

学校编码: 10384

学号: 200428049

分类号 _____ 密级 _____

UDC _____

厦 门 大 学

硕 士 学 位 论 文

基于智能人机交互的批量图像主动分割算法研究

Active Segmentation for Groups of Images

吴琦颖

指导教师姓名: 李翠华 教授

杨 琼 博士 (微软亚洲研究院)

专 业 名 称: 计算机应用技术

论文提交日期: 2007 年 5 月 日

论文答辩时间: 2007 年 6 月 日

学位授予日期: 2007 年 月 日

答辩委员会主席: _____

评 阅 人: _____

2007年 月

厦门大学学位论文原创性声明

兹提交的学位论文，是本人在导师指导下独立完成的研究成果。本人在论文写作中参考的其他个人或集体的研究成果，均在文中以明确方式标明。本人依法享有和承担由此论文产生的权利和责任。

声明人（签名）：

年 月 日

厦门大学博硕士学位论文摘要库

摘要

基于语义的图像分割问题是实现图像个性化编辑、编解码、个性化检索以及图像理解等应用的关键问题。数码图像采集设备的普及，导致了海量图像数据的生成，人们对于图像分割的需求，也从原先单一图像正逐渐拓展为批量图像。但是如果使用以往的方法实现批量图像的有效分割，人机交互的成本非常昂贵。

本文提出了一个全新的主动分割框架用以解决批量图像的分割问题。该框架努力地挖掘人机“合作”分割的有效方式，期望用尽可能少的人机交互达到尽可能好的分割效果，主要通过如下三个方面实现挖掘工作：

1. 让用户进行手动分割或是修正的图像都是经过主动挑选的典型图像；
2. 用户提供的有限的语义信息，被用来主动地挖掘出图像之间的关联关系，并指导相关图像的自动分割；
3. 用户对代表性图像分割结果的修正信息，被用于主动地修正其他相似图像的分割结果。

具体地，本文的主要工作和创新点有：

1. 提出了层级的方式用于组织批量图像。该方式是建立在相似图像分组的基础上、以图像组为单位的批量图像表示方式。特别地，用融合的特征词袋模型实现相似度的衡量，提高鲁棒性和精确性。
2. 提出了基于前/背景分类的分割语义传播算法。在目标和场景分类的基础上，以图像组为单位实现语义在相关联图像间的传播。特别地，在前/背景类别部分已知的情况下，用迭代的graph-cut分割算法实现前景和背景语义信息的估计。
3. 提出了一个基于样例的分割错误修正算法。通过在相似的分割结果中逐步修正前景及背景，达到图像的主动修正目的。

实验结果表明，本文提出的主动分割框架能够在保证批量图像分割质量的前提下，显著地减少用户的交互量，特别地，批量图像之间的关联性越强，智能人机交互的作用越显著。

关键词：图像分割；目标/场景分类；主动学习。

厦门大学博硕士论文摘要库

Abstract

Semantic segmentation plays a very important role in many applications, including image editing, coding, retrieval and understanding. Meanwhile, a marvelous demand for semantic segmentation of groups of images is growing in both fashionable people and researchers, since digital camera and videos are becoming so popular in the modern world, which generates a lot of photos and videos. Unfortunately, a lot of user interaction efforts are needed for obtaining good results in existing works.

In this paper, we propose an active segmentation framework for groups of images. This framework aims to release user's burden for achieving satisfactory segmentation results by making great effort to provide the most effective collaboration method between user and machine.

There're three main ways for exploring the collaboration method:

1. Typical images are actively selected as a query to request the user for labeling semantics and refining errors;
2. The semantic information is propagated actively for the segmentation of related images;
3. The refinement information is propagated for similar images actively.

Furthermore, the main contributions of this thesis include:

1. A hierarchical structure is proposed to organize those groups of images. Similar images are grouped at first, and the groups of similar images are designed as the high level description of all the images. Specifically, combined features are modeled as bag of words to enhance robustness and distinction.
2. A semantic propagation algorithm based on foreground and background classification is proposed. The system implements the semantic propagation based on object and scene classification techniques. Specifically, when the segmentation class information is partially known, the semantic information is estimated by iterative graph-cut algorithm.
3. An example based segmentation refinement algorithm is proposed to implement

the error refinement in similar images. Typically, the refinement is done iteratively.

