

学校编码: 10384

分类号_____密级_____

学号: X2010230106

UDC_____

廈門大學

工程硕士学位论文

初唐楷书字体检索系统的研究与设计

Research and Design of Retrieval System for
Early Tang Dynasty Regular Script Font

杨树森

指导教师姓名: 陈海山 教授

专业名称: 软件工程

论文提交日期: 2012 年 10 月

论文答辩时间: 2012 年 12 月

学位授予日期: 年 月

答辩委员会主席: _____

评 阅 人: _____

2012 年 10 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,本学位论文为()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

随着社会的发展，数字资源的使用越来越受到人们的重视，如何高效的利用数字资源成为当今的热点话题。图像检索技术的出现使得数字资源的高效利用成为可能，该项技术是一种基于内容的综合检索技术，它综合了多种学科，这种技术现在大多应用于图书馆的目录检索和专业图库方面的检索。虽然基于形状特征的图像检索技术的应用有很多其他相关技术无法比拟的优点，但是由于物体具有稳定性的特征以及目前还没有对该项技术进行有效的定义，这使得这种技术在研究过程中遇到了很多的难题，但其仍是人们当前研究的热点问题。

本文通过对初唐时期楷书字体的分割、特征提取、聚类 and 典型度量方法一系列步骤，建立一个有关初唐楷书字体的特点和字形独立数据库，并且运用图像检索技术对初唐楷书字体进行标签化的索引。初唐楷书在我国的书法史中占有重要的地位，以欧阳询、虞世南、褚遂良、薛稷为代表的初唐楷书四大家，其作品风格基本体现了初唐整个楷书的笔法格调及流向。对初唐楷书形状特征的提取和相似性测度进行研究，对于保护和发扬我国的传统书法文化有着重要的意义。

本文以楷书图库为研究对象，在特征提取及描述中，利用初唐楷书字体分割、字体字形特征的提取、聚类法和典型度量方法等技术获得目标区域，用不变矩描述区域。实验表明，在对图像的内部和外部进行处理的过程中，当使用聚类法以及不同度量方法时，这些方法不仅能够达到滤波的作用，还能够使不连续的边缘点连接起来。

关键词：初唐楷书；字体检索；图像识别

Abstract

With the development of the society, the usage of the digital resources is more and more important. It's a hot issue how to use the digital resources efficiently. The emergence of the image retrieval technology makes it possible of efficient use of digital resources. The technology is a content-based retrieval technique. It is combined with many other technologies. It's mainly used in professional library and the library catalog search. The shape does not vary with changes in image color change is stable characteristics of an object, and at this stage there is no definition as to its accuracy, shape feature based image retrieval technology has become a hotspot and difficulty of the present research.

By taking the early Tang period regular script font segmentation, feature extraction, clustering and measure this series of steps to establish the characteristics of a regular script in early and shaped an independent database, and skilled use of image retrieval technology, easy for beginners Tang regular script font for the label index function. the early Tang regular script occupies an important position in the history of our country to express the early Tang regular script represents the entire regular script style of the early Tang, Ouyang Xun, Yen Chen, Liu and public rights, Zhao Mengfu, extraction and description of the shape characteristics of the early Tang regular script and similar measure research has important significance for the protection and carry forward the traditional calligraphy culture that I had.

Regular script library, feature extraction and description of the early Tang regular script font, segmentation, and the font glyph feature extraction, clustering method, and different measurement methods and other techniques the target area, use the same moment to describe the area. the experimental results show that the clustering method and a different measure of processing internal and external image of the filtering effect of connection discontinuous edge points; weighted Euclidean distance, which reduces the higher order moments of the amplitude fluctuations of search results impact.

