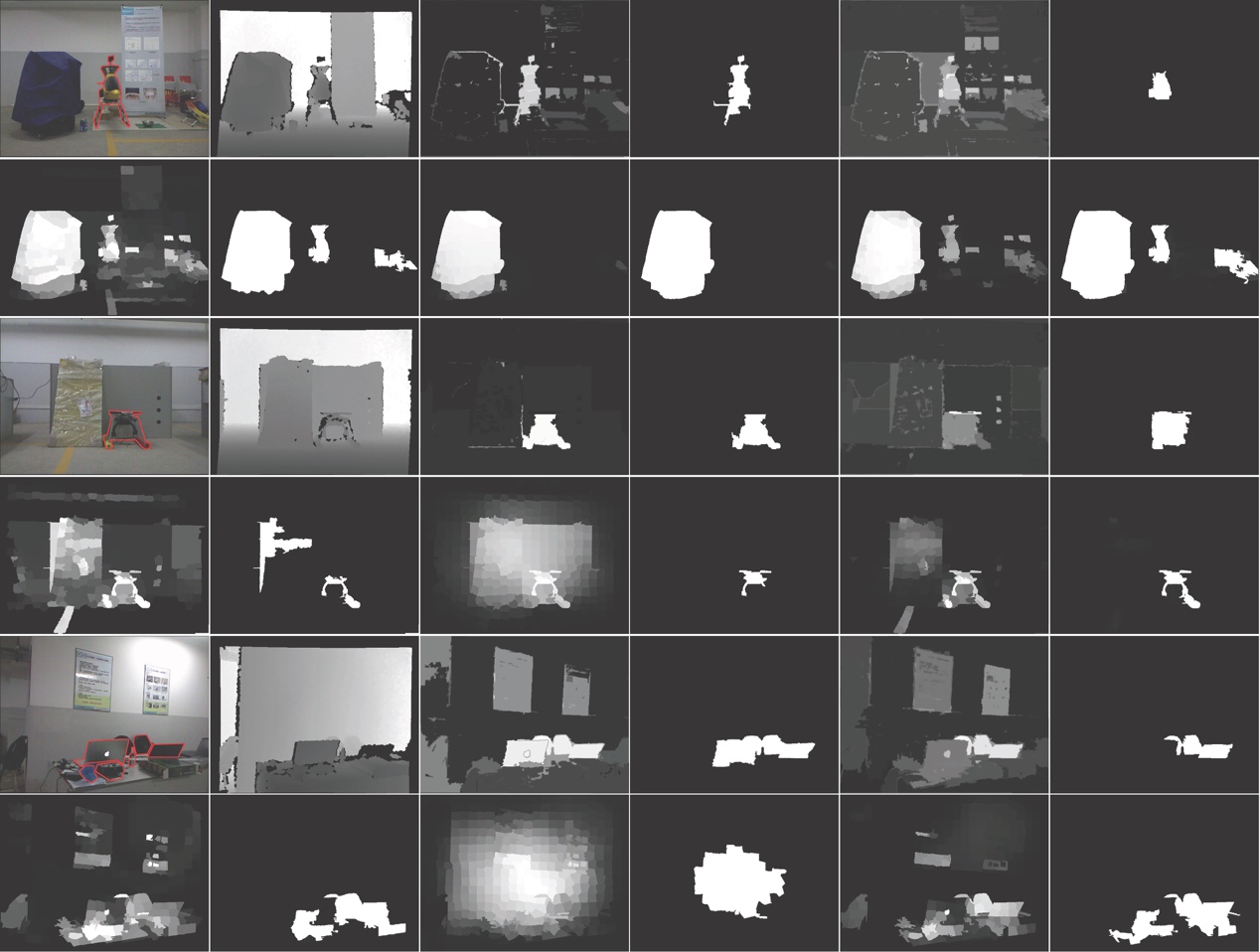


图3-13 在真实机器人系统上检测感兴趣对象的例子，从左到右分别是原图、

深度图以及采用MRS、RC、GS、MR、SO方法得到的显著性图

3.5 本章小结

本章提出了一种新的基于多模态数据计算区域一致显著性图的方法。提出的方法包含一种简单有效自适应多模态过分割方法去精确分割场景。基于这些得到的精确的过分割区域，结合全局对比度以及前景背景信息去计算显著性值。对比已有最好的方法，我们提出的方法获得了更好的性能。同时我们发现，在室内场景中平面结构可以用作背景先验去提升显著性计算的准确性，这是因为我们可以消除混乱场景中背景产生的影响。此外，通过检测焦点区域获得的前景先验也对计算显著性有正面的影响。我们提出的方法在有大量颜色图像和深度图像的公共数据集上进行了评估测试。实验结果表明我们提出的方法在准确率和召回率上比其他方法要好。

# 第？章 结论与展望

### ？.1 结论

简要回顾所做的工作，包括为什么要做这件研究工作，采用什么方法，做了哪些事，取得了哪些结果，是否有实验、仿真或实际应用，效果如何？对推进本学科发展有什么作用？要注意不要完全复制摘要，既有类似之处，但也不完全相同。

