

学校编码: 10384

分类号 _____ 密级 _____

学 号: 24320141152403

UDC _____

厦 门 大 学

硕 士 学 位 论 文

高质量图像的渐进式并行检索算法研究

Research on Progressive and Parallel Retrieval Algorithm for
High Quality Images

张光裕

指导教师: 林俊聪 副教授

专业名称: 计算机科学与技术

论文提交日期: 2017 年 4 月

论文答辩日期: 2017 年 5 月

学位授予日期: 201 年 6 月

指导教师: _____

答辩委员会主席: _____

2017 年 4 月

厦门大学学位论文原创性声明

本人呈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其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

- () 1.经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。
- () 2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

摘 要

由于图像采集技术的多样性和普及性,数据驱动的图片分析和编辑方法变得越来越热门。然而图像数据的爆炸性增长也带来了许多挑战,其中之一就是如何帮助人们快速有效地从众多素材中检索出所思所想的图像。对于一个图像检索方法而言,首先要解决的一个问题是提供什么样的线索从而返回匹配的图像。单一的线索往往很难完整地表达用户的意图,因此需要多种线索组合。其次基于内容的匹配过程耗时较长,在海量图像数据情况下,多线索组合的模式在实时交互方面无疑是一个很大的挑战。另一方面,图像检索的目的往往是为了进一步的图像分析、编辑以及数字内容创作等提供素材。目前的检索方法大都不能保证检索结果的图像质量,而这却是许多图像分析和编辑方法所要求的。

本文针对目前检索方法无法保证检索结果的图像质量以及面向海量图像数据的特征匹配耗时的问题,提出了一种高质量图像的渐进式并行检索算法。候选图像根据一系列检索线索被渐进地排除:首先我们根据图像标签进行快速的文本比较,从海量图像库中获取待匹配图像;然后计算每张图像的显著性值,进一步得到高显著性区域,并过滤掉显著性得分低的对象;接着对这些显著性对象从三分点、对角线准则、视觉平衡三个美学标准来评价,过滤掉不符合美学标准阈值的部分;接下来计算每个显著性对象的边界清晰度并过滤掉清晰度低的对象;对剩余的对象扩张周围区域并计算此区域的分块数量,数量越大表示复杂度越高,然后过滤掉复杂度过高的图像;对于每个剩余的图像,首先对高显著性区域进行形态学扩张,并以此区域为指导将场景元素分割出来;最后我们衡量用户绘制的轮廓和场景元素轮廓之间的一致性,并将最符合用户意图的结果集返回。整个算法过程涉及到多个基于图像特征内容的计算和匹配步骤,实时性将成为一个关键问题。因此我们将整个框架通过 MapReduce 来实现并行化,以提高检索性能。本文通过各种实验来验证算法的检索性能、准确性以及检索结果的图像质量。并且我们实现一个图像合成原型系统来展示本文提出的算法在数字内容创作等方面的应用前景。

关键词: 图像检索; 高质量; 渐进

Abstract

Thanks to the diversity and popularity of image acquisition technique, data driven methods for image analysis and edit become popular. However, the explosive growth of images also presents challenges. To help users retrieve expected images quickly and effectively is one of the most difficult. The retrieve cue is the first problem when designing an image retrieval method. It is difficult for a single cue to express the user's intention completely, thus combining multiple cues becomes necessary. Another problem is that the content-based matching process takes a long time. In the case of massive image data, multi cue model is undoubtedly a great challenge for real-time interaction. On the other hand, the purpose of image retrieval is usually to provide material for further image analysis, editing and digital content creation. However, most of the current retrieval methods cannot guarantee the quality of the image which is usually required for many image analysis and editing tasks.