Experimental results show that the active segmentation framework can significantly reduce the user interaction efforts and reserve the segmentation quality in groups of images. Especially, if the images are highly related, the results are much better.

Keywords: Image Segmentation; Object/Scene Classification; Active Learning.

厦门大学博硕士学位论文摘要库

目 录

第一章 绪 论	1
1.1 研究背景和意义	1
1.2 图像分割发展现状	3
1.3 课题的研究目标	1
1.4 论文结构安排	7
第二章 主动分割的总体研究框架	9
2.1 主动学习机制	9
2.2 框架的设计依据	10
2.2.1 主要观察	10
2.2.2 自顶向下图的描述及分析	12
2.3 总体研究框架	14
2.4 本章小节	18
第三章 主动询问语义信息	19
3.1 词袋(Bag of words)模型	19
3.1.1 Bag of words 思想	19
3.1.2 词典建立的一般步骤	20
3.2 融合的图像特征词典	22
3.3 相似图像的分组	25
3.3.1 基于融合特征的图像分组	25
3.3.2 实验结果	28
3.4 典型语义的主动询问	29
3.4.1 学习规则	29
3.4.2 实验结果	31
3.5 本章小结	32
第四章 语义信息的主动传播	33
4.1 目标/场景的分类	33
4.2 基于语义信息的图像组分类	34

4.3 前/背景分割问题的求解	39
4.4 语义信息在相关联图像间的传播	41
4.5 实验结果	45
4.6 本章小结	47
第五章 分割错误的主动修正及增量迭代	49
5.1 分割模型的主动学习机	50
5.1.1 学习规则	50
5.1.2 实验结果	52
5.2 分割错误的主动修正	53
5.2.1 错误修正信息的传播	53
5.2.2 传播后的实验结果	56
5.2.3 区域级的自动分割错误修正	57
5.3 增量迭代过程	58
5.3.1 增量迭代机制	58
5.3.2 实验结果	59
5.4 本章小结	61
第六章 系统的可用性分析	63
6.1 用户交互数据分析	63
6.2 图像分割质量分析	64
6.3 鲁棒性分析	65
第七章 结束语	67
7.1 主要工作总结	67
7.2 未来工作展望	68
7.3 小结	69
研究生期间发表论文及参与项目	71
致 谢	73
[参考文献]	75

Contents

Chapter 1 Introduction	1
1.1 Background and Significance	1
1.2 Status of Image Segmentation.....	3
1.3 Research Objective	1
1.4 Arrangement of the Thesis	7
Chapter 2 Overview of the Active Segmentation Framework	9
2.1 Active Learning	9
2.2 Design of the framework	10
2.2.1 Key Observation	10
2.2.2 Top-down Graph Description and Analysis	12
2.3 Overview of the framework	14
2.4 Summary.....	18
Chapter 3 Active Query for Semantics.....	19
3.1 Bag of words Model	19
3.1.1 Idea of Bag of words.....	19
3.1.2 Methods for Codebook Construction.....	20
3.2 Codebook of Combined Image Feature	22
3.3 Similar Image Grouping.....	25
3.3.1 Combined Feature based Image Grouping	25
3.3.2 Experimental Results	28
3.4 Active Query for Typical Semantics	29
3.4.1 Active Query Rule.....	29
3.4.2 Experimental Results	31
3.5 Summary.....	32
Chapter 4 Active Propagation of Semantics	33
4.1 Object/Scene Classification.....	33
4.2 Semantic Classification based Image Grouping.....	34