### ？.2 展望

可结合技术发展趋势，分析本文尚存在的问题，简要说明下一步可如何做以解决这些存在的问题，同时展望一下该方向的发展前景。

# 参 考 文 献（小二号黑体居中，字间空两格）

参考文献应按文中引用出现的顺序列全，附于文末。学位论文中列出的参考文献必须实事求是，论文中引用的必须列出，没引用的一律删去。

参考文献字体用5号，中文用宋体，英文用Times New Roman，行间距选1.4倍。

根据GB 3469 规定，以单字母方式标识各种参考文献类型：

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 参考文献  类型 | 专著 | 论文  集 | 析出  论文 | 报纸  文章 | 期刊  文章 | 学位  论文 | 报告 | 标  准 | 专  利 | 其他  文献 |
| 文献类型  标识 | M | C | A | N | J | D | R | S | P | Z |

对于数据库，计算机程序及光盘图书等电子文献类型的参考文献，以下列字母为标识:

|  |  |  |  |
| --- | --- | --- | --- |
| 参考文献类型 | 数据库 (网上) | 计算机程序 (磁盘) | 光盘图书 |
| 文献类型标识 | DB (DB/OL) | CP (CP/DK) | M/CD |

具体格式如下：

专著格式：

[序号] 著者. 书名[M], 版次(可选项). 出版地: 出版者, 出版年份.

中英文作者均采用姓在前，名在后的形式，例：

* 1. 殷剑宏. 吴开亚, 图论及其算法[M]. 合肥：中国科学技术大学出版社, 2004.
  2. 党跃武，谭祥金. 信息管理导论[M], 第二版. 北京：高等教育出版社, 2006.
  3. Pedrycz W, Gomide F. An Introduction to Fuzzy Sets [M]. MA: MIT Press, 1998.

期刊格式：

[序号] 作者. 题名[J]. 期刊名称, 出版年份，卷号(期号): 起止页码.

中英文作者均采用姓在前，名在后的形式。例：

1. 廖建新. 移动智能网技术的研发现状及未来发展[J]. 电子学报, 2003, 31(11): 1725-1731.
2. 冯永新, 刘治国, 姜月秋. 一个应用于移动自组网络管理的簇生成算法[J]. 软件学报, 2003, 14(1): 132-138.
3. V.Wong and V.Leung， Location management for next generation personal communication networks [J]. IEEE Network, 2000, 14(10): 18–24.
4. Tu N A, Dinh D L, Rasel M K, et al. Topic modeling and improvement of image representation for large-scale image retrieval[J]. Information Sciences, 2016, 366: 99-120.
5. Borji A, Sihite D N, Itti L. An Object-Based Bayesian Framework for Top-Down Visual Attention[J]. AAAI. 2012; pp 1529-1535.
6. Perazzi, Federico. Saliency filters: Contrast based filtering for salient region detection[J]. Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012: 733-740
7. Goferman, Stas, Lihi Z, and Ayellet T. Context-aware saliency detection[J]. IEEE transactions on pattern analysis and machine intelligence 34.10 (2012): 1915-1926.

论文集格式：

[序号] 作者. 题名[A]. 编者(可选项). 文集名[C]. 出版者地点: 出版者, 出版年份, 起止页码(可选项).

例：

1. 赵秀珍. 关于计算机学科中几个量和单位用法的建议[A]. 中国高等学校自然科学学报研究会编. 科技编辑学论文集[C]. 北京: 北京师范大学出版社, 1997, 125-129.
2. Bao F, Deng R H, Mao W. Efficient and practical fail exchange protocols with off-line TTP [A]. Proc of the 1998 IEEE Symposium on Security and Privacy [C]. Oakland: IEEE Computer Press, 1998, 77-85.

专利格式：

[序号] 专利申请者. 专利名称[P]. 国别: 专利号, 发布日期.

例：

1. 姜锡洲. 一种热温外敷药制备方案[P]. 中国专利: 881056073, 1989-07-26.

报纸格式：

[序号] 作者. 题名[N]. 报纸名, 年-月-日（版次）

例：

1. 国务院新闻办公室. 中国的粮食问题[N]. 人民日报, 1996-10-25(2)

网络文献格式：

[序号] 作者. 电子文献题名[电子文献及载体类型标识]（任选）. 电子文献的出处或可获得地址, 发表或更新日期/引用日期（任选）

例：

1. 王明亮. 关于中国学术期刊标准化数据库系统工程进展[EB/OL].

http://www.cajcd.edc.cn/pub/wml.txt/980810-2.html, 1998-08-16/1998-10-04

标准：

[序号] 起草责任者(可略). 标准代号 标准顺序号 发布年, 标准名称[S]，出版地(可略)，出版者(可略)，出版年（可略）

例：

1. 全国量和单位标准化技术委员会. GB 3100～3102-93 量和单位. 北京: 中国标准出版社, 1997
2. GB/T 16159, 汉语拼音正词法基本规则[S].

# 

# 附 录（小二号黑体居中）

主要列入正文内过分冗长的公式推导；以便查读所需的辅助性数学工具或表格；重复性数据图表、程序全文及说明。

论文的附录依次为附录1，附录2等。附录中的图和公式另编排序号，与正文分开。

# 

# 致 谢（小二号黑体居中）

对给予各类指导、支持、和协助完成研究工作，以及提供各种条件的单位及个人表示感谢。致谢应实事求是，切忌浮夸与庸俗之词。

# 

# 攻读学位期间参加的科研项目和成果

**参加的科研项目**

[1] 国家自然科学基金：名称（项目编号），

**录用和发表的论文**

[1] 按论文发表的时间顺序，列齐本人在攻读学位期间发表或已录用的学术论文清单（格式参照参考文献的写法），作者单位必须是浙江工业大学。对论文的盲审版本，须删去作者名。

[2] 投稿的论文不在此列。

**专利**

[1]