In this dissertation, we propose a progressive retrieval method for high quality images, targeting on the problems that the existed retrieval methods cannot guarantee the image quality and the time consuming of the feature matching for the massive material. Images are progressively filtered out according to a series of criteria: we firstly pick up candidates from the huge image database according to a quick textual comparison with images' tags; then we calculate the high saliency objects of each image, and filter out those objects with low saliency scores; for each candidate image, we evaluate an aesthetic score from the aesthetic standards of the three points, diagonal rule and visual balance, and filter out the part that does not meet the threshold of aesthetic standards; next, we calculate the boundary clarity of each salient object and filter out the object with low clarity; for the remaining objects, we expand the surrounding area and count the number of segment in this area by a graph cut method; The greater the number, the higher the complexity, and then filtering out the image that is too complex; for each of the remaining images, we firstly expand the

high saliency region by morphological dilation and then cut the scene item according this expanded region. Finally, we measure the consistency between the user drawn contour and the scene item contour, and return the result set that best matches the user's intent. The whole algorithm involves a lot of calculation and matching steps based on the image feature, so performance will become a key issue. We parallelize the whole framework in a MapReduce framework to further improve the performance. We validate the performance, accuracy of the algorithm and quality of retrieval result in various experiments. We also demonstrate the potential of our technique with an image synthesis prototype system.

Key Words: Image Retrieval; High Quality; Progressive

目录

第一章 绪论	1
1.1 研究背景和意义	1
1.1.1 研究背景.....	1
1.1.2 研究意义.....	3
1.2 国内外研究现状	3
1.3 论文研究内容	5
1.4 论文组织结构	6
第二章 相关技术及理论介绍	8
2.1 基于内容的图像检索	8
2.2 基于图表示的图像分割	10
2.3 图像美学构图理论	13
2.4 MapReduce 框架	16
2.4.1 总体架构.....	17
2.4.2 执行过程.....	19
2.5 本章小结	21
第三章 算法框架设计	22
3.1 总体流程设计	22
3.2 高质量候选图像的过滤	25
3.3 前景对象提取	27
3.4 轮廓匹配	30
3.5 本章小结	33
第四章 渐进的图像内容过滤	34
4.1 渐进过滤策略	34
4.2 对象清晰度	35
4.3 对象可分离度	39

4.4 本章小结	41
第五章 基于 MapReduce 框架的并行实现	42
5.1 总体设计	42
5.2 数据准备	44
5.3 候选图像收集	46
5.4 候选图像渐进过滤	47
5.5 本章小结	49
第六章 实验过程及结果分析	50
6.1 整体实验方案与环境	50
6.2 检索性能分析	51
6.3 检索结果质量分析	53
6.4 本章小结	57
第七章 总结与展望	58
7.1 总结	58
7.2 展望	59
参考文献	60
攻读硕士研究生期间发表的学术论文	64
致 谢	65

Contents

Chapter 1 Preface.....	1
1.1 Background and Significance	1
1.1.1 Background.....	1
1.1.2 Significance.....	3
1.2 Research Status	3
1.3 Research Content	5
1.4 Structure of Paper.....	6
Chapter 2 Introduction to Related Technology and Theory.....	8
2.1 Content Based Image Retrieval.....	8
2.2 Graph-based Image Segmentation	10
2.3 Image Aesthetic Composition Theory	13
2.4 MapReduce Framework.....	16
2.4.1 Overall Architecture.....	17
2.4.2 Execution Procedure	19
2.5 Summary.....	21
Chapter 3 Algorithm Framework Design	22
3.1 Overall Design	22
3.2 High Quality Candidate Image Filtering	25
3.3 Foreground Object Extraction	27
3.4 Contour Matching.....	30
3.5 Summary.....	33
Chapter 4 Progressive Image Content Filtering	34
4.1 Strategy of Progressive Filtering	34
4.2 Object Clarity.....	35
4.3 Object Separability	39
4.4 Summary.....	41
Chapter 5 Parallel Implementation Based on MapReduce	