Key words: Early Tang Regular Script; Font Retrieval; Image Recognition

目 录

第 1 章 绪论	1
1.1 研究背景及意义	1
1.1.1 研究背景	1
1.1.2 研究意义	1
1.2 国内外研究状态	2
1.3 研究目标与内容	6
1.4 本文组织结构	7
第 2 章 初唐楷书字体分割	8
2.1 分割原理	8
2.1.1 分割目的	8
2.1.2 分割原则	8
2.2 分割方法	9
2.2.1 边缘检测	9
2.2.2 阈值分割	14
2.3 分割作用	18
2.3.1 分割的前提	18
2.3.3 分割的条件	18
2.3.4 分割的作用	19
2.4 本章小结	19
第 3 章 字体字形特征的提取	20
3.1 提取楷书字体特征	20
3.1.1 初唐楷书字体分析	20
3.1.2 提取字体特征	20
3.1.3 提取时注意事项	21
3.2 提取楷书字形特征	25

3.2.1 形状的特征与提取	27
3.2.2 提取横画	27
3.2.3 提取点画、长撇与捺笔	28
3.2.4 提取竖画	28
3.2.5 提取提笔、勾笔	28
3.3 本章小结	28
第 4 章 聚类方法研究	29
4.1 聚类方法	29
4.1.1 聚类分析法简介	29
4.1.2 聚类分析法依据	29
4.1.3 三种聚类法分析	29
4.1.4 三种聚类法比较	32
4.2 典型度量方法	32
4.2.1 度量方法分类	32
4.2.2 明式距离	33
4.2.3 直方图相交	34
4.2.4 马氏距离	34
4.3 聚类与度量结合	34
4.4 本章小结	37
第 5 章 检索系统设计	38
5.1 初唐楷体	38
5.1.1 楷书的历史演变	38
5.1.2 初唐楷体的特点	39
5.2 系统功能设计	40
5.2.1 系统功能结构	40
5.2.2 主要功能模块	40
5.3 楷体信息检索	43
5.3.1 特征提取描述	43
5.3.2 反馈机制和索引	45

5.4 实验结果分析.....	46
5.5 本章小结.....	49
第 6 章 总结与展望.....	50
6.1 总结.....	50
6.2 展望.....	50
参考文献.....	52
致谢.....	53

厦门大学博硕士学位论文摘要库

Contents

Chapter 1 Introduction	1
1.1 Research Background and Significances	1
1.1.1 Research Background.....	1
1.1.2 Research Significances.....	1
1.2 Research Status at Home and Abroad	2
1.3 Research Objectives and Contents	6
1.4 Outline of the Dissertation	7
Chapter 2 Segmentation of the Early Tang Dynasty Fonts	8
2.1 Segmentation Principle	8
2.1.1 Regular Script Font Segmentation Purpose.....	8
2.1.2 Regular Script Font Segmentation Principle.....	8
2.2 Segmentation Method	9
2.2.1 Edge Detection.....	9
2.2.2 Threshold Segmentation.....	14
2.3 Segmentation Role	18
2.3.1 Segmentation Premise.....	18
2.3.3 Segmentation Conditions.....	18
2.3.4 Segmentation Role.....	19
2.4 Summary	19
Chapter 3 Extraction of Font Characteristics	20
3.1 Extract Regular Script Font Characteristics	20
3.1.1 Analysis of Regular Script Font.....	20
3.1.2 Extract Font Characteristics.....	20
3.1.3 Attention of the Extraction.....	21
3.2 Extract Regular Script Shape Characteristics	25
3.2.1 Feature and Extraction of the Shape.....	27

