

学校编码: 10384

学 号: 200428036

分类号 \_\_\_\_\_ 密级 \_\_\_\_\_

UDC \_\_\_\_\_

厦 门 大 学

硕 士 学 位 论 文

# 基于 SIFT 特征的图像检索技术研究

## Research on Content-Based Image Retrieval

### Using SIFT

吴锐航

指导教师姓名: 李绍滋 教授

专业名称: 计算机应用技术

论文提交日期: 2007 年 5 月

论文答辩时间: 2007 年 6 月

学位授予日期: 2007 年 月

答辩委员会主席: \_\_\_\_\_

评 阅 人: \_\_\_\_\_

2007 年 5 月

# 厦门大学学位论文原创性声明

兹呈交的学位论文，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）： 吴 锐 航

2007 年 6 月 1 日

# 厦门大学学位论文著作权使用声明

本人完全了解厦门大学有关保留、使用学位论文的规定。厦门大学有权保留并向国家主管部门或其指定机构送交论文的纸质版和电子版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

本学位论文属于

- 1、保密（ ），在            年解密后适用本授权书。
- 2、不保密（）

（请在以上相应括号内打“√”）

作者签名： 吴 锐 航      日期： 2007 年 6 月 1 日

导师签名：                      日期：      年    月    日

## 摘 要

基于内容的图像信息检索是多媒体信息检索领域中重要的组成部分,对图像的内容进行准确快速的描述一直都是图像检索技术中研究的重点和难点。传统的图像特征提取方法,基本上是围绕图像的颜色、纹理、形状和空间关系来展开的。本文提出一种基于 SIFT 特征的新的图像信息检索算法。SIFT 特征向量是一种图像局部特征向量。它对于图像的尺度缩放、旋转、平移以及一定程度的仿射和光照变化具有良好的不变性。本文的主要内容如下:

1. 系统分析和总结 SIFT 特征向量的特点,尝试将它应用到基于内容的图像信息检索中,并且改进传统的图像相似度度量方法。实验证实,改进后的图像距离度量更适合基于 SIFT 特征的图像检索。
2. 采用主成分分析的方法对 128 维的 SIFT 特征向量进行降维。这种主成分分析主要是通过 K-L 变换来实现的。它消除了各分量之间的相关性,大大降低了算法的计算量。
3. 借鉴国际上公认的高维索引技术——基于向量近似的索引技术对降维后的特征向量进行量化近似,加速相似向量搜索过程中的数据过滤速度。
4. 通过实验进一步证实将 SIFT 特征提取方法应用在基于内容的图像检索中的可行性,并简单分析了它的应用前景。

**关键字:** SIFT 特征向量; 主成分分析; 向量近似搜索

## Abstract

Content based image retrieval (CBIR) is one of the fields of Multimedia information retrieval. The difficulty of CBIR is to properly express the contents of the images. Current CBIR systems generally make use of lower-level features like color, texture, shape and space relationship. This paper presents a new approach to extract image features for CBIR, which based on the SIFT features. SIFT features are distinctive local image features. The features are invariant to image scaling, translation, and rotation, and partially invariant to illumination changes and affine. The main points of this paper are as follows:

1. Systemically analyzed and concluded the character of SIFT features, innovately apply it into CBIR, and improved the general measurement for image similarity. Experiments have showed that the improved distance measurement is better fit for the image similarity presented in this paper.
2. Apply principal components analysis (PCA) to reduce the 128 dimensional vectors to lower dimension for easy data analysis. It is carried out by K-L transform which has the distinction of being the optimal linear transformation for keeping the subspace that has largest variance.
3. Referring to the vector approximation based indexing for non-uniform high dimensional data sets approach, the speed of filtering datasets has been accelerated.
4. Experimental results show that the robust image retrieval can be achieved by apply SIFT features into CBIR. The potential applications have been generalized simply at the end of the paper.