Framework	42
5.1 Overall Design	42
5.2 Data Preparation.....	44
5.3 Candidate Image Collection	46
5.4 Candidate Image Progressive Filtering.....	47
5.5 Summary.....	49
Chapter 6 Experiment Procedure and Results Analysis	49
6.1 Overall Experiment Schema and Environment	50
6.2 Retrieval Performance	50
6.3 Quality Evaluation of Retrieval Results	51
6.4 Summary.....	53
Chapter 7 Conclusions and Prospects.....	58
7.1 Conclusions.....	58
7.2 Prospects	59
References.....	60
Papers Published During Postgraduate.....	64
Acknowledgements	65

第一章 绪论

1.1 研究背景和意义

1.1.1 研究背景

随着科技进步，图像获取设备日益普及。除了相机外，手机、平板电脑等随身设备也都具备图像获取能力，这极大的方便了人们去获取各种各样的图像信息以记录经历，表达观点。有调查表明，当今我们获取的社会信息中，60%到70%是以图像的方式，我们已经置身于“读图时代”，对图像的解读成为主流阅读方式之一。图像采集的方便带来了新的机遇。数据驱动的图像分析和编辑方法变得越来越热门，并在许多传统的问题上取得了相当不错的进展。比如，我们可以很容易地将来自不同源图像的部分内容组合成新的图像，如图 1-1 所示。用户通过绘制线条草图和添加标签的方式，在图像数据库中搜索出期望的图像，并将多个图像中的目标元素合成在一起以更好的表达自己的想法，同时也达到快速创作的目的，降低创作门槛。