4.3 Problem of Foreground/Background Segmentation and its solution...	39
4.4 Semantic Propagation.....	41
4.5 Experimental Results.....	45
4.6 Summary.....	47
Chapter 5 Active Refinement and Incremental Iteration.....	49
5.1 Active Learner of Segmentation Model	50
5.1.1 Active Query Rule.....	50
5.1.2 Experimental Results	52
5.2 Active Refinement of Segmentation Error	53
5.2.1 Propagation of Error Refinement.....	53
5.2.2 Experimental Results after Propagation.....	56
5.2.3 Regional Automatic Error Refinement	57
5.3 Incremental Iteration.....	58
5.3.1 Incremental Iteration Strategy.....	58
5.3.2 Experimental Results	59
5.4 Summary.....	61
Chapter 6 Usability Analysis	63
6.1 Analysis of User Interaction.....	63
6.2 Analysis of Segmentation Quality.....	64
6.3 Robustness Test	65
Chapter 7 Conclusions and Future Works.....	67
7.1 Conclusions.....	67
7.2 Future Works.....	68
7.3 Brief Summary	69
Publications and Projects	71
Acknowledgements.....	73
[Reference]	75

第一章 绪论

图像分割是图像处理与计算机视觉领域的一个经典问题，它在图像编辑、内容分析、编解码以及基于内容的图像检索等领域都有着非常广泛的应用。在几十年的研究发展过程中，图像分割研究工作一直在不断地创新，其技术也渐渐地走向成熟实用化。另一方面，当今信息时代，图像采集设备的大量普及，导致了海量图像数据的生成，分割的图像范围也从原先的单一图像向批量图像拓展。为了解决批量图像的分割问题，本文拟提出一套主动图像分割框架，挖掘出人机“合作”分割的有效方式，以期用尽可能少的人工操作达到尽可能好的分割效果。本章主要对主动图像分割的研究背景、意义、研究重点以及全文的结构安排作概要介绍。

1.1 研究背景和意义



图1.1 普遍的海量图像

当今，数字照相机和摄像机的造价日益低廉，数以万计的家庭拥有了照相机、摄像机等数字设备。与此同时，数字存储卡的容量也在日益增大，以1G存储卡为例，可存储1千多张700万像素的图像。一次朋友聚会，一次出行旅游，数百张图像往往随之而生，如图1.1所示。

然而，不断地按下快门，捕捉精彩瞬间，仅仅是当今时尚人士兴趣的一小部分而已。拥有了海量的图像，富有创造性的人们，往往更感兴趣地是如何能够有效地管理图像，在图像上做出个性化的处理。譬如，在一次旅行之后，人们常常期望以旅行照片为素材，制作一份精美的合成相册，为难忘的旅行再添一笔美好的色彩。又譬如，拍摄技巧和设备条件欠缺的人们，常常希望通过后处理，批量地对前背景自定义风格化出专业水准的图像。再比如说，从事目标识别或是机器学习的研究人员是多么地渴望从大量的图像中获取分割好的目标样本。更别提，对视频内容分析和合成特效感兴趣的人们，是多么迫切地希望能够实现视频目标的有效分割。

人们对于分割对象的定义常常是个性化的，用户的需求千差万别。人们的个性化需求，使得在任务开始之前，由用户指定分割的语义信息变得不可避免，有人形象地将之称为抠图（cutout）。

要实现大量图像的自定义分割，常规的做法是一张一张的“抠”过去。而即使使用含有先进抠图技术的图像图形编辑工具Microsoft Expression Design^[1]进行抠图，一个熟练的抠图人员完整地抠一张图像平均需花费30秒，而抠一百张图片则需要花费将近一个小时的时间，如此大量的交互代价让普通人望而生叹。

进一步地，在这些海量的图像中，存在着大量彼此关联的图像。这种相互关联关系常常表现为：(1)前景相似（比如同一个人在很多景点拍的照片）；(2)背景相似（比如朋友们轮流在同一个漂亮的景点留下的纪念）；(3)或者是前景和背景都很相似（比如小孩子在花园前的多个姿势、多个表情）。这些相互的关联关系使得用户在一张图像上所作的某一抠图操作，往往可能要在好几张图像上重复。这些大量的高重复性操作，极大地挑战着用户的耐心。

如果能够让人们从这种大量高重复性的交互中解脱出来，无疑将更大地激发人们的创造欲望，集中精力在更有价值的工作上。

直观地想，如果能够让机器主动地去分析所有待分割图像，从中挑选出具有

代表性的图像，让用户进行分割语义信息的标注，并用这些交互信息指导其他相关联图像的分割，那么，分割的自动化程度将有希望得到提高。

这里我们称机器主动分析待分割图像、并主动向用户询问语义信息，从而达到批量图像的分割过程为**主动分割**。主动分割的思想，是在新的分割应用需求背景下提出的全新思想。