**Key Words:** SIFT features; Principal Components Analysis; Vector Approximate Searching

## 目 录

摘 要.....	i
Abstract.....	ii
<b>第 1 章 绪 论.....</b>	<b>1</b>
1.1 基于内容的图像信息检索概述.....	1
1.2 国内外发展概况.....	1
1.3 基于内容的视觉信息检索的基本工作框图和功能模块.....	2
1.4 基于内容的视觉信息检索的应用领域.....	4
1.5 基于内容的图像检索面临的挑战.....	5
1.6 本文的主要研究内容与组织结构.....	5
<b>第 2 章 SIFT特征向量提取.....</b>	<b>7</b>
2.1 图像的特征向量简述.....	7
2.2 图像的多尺度表示.....	8
2.2.1 图像金字塔.....	8
2.2.2 高斯金字塔.....	10
2.2.3 高斯滤波.....	10
2.3 SIFT特征提取算法.....	11
2.3.1 高斯差分DoG滤波.....	11
2.3.2 尺度空间的极值检测.....	12
2.3.3 关键点的位置确定.....	14
2.3.4 关键点的方向参数.....	16
2.3.5 SIFT特征向量描述.....	17
<b>第 3 章 向量的主成分分析.....</b>	<b>19</b>
3.1 预备知识.....	19
3.1.1 “维数灾难”现象.....	19
3.1.2 多媒体数据降维：线性变换.....	19
3.1.3 图像的正交变换.....	20
3.2 主成分分析简介.....	21
3.3 K-L变换.....	23
3.3.1 K-L变换的计算过程.....	24
3.3.2 K-L变换的性质和特点.....	25
<b>第 4 章 向量的近邻搜索.....</b>	<b>26</b>
4.1 预备知识.....	26

4.1.1	向量间的距离度量 .....	26
4.1.2	k最近邻 .....	27
4.2	相关研究 .....	27
4.3	索引结构 .....	28
4.3.1	VA-File索引结构 .....	28
4.3.2	VA <sup>+</sup> -File索引结构 .....	30
4.3.3	近邻搜索算法 .....	34
<b>第5章 实验及结果分析 .....</b>		<b>36</b>
5.1	检索性能指标和评价准则 .....	36
5.1.1	查全率(recall)和准确度(precision) .....	36
5.1.2	检索效率(efficiency of retrieval) .....	37
5.1.3	检索评分 (retrieval score) .....	37
5.2	图像相似性度量方法 .....	40
5.2.1	传统的方法 .....	40
5.2.2	改进的方法 .....	40
5.3	算法工作流程图 .....	42
5.4	实验环境 .....	43
5.5	实验 .....	43
5.5.1	distRatio的确定 .....	43
5.5.2	算法不变性测试 .....	44
5.5.3	在真实数据库中的测试 .....	47
5.6	算法的分析 .....	50
<b>第6章 总结和展望 .....</b>		<b>52</b>
<b>参 考 文 献 .....</b>		<b>54</b>
<b>研究生期间发表的论文 .....</b>		<b>58</b>
<b>致 谢 .....</b>		<b>59</b>

## Contents

<b>Abstract in Chinese .....</b>	<b>i</b>
<b>Abstract.....</b>	<b>ii</b>
<b>Chapter 1: Introduction.....</b>	<b>1</b>
<b>1.1 What is CBIR? .....</b>	<b>1</b>
<b>1.2 Developments at Home and Abroad .....</b>	<b>1</b>
<b>1.3 Working Flow Chart and Function Modules.....</b>	<b>2</b>
<b>1.4 Scope and Application.....</b>	<b>4</b>
<b>1.5 Challenges .....</b>	<b>5</b>
<b>1.6 Research Objectives and Thesis Organization .....</b>	<b>5</b>
<b>Chapter 2: SIFT Feature Extraction .....</b>	<b>7</b>
<b>2.1 Introduction to Image Feature .....</b>	<b>7</b>
<b>2.2 Multi-Resolution Images.....</b>	<b>8</b>
2.2.1 Image Pyramid.....	8
2.2.2 Gaussian Pyramid.....	10
2.2.3 Gaussian Filtering.....	10
<b>2.3 SIFT Feature Extraction.....</b>	<b>11</b>
2.3.1 Difference of Gaussian .....	11
2.3.2 Detection of Scale Space Extrema .....	12
2.3.3 Accurate Keypoint Localization.....	14
2.3.4 Orientation Assignment .....	16
2.3.5 SIFT Feature Descriptor .....	17
<b>Chapter 3: Principal Component Analysis for SIFT Vectors ....</b>	<b>19</b>
<b>3.1 Related Work .....</b>	<b>19</b>
3.1.1 Curse of Dimensionality.....	19
3.1.2 Media Data Reduction: Linear Transformation.....	19
3.1.3 Orthogonal Transform for Image.....	20
<b>3.2 Introduction to Principal Component Analysis .....</b>	<b>21</b>
<b>3.3 K-L Transform.....</b>	<b>23</b>
3.3.1 Computation of K-L Transform.....	24
3.3.2 Property of K-L Transform.....	25
<b>Chapter 4: Nearest Neighbor Searching of SIFT Vectors .....</b>	<b>26</b>
<b>4.1 Related Work .....</b>	<b>26</b>