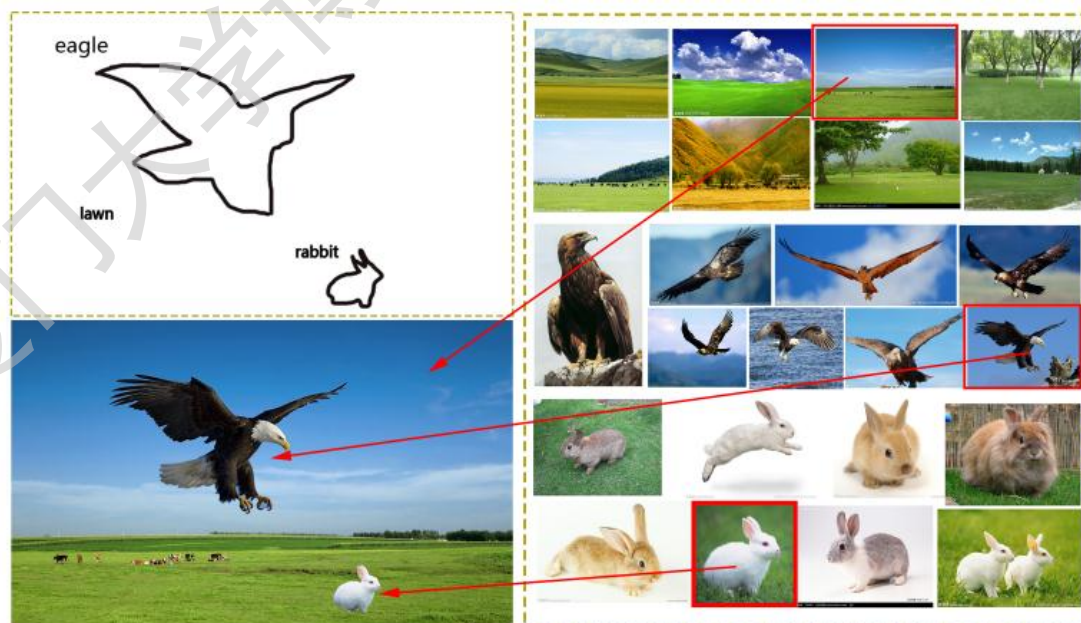


图 1-1. 合成来自不同图像的对象

然而，图像数据的爆炸性增长也带来了许多挑战，其中之一就是如何帮助人们快速有效地从海量素材数据库中检索出期望的图像。这也正是图 1 中所展示的数字内容创作应用中的关键问题。对于海量的图像素材而言，直接浏览的方式显然不切实际，因此一种有效的图像检索方法是人们所迫切渴望的。准确性、有益性和快速性是一个图像检索系统的关键要求。最初的方法是首先为图像添加文本标签，用来描述这幅图的内容，而后使用文本检索方法实现图像的检索。由于图像数据库的容量越来越大，我们很难为图像提供一个合适并且完备的标签集，并且标注的人工成本大且难以进行后续的管理。因此人们提出一种新的图像检索技术，基于内容的图像检索（Content-Based Image Retrieval，以下简称 CBIR）。这种技术利用图像特征来作为检索线索，不需要人工干预，简单、准确。基于内容的图像检索首先必须解决的一个问题是：提供什么样的线索从而让检索系统返回匹配的图像。单一的线索往往很难完整地表达用户的意图，因此需要多种线索组合。另一个问题在于基于内容的匹配过程耗时较长。对于动辄成千上万张的待匹配图像来说，即使是单一的检索线索，无疑都是一个很大的挑战，更不用说多线索组合的情况。多线索的检索方法可以有效地提高准确率，但是这需要消耗额外的时间来匹配每一种线索。在面向海量图像数据库的情况下，性能是另一个不容忽视的问题，特别是对于许多交互式应用，低检索实时性将导致可用性与用户体验度降低。并且图像数据库由于其容量之大，一般是分布式存储在多台机器上，传统的集中式的处理机制将导致巨大的数据传输，从而可能使得网络崩溃或者是很高的传输时间代价。因此，由于精度和效率等方面的因素制约，检索的实用性不尽如人意，基于内容的图像检索离实际应用仍然有一段距离。另一方面，图像检索的目的往往是为了进一步的图像分析、编辑以及数字内容创作等提供素材。目前的检索方法大都不能保证检索结果的图像质量，而这却是许多图像分析和编辑方法所必须的。许多应用对输入的图像数据的质量有一定的要求。比如在图像合成应用中，检索到的对象可能太小以至于无法在新的合成图像中作为目标元素，又或者检索到的对象周围的背景可能太复杂，导致无法从源图像中分离出来。因此，如何在海量图像素材库中快速有效地检索出用户所思所想的素材，并且保证检索结果的图像质量以便用于进一步的数字内容创作是一个迫切需要研究的方

向。

1.1.2 研究意义

基于内容的图像检索是计算机视觉领域下一个非常具有研究意义的分支,主要关注于大规模的数字内容。本文针对目前检索方法无法保证检索结果的图像质量以及面向海量素材的检索效率低的问题,研究一种不仅能够保证检索结果的图像质量,并且检索实时性高的图像检索方法。主要研究如何通过一些科学的质量标准有效地过滤掉图像质量低的候选图像,确定行之有效的线索以达到准确检索的目的。同时,为了解决多线索检索中特征匹配非常耗时问题,研究如何通过并行的方式进行性能优化。其研究意义体现在这些方面:

(1) 科学研究

本文涉及到的一些技术如基于内容的图像检索、图像质量评估等都是当前学术热点,在图像质量评价、检索匹配度、检索效率等方面仍有待改进,有必要进行专门的探讨和研究。

(2) 工程应用

本研究提出的方法可以为众多的图像分析、编辑以及数字内容创作应用提供高品质的图像素材,并且能够保证良好的检索反馈实时性,使得这些应用具有较高的用户体验。本方法内嵌于这些应用中,为用户提供符合他们所思所想的高品质图像素材,以便他们再进一步的创作,很大程度地降低图像创作的门槛,让普通用户也可以在“读图时代”直观、自如的表达出他们内心的想法,让科技走入生活,改变生活。并且本方法的操作过程非常简单,用户只需提供欲检索的物体的大致轮廓以及对应的标签,即可为用户呈现期望的结果,故也能很大程度上提高平面设计的工作效率,快速可视化灵感,间接的促进文化创意产业发展。