设计实现这个思想的框架，是对以往图像分割研究的一种提升。单一图像的分割、两两关联图像的分割以及一组关联图像的分割课题，都成为了这个框架的子问题。

进一步地，这个框架的设计虽然源于图像分割需求，却不仅仅局限于图像分割技术。如何分析图像的组成，如何确定图像之间的关联关系，如何分析分割错误并将错误修正信息主动传播，也是这一框架中的关键技术，而这些又将涉及到图像特征表示、物体的描述、目标/场景分类及主动式机器学习等众多图像处理、计算机视觉和机器学习领域的经典问题。

实质上，主动分割的思想，就是一种人机最佳合作分割方式的探寻。我们知道，如果把分割工作完全交给人来手动完成，人们太伤神费力，而如果要想让计算机实现全自动的分割也很困难，因为目前还没有一个算法可以做到对普适的图像进行满意的分割，更何况计算机无从了解当前的用户想要割的是什么。是否能够将计算机的高效计算能力和人们的主观能动性结合起来，达到一个效率和质量的优化平衡呢？主动分割框架正是应这样的想法而提出的，其核心内容是让计算机主动地去分析批量图像的分割任务，不懂的时候向人们询问，由人们适时地提供个性化的前/背景分割信息，尽可能地将有限的信息加以传播扩散，从而实现自动的分割以及修正，将上述的过程加以反复，直至所有的图像都实现了有效的分割。

1.2 图像分割发展现状

图像分割一直以来都是数字图像处理领域的一大基本问题。在几十年的研究过程中，图像分割技术经历了从一开始的自底向上方法，到自顶向下方法，再到现在广泛使用的自底向上和自顶向下相结合的方法的演变。

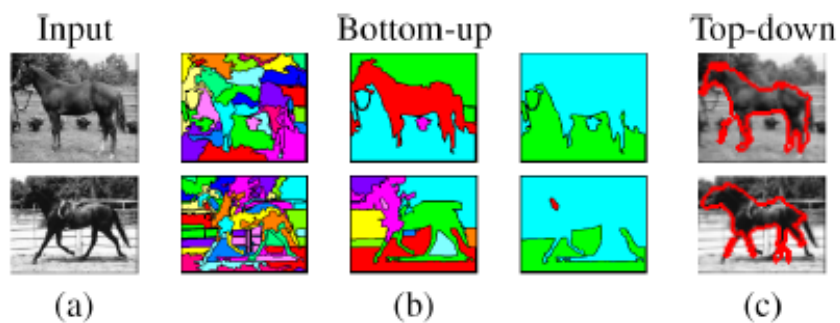


图1.2 (a)输入. (b)在三个不同分辨率下的自底向上分割结果. (c)自顶向下分割结果.^[2]



图1.3 (a)Lazy-snapping交互方式与分割结果^[19]. (b)Grab-cut交互方式与分割结果^[20].



图1.4 (a)原视频截图. (b)video object cutout分割出来的视频对象^[22].

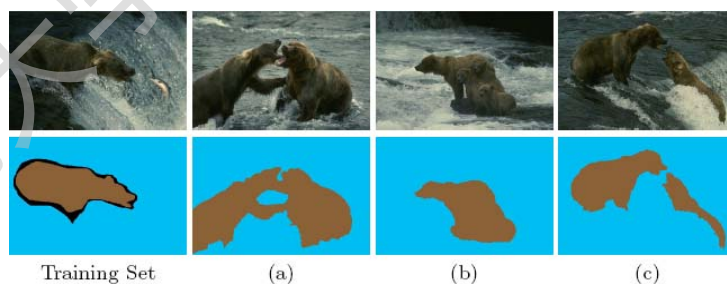


图1.5 基于样例的图像分割^[23].



图1.6 输入前景相同的图像对, 同步分割输出的分割结果^[24].

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库