4.1.1	Distance Measurement .....	26
4.1.2	k Nearest Neighbors .....	27
<b>4.2</b>	<b>Related Research .....</b>	<b>27</b>
<b>4.3</b>	<b>Indexing .....</b>	<b>28</b>
4.3.1	VA-File Indexing Structure .....	28
4.3.2	VA <sup>+</sup> -File Indexing Structure .....	30
4.3.3	Nearest Neighbors Searching .....	34
<b>Chapter 5: Experiment and Result Analysis .....</b>		<b>36</b>
<b>5.1</b>	<b>Retrieval Measurement and Judgment .....</b>	<b>36</b>
5.1.1	Recall and Precision .....	36
5.1.2	Efficiency of Retrieval .....	37
5.1.3	Retrieval Score .....	37
<b>5.2</b>	<b>Similarity Measurement between Images .....</b>	<b>40</b>
5.2.1	General Measurement .....	40
5.2.2	Improved Measurement .....	40
<b>5.3</b>	<b>Flow Chart .....</b>	<b>42</b>
<b>5.4</b>	<b>Experimental environment .....</b>	<b>43</b>
<b>5.5</b>	<b>Experiment .....</b>	<b>43</b>
5.5.1	distRatio Determination .....	43
5.5.2	Testing the Invariance .....	44
5.5.3	Testing in Real Databases .....	47
<b>5.6</b>	<b>Analysis of Algorithm .....</b>	<b>50</b>
<b>Chapter 6: Conclusions and Recommendations for Future Work ...</b>		<b>52</b>
<b>References .....</b>		<b>54</b>
<b>Papers Published .....</b>		<b>58</b>
<b>Acknowledgements .....</b>		<b>59</b>

## 第1章 绪论

本章是对全文和主题的概括介绍,简要介绍了基于内容的图像信息检索的背景、国内外的的发展状况。同时大致介绍了基于内容的图像信息检索的基本工作框图和各个功能模块,罗列了它的主要应用领域。最后概括了这项技术面临的挑战,并给出了全文的主要研究内容和组织结构。

### 1.1 基于内容的图像信息检索概述

基于内容的图像信息检索是由于视觉信息的飞速膨胀而得到关注并被提出来的 [1]。

基于内容的图像信息检索 (Content-Based Image Retrieval, CBIR), 是计算机视觉在图像检索中的应用。它的目的是从视觉数据库中快速地提取出与一个查询相关的图像或图像序列。“基于内容”是指利用图像的视觉信息, 而不是传统的视觉信息检索方案。它借助对视觉媒体从低层到高层进行处理、分析和理解的过程获取其内容并根据内容进行检索。它所解决的是图像在没有“索引”、“目录”和“摘要”等人为输入信息下的检索问题。

CBIR 一词产生于 1992 年, 并在国际上使用。CBIR 技术最初来源于统计学、模式识别、信号处理, 以及计算机视觉。

传统的基于使用文字标识符的图像信息检索方案的固有的局限性, 使得 CBIR 技术越来越多地受到关注。传统的方法不能完整地概括图像的视觉内容, 并且需要事先人为的给图像数据库中的图像加注标签, 因而在很多情况下不能满足使用的要求。为了解决上述问题, 需要全面地、一般性地和客观地提取视觉内容。实际上人们利用图像视频并不是仅根据他们的视觉质量而更重要的是根据他们的视觉内容 [2], 所以根据内容进行检索才可能有效地获得所需的视觉信息, 同时在掌握视觉内容的基础上数据库中的信息才可能有效地得到管理。