1.2 国内外研究现状

图像可以表达丰富的信息,因此常作为多媒体信息的载体,在众多领域得到广泛的应用^[1]。由于图像数据的非结构化性以及爆炸性增长,如何准确并且快速地从海量的图像素材库中检索出用户所思所想的图像,成为了一个难题,受到了

许多学者的关注。准确性、有益性和快速性是一个图像检索系统的关键要求。最初的方法是首先为图像添加文本标签，用来描述这幅图的内容，而后使用文本检索方法实现图像的检索。基于文本的检索方法主要的优点是容易理解，简单并且检索速度快。但是缺点也同样明显：（1）需要为每一张图像添加描述的文本标签，因此需要较多的人力，尤其是针对海量的图像素材；（2）人工添加图像标签带有较强的主观性，不同的人对图像的认知不同，且后期的更新和管理不方便；（3）图像所表达的信息丰富，人们很难添加完备的标签以完整地表达图像内容；（4）由于不同语言和文化的影响，对图像的标签很难统一。因此人们提出一种新的图像检索技术，即基于内容的图像检索。这种技术利用图像特征作为检索线索，不需要人工干预，简单、准确。基于内容的图像检索首先必须解决的一个问题是：提供什么样的线索从而让检索系统返回匹配的图像。最初的研究通常采用以图搜图的方式，即提供一张样图，让系统返回相似的图片^[2-5]。尽管想法很简单明了，但却不太切合实际，因为很多时候用户正是手上没有合适的图片才需要去检索，这本质上是个“鸡生蛋，蛋生鸡”的问题。随后，研究人员尝试根据用户手绘的线条来检索满足他们需要的图片^[6-8]。这种方法极大的方便了用户描述他们想要的图片内容，近年来颇为流行。虽然基于草图的方法让用户不必关心精确度、方向、尺寸、纹理和颜色等，但他们大多遇到草图本身绘制不准确或耗时的轮廓匹配的问题。并且这种方法也有一些问题：用户画的线条一般较为粗糙，即便对于人来说都有多种解读，更不用说机器。因此手绘线条的解析就成了一个大家关注的悬而未解的问题。其他一些学者提出了多条件的包括基于内容和文本不同程度混合的方法。在文献[9]中，Kaster 等人描述了一个多模式系统，包括用手势和语音来查询和交互反馈。Assfalg 等人^[10]使用 3D 模型作为线索捕捉图像中的对象的空间排列。

除了线索的提供之外，CBIR 面临的另一个重要的挑战在于基于内容的匹配过程耗时较长。对于动辄成千上万张的待匹配图片来说，这无疑是一个很大的挑战。对于基于内容的检索中特征匹配耗时长的问题，一些学者选择将关键的步骤进行并行处理比如用 GPU^[11-13]，但是图像数据库一般是分布式存储在多台机器上，这种集中式的机制将导致巨大的数据传输，从而可能使得网络崩溃或者是很高的

传输时间代价。MapReduce 框架的出现，提供了另一种方式的并行化处理。MapReduce 是一种最初由谷歌提出来的并行计算框架^[14]，通过一个普通机器集群来处理海量的数据。受高级函数式编程语言的启发，MapReduce 提供了一个抽象的方法，然后让程序员自己去实现“mapper”（特定的记录的计算）和“reducer”（特定的结果的聚合）。这个框架将透明地在集群上执行其它的步骤，包括调度，故障处理等。这种方式被成功地应用到许多应用中比如各种各样的计算机视觉任务^[15-18]，场景渲染^[19,20]以及医学领域^[21,22]等。