### 1.2 国内外发展概况

1992 年, “基于内容的图像检索(content-based image retrieval, CBIR)” 开始在国际上使用, 国内也很早就开展了相关研究。10 年来, 这方面的研究和应用得

到长足的进展。基本上可以分为以下两代：

第一代，视觉信息检索系统允许通过字符串属性访问图像和视频。用字符串表达与内容无关的元数据，使用如相关模型、框架模型、面向目标的模型等表达方案。与内容相关的元数据和内容描述元数据通过关键字或剧本描述。检索可基于在文字领域工作的传统的搜索引擎，使用传统的查询语言（如 SQL），或全文检索。

第二代，新一代的视觉信息检索系统支持视觉内容的检索。对视觉信息的访问不仅在概念层利用关键字进行（与在文字领域类似），而且也在感知层利用对视觉内容的客观测量和合适的相似模型进行。由于查询通过视觉范例来表示，所以用户使用的灵活性得到了很大的提高。

### 1.3 基于内容的视觉信息检索的基本工作框图和功能模块

基于内容的检索有比较一致的基本工作框架，因为它们都要完成对视觉信息的处理、分析和理解。在这个基本工作框架下，为完成检索任务还需要一些功能模块。在每个功能模块里，需要采用一系列技术来完成相应的任务。这些技术里，既包括基本的、对检索图像和视频都需要的通用技术，也包括针对图像和视频的不同特点需要采用的以及考虑特定应用所要用到的特定技术。

#### 1. 基本框架

基于内容的视觉信息检索是一种重要的和关键的多媒体信息检索处理技术。一般可把基于内容的视觉信息检索系统看作是介于信息用户和（多媒体）数据库之间的一种信息服务系统。用户通过它可按自然的方式从库中提取满足所需要内容的媒体数据。这类系统的通用框架如图 1-1 所示，主要有五个模块组成，见下图虚线框内。

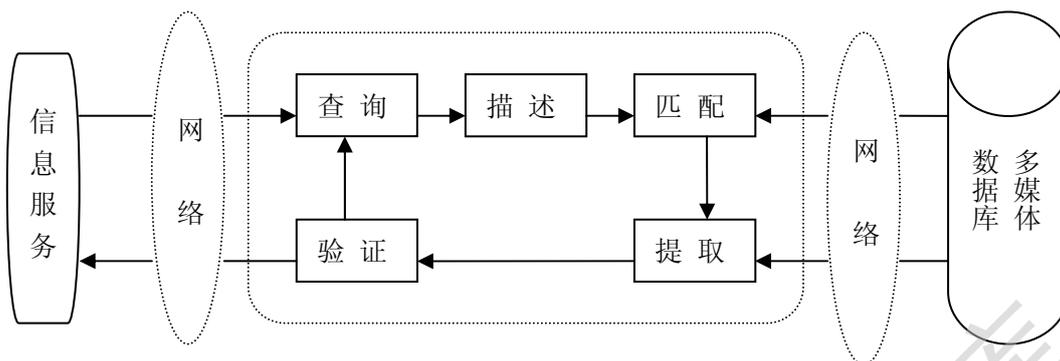


图 1-1 基于内容的视觉信息检索系统基本框图

用户发出查询要求，系统将查询要求转化为计算机内部描述，并借助这些描述与库中信息进行匹配，提取出需要的信息数据，用户对此验证后可直接使用或借以改进查询条件并开始新一轮检索。

由于基于内容的多媒体信息检索在信息用户和多媒体数据库不直接连通的情况也可以借助各种网络以服务器和客户机的形式远程进行，所以图中在信息用户与检索系统以及检索系统与多媒体数据库间画出了网络。

## 2. 模块功能简介

下面对各模块的主要功能给予概括解释。

### (1) 查询模块

其主要功能是对用户提供多样的查询手段，以支持用户根据不同应用进行各种类型的查询工作。换句话说，用户要进行查询，先提出要求条件，这些要求条件主要基于对图像内容的描述。例如用户可提出感兴趣目标的几何形状或所需要图像的背景颜色等作为要求条件。

### (2) 描述模块

其主要功能是将用户的查询要求转化为对图像内容的比较抽象的内部表达和描述，即通过对图像的分析，从而以一定的、计算机可以方便表达的数据结构建立对图像内容的描述。这个模块在图像数据库建库时也需对每幅图像进行。

### (3) 匹配模块

其主要功能是在图像数据库中搜索所需的图像内容。因为对被查询图像建立的表达描述也已对图像数据库中的图像建立了，所以将对查询图的描述与图像数

数据库中被查询图的描述进行内容匹配和比较就可以确定它们在内容上的一致性和相似性。这个匹配的结果将传给提取模块。

#### (4) 提取模块

其主要功能是根据匹配的结果在图像数据库中对感兴趣的图像定位,并在内容匹配的基础上将图像数据库中所有满足给定要求条件的图像自动地提取出来以让用户使用。如果事先对图像数据库建立了索引,这样在提取时就可提高效率。

#### (5) 验证模块

其主要功能是帮助验证如上提取的图像是否满足用户要求。根据目前技术水平和设备条件,在自动查询和提取的基础上用户还需要最后验证结果的手段。如果验证效果不满意,新一轮的查询可通过修改查询条件重新开始。

实际中对每个模块都有许多具体技术可以采用。由于图像和视频各有其特点,所以对图像和视频检索所采用的技术也有所不同。

### 1.4 基于内容的视觉信息检索的应用领域

基于内容的多媒体信息检索建立在对多媒体信息内容处理、分析和理解的基础上,且在多媒体应用中占据重要的位置。

#### 1. 应用领域

基于内容的视觉信息查询具有广泛的应用领域 [3], 例如:

- ①. 遥感和地球资源信息的管理和共享;
- ②. 电子图书馆、艺术博物馆、远程教育;
- ③. 医学分析、研究和远程医疗系统, 特别对远程会议, 基于内容的图像查询可提高效率并节省大量的网络传输费用;
- ④. 专利检索、商标注册管理;
- ⑤. 人口户籍管理、档案查询, 以及公安、安全监视系统;
- ⑥. 数字视频的协同生产和编辑以及传输播发;
- ⑦. 地图地理信息系统、地理事件分析;
- ⑧. 军事侦察作战等应用;
- ⑨. 服装设计: 服装设计师可根据其设计的要求提出需要的图案进行选择;
- ⑩. 建筑设计、内部装潢。

以上各方面对基于内容的图像检索系统的需求是大量的。随着图像获取方式的改善和增强, 各类图像库的建立和扩展, 这种需求将变得更为迫切。

## 1.5 基于内容的图像检索面临的挑战

基于内容对视觉信息的检索是一个挑战性的研究, 目前方兴未艾。它涉及许多领域, 必须要考虑许多问题, 包括从获取和表达原始数据的方法, 获取和表达原始数据中的信息, 到通过对这些数据和信息的处理、分析和理解以提供对内容的访问以及发送、显示和操作这些内容等。这里的每一个步骤都受到数据种类、应用领域和用户(对系统有要求的人或信息代理)的影响。由于这些问题的存在, 基于内容的视觉信息检索的研究和应用还面临许多挑战。例如, 如何建立语义层次的描述, 使用户自然地与检索系统交互, 又例如如何挖掘更多的实际应用场合, 是视觉信息检索技术真正使用起来, 都是推动基于内容的视觉信息检索所必须解决的问题。

当前研究图像检索的重点和难点集中在如何在抽取的底层特征和图像内容所表示的语义特征间建立很好的联系, 直到今天, 这仍然是基于内容图像检索的一个没有解决的问题。也就是说, 由于我们所拥有的特征并不能很好地体现图像真正的语义信息, 以至于检索的结果往往不能令人满意 [1]。

## 1.6 本文的主要研究内容与组织结构

本文以多尺度的图像特征 SIFT (Scale Invariant Feature Transform) 为基础, 提出一种新的 CBIR 检索算法, 它适用于从不同视角观察得到目标或场景图像的检索。