虽然也有学者在基于 MapReduce 的图像检索方面做了一些工作，但他们更多关注的是一些常规任务如索引结构的构建。与目前的检索算法不同的是，我们独特的多线索检索策略是为图像创作应用量身定做的，对检索结果的实时反馈要求高。因此，新的模式需要在 MapReduce 框架下设计，以提高检索的交互性。

1.3 论文研究内容

本文根据前面提出的研究背景和意义以及研究现状，针对目前图像检索方法无法保证检索结果的图像质量以及在面向海量素材的特征匹配耗时长的问题，提出了一种高质量图像的渐进检索算法。该算法除了能够达到传统的图像检索目的，而且特别适用于数字内容创作应用中。因为算法最终输出的结果图像品质高，能够很好地满足这些应用对输入图像的质量要求，方便快捷地提供了满足用户意图的创作素材。并且由于本算法采用渐进式过滤的策略，同时研究了使用 MapReduce 框架实现了计算的并行化，保证算法效率，所以能够很好地满足这些应用的实时交互的要求，实用性高。

该算法的基本思想是图像通过各种条件被渐进地过滤出：首先我们根据图像标签进行快速的文本比较，从海量图像素材库中获取待匹配图像；然后这些候选图像通过一系列的质量检查，以进一步剔除那些图像质量低的部分。而如何确定评价图像品质的标准，是一个关键的问题，以便不仅能够尽可能地在早期过滤掉质量低的图像，而且剩余的图像能够符合人类审美习惯并且能够满足进一步的图像分析、编辑方法对输入图像的一般要求。最后我们将图像的前景元素通过图像分割技术提取出来，并与用户绘制的线条草图进行最耗时的轮廓匹配步骤，从而

得到最终的结果列表。由于一系列用于评价图像品质的标准都是由图像的各种特征组成，这些特征的计算和匹配都是计算量大的任务，非常耗时，所以如何保证算法的实时性也是一个重要的研究内容。

1.4 论文组织结构

本篇论文一共分为七个章节。每个章节安排的具体内容如下：

第一章为绪论。主要介绍本文的研究背景和研究意义。并对目前基于内容的图像检索和 MapReduce 框架在计算机视觉领域应用的研究现状做出简单介绍，最后描述本文的主要研究内容和整篇论文的主要结构安排。

第二章为相关技术和理论介绍。首先介绍了基于内容的图像检索的一般工作流程和图像检索系统的总体框架；其次具体介绍了图像分割算法，此算法在本文算法中用于计算图像显著性对象的背景复杂度。最后对 MapReduce 框架以其一般编程思想进行了简要介绍，为后文的渐进检索算法的并行实现做好铺垫。

第三章为算法框架设计。首先研究哪些线索可以有效地过滤低品质的图像，并且匹配用户检索意图，接着基于这些线索，设计一种高效的组合方式，渐进地过滤图像，以便尽早地过滤掉较多不符合标准的候选图像，提高检索性能。并且具体介绍了算法中涉及的前景对象提取和轮廓匹配步骤。

第四章为渐进的图像内容过滤。详细介绍了保证检索结果图像质量的过滤策略，从理论依据到数学公式定义等方面对检索对象清晰度和对象可分离度两类策略开展阐述。

第五章为基于 MapReduce 框架的并行实现。针对上一章的过滤算法中各种特征的计算和匹配步骤耗时长的问题，利用 MapReduce，分为文本检索和特征过滤两阶段进行并行实现，提高整个检索算法的效率，以满足实时性要求。

第六章为实验过程及结果分析。首先对本文算法的整体实验方案和实验环境进行了简要描述，并分别从检索性能和检索结果的图像质量两方面进行了验证本算法的有效性。最后基于本文检索算法，实现了一个图像合成原型系统，进一步证明本算法在数字内容创作方面的应用价值。

第七章为总结与展望。对本文工作进行总结，并分析了尚未解决的问题，并

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.

厦门大学博硕士学位论文摘要库