本文的主要研究工作分为以下几点: 1.系统的分析和总结图像的 SIFT 特征向量的特点, 创新性地将其应用到基于内容的图像检索算法中, 并根据 SIFT 特征向量的特点, 改进传统的图像间相似距离的度量方法, 使得改进后的相似度量更适合于基于 SIFT 特征的图像检索算法; 2.为了降低特征向量匹配搜索时的计算复杂度, 采用主成分分析的方法对特征向量降维; 3.为了加速算法的数据过滤能力, 借鉴了向量近似的方法对特征向量进行索引。在上述的工作中, 主要的创新点有三: 1.将 SIFT 特征向量应用到基于内容的图像检索中; 2.采用主成分分析

的方法对 128 维的 SIFT 特征向量降维；3.改进传统的图像间相似距离的度量方法，使其更适合于基于 SIFT 特征的图像检索算法。

本文的组织结构如下：

- 第一章，绪论，对 CBIR 背景及国内外发展的现况进行概述，介绍了 CBIR 系统的基本工作框图和各个功能模块，提出了这项新的技术研究的重要意义及面临的挑战，并在最后简要概括本文的主要研究内容和组织结构。
- 第二章，SIFT 特征向量提取，介绍多尺度图像描述方法，分析研究一种对尺度空间、图像缩放、旋转甚至是仿射不变的图像局部特征描述算子——SIFT 特征向量，揭示它在 CBIR 中的应用可能性。
- 第三章，向量的主成分分析，对 128 维的 SIFT 特征向量进行主成分分析，具体的做法是利用 K-L 变换来消除各分量之间的相关性，降低数据处理的复杂度。
- 第四章，近邻搜索，借鉴高维索引技术——基于向量近似的索引技术加速数据过滤。
- 第五章，实验及结果分析，首先介绍图像检索算法的性能指标和评价准则，然后对前文中构建的基于 SIFT 特征的图像检索的算法模型进行实验测试，同时分析算法的优缺点。
- 第六章，总结和展望。对本文的工作进行了总结，分析了这项技术的应用前景，提出了进一步的工作目标。

## 第2章 SIFT 特征向量提取

检索中最常用的模型是基于特征向量的，本章简要介绍了图像特征向量的特点，以及常见的图像特征描述方法，然后给出本文算法中应用的图像特征——SIFT 特征的提取方法。

### 2.1 图像的特征向量简述

如果要建立一个识别不同种类对象的系统，首先必须要确定应测量对象的哪些特征以产生描述参数。被测量的这些特殊的属性称为对象的特征，而所得的参数值组成了每个对象的特征向量。适当地选择特征是很重要的，因为在识别对象时它是惟一的依据。良好的特征应具有 4 个特征 [4]:

1. 可区别性 对于属于不同类别的对象来说，它们的特征值应具有明显的差异。
2. 可靠性 对同类的对象，特征值应比较相近。
3. 独立性 所用的各特征之间应彼此不相关。相关性很高的特征可以组合起来使用，而不应该作为单独的特征使用，以减少噪声干扰。
4. 数量少 信息检索系统的复杂性随系统的维数（特征的个数）迅速增长。较多的特征向量个数将带来系统的复杂性和检索效率的低下。

目前在图像检索系统中广泛应用到的图像的底层特征有颜色、纹理、形状、空间关系等，这些特征对于图像检索有着不同的作用，但是也同时存在着不足。如颜色是一种全局特征，它对图像或图像区域的方向、大小等变化不敏感，所以颜色特征不能很好的捕捉图像中对象的局部特征，也不能很好表达颜色的空间分布信息。纹理也是一种全局特征，但纹理只是物体表面的一种特性，并不能完全反映物体的本质属性，而且其中一点严重的不足是，当物体受到光照、反射等影响时，从图像中反映出来的纹理并不一定是物体真实的纹理。这些虚假的纹理都会对检索造成“误导”。全局的特征描述子并不能很好的适应对目标感兴趣的图像的检索。基于形状的特征常常可以利用图像中感兴趣的目标进行检索，但是形状特征的提取又常常受到图像分割效果的影响。空间关系特征可以加强对图像内

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](mailto:etd@xmu.edu.cn) for delivery details.

厦门大学博硕士学位论文摘要库