3.2.2 Extract Parallel Stroke Action	27
3.2.3 Extract Point, Left Hook and Stroke Action	28
3.2.4 Extract Vertical Stroke Action	28
3.2.5 Extract Hook Stroke Action	28
3.3 Summary	28
Chapter 4 Research on Clustering Methods.....	29
4.1 Clustering Methods	29
4.1.1 Brief of Cluster Method.....	29
4.1.2 Basis of Cluster Analysis Method	29
4.1.3 Analysis of Three Kinds of Clustering Methods	29
4.1.4 Comparison of Three Kinds of Cluster Methods	32
4.2 Typical Measure Methods	32
4.2.1 Measure Methods	32
4.2.2 Ming Distance	33
4.2.3 Histogram Intersect.....	34
4.2.4 Mahalanobis Distance	34
4.3 Combination of Clustering and Measurement	34
4.4 Summary	37
Chapter 5 Retrieval System Design	38
5.1 Early Tang Dynasty Regular Script	38
5.1.1 Evolution of Regular Script History.....	38
5.1.2 Characteristics of Regular Script.....	39
5.2 System Function Design	40
5.2.1 System Function Structure	40
5.2.2 Main Function Modules	40
5.3 Regular Script Information Retrieval	43
5.3.1 Description of Feature Extraction.....	43
5.3.2 Feedback Mechanisms and Index.....	45
5.4 Analysis of Experiment Results.....	48

5.5 Summary	49
Chapter 6 Conclusions and Prospects	50
6.1 Conclusions	50
6.2 Prospects.....	50
References	52
Acknowledgements.....	53

厦门大学博硕士学位论文摘要库

第 1 章 绪论

1.1 研究背景及意义

1.1.1 研究背景

当今世界已全面进入网络时代，网络及多媒体技术皆飞速发展，并铺天盖地地渗透到了各个领域，人们因此接触到了各种各样的数字图像信息。然而让人们感到新奇的同时又苦于无法将这些数字图像信息收集与检索，因为这些图像信息分散在世界各地，缺乏有效的组织方式。为了解决这一难题，研究一个高效、准确的图像检索系统并将之推广应用就显得尤为必要。早在上世纪 70 年代，图像检索技术便以一种很具有发展潜力的技术进入人们的视野，这种技术是一种全新的领域，同时也是备受人们看好的领域。图像检索技术刚开始使用的是文本索引技术，这种技术就是根据文件的内容手动的把一些关键字标示出来，然后可以根据这些被标示的关键字信息建立起索引，这样就把图像检索变换成了文本检索。然而，现代社会的快速发展尤其是计算机技术的快速发展，使得人们对这种利用手工标示的图像检索不能感到满意。因此，到了九十年代，基于内容的图像检索（CBIR）这一技术得到了广泛的研究。CBIR 这一技术看其字面意思就可以猜出，这种技术在对图像信息进行处理过程中，主要是通过从低层到高层进行的。我们这里所指的内容信息包括很多方面，例如图像的颜色、形状、纹理、语义等等内容。CBIR 主要是通过分析图像本身所包含的一些内在的特征以及图像之间的联系来进行检索，它是将计算机图像处理技术和数据库技术结合起来的产物，是相较于早期图像检索技术的一大飞跃。从目前研究和应用的现状来看，基于内容的图像检索技术目前主要应用在专业图库以及图书馆的目录检索等领域。

可是基于图像检索的字体检索系统，有时候不能被用户快速准确检索，因此国内外学者对此进行了广泛的研究。初唐楷书作为中国书法文化的组成部分，在历史中占有重要的地位，从初唐楷书字体的特征等方面进行研究，设计出一种能够快速检索的系统，对文化传承具有重要的现实意义。

1.1.2 研究意义

形状特征是图像的重要特征之一。我们都知道，图像的形状信息是不会随着

图像颜色的变化而随之改变，这是物质稳定的特征。运用这一点特征，我们可以利用图像的形状特点，研究出更便于使用且愈加明显的各类图像库的检索。有些图像的颜色和纹理都不够丰富，运用颜色、纹理来研究图像检索有着很大的缺陷，这也就更加突出了形状特征的优越性。从目前的研究领域来看，基于形状的图像检索已经有了非常广泛的应用，比如指纹识别、人脸识别、医学影像、数字图书馆等等，形状特征在这些领域里都起着至关重要的作用。本文的研究目的就是要通过提取和描述初唐楷书字体的形状特征，再利用这些形状特征之间的某些相似度来实现对图像的检索。但是，图像的形状特征很难被准确的加以描述，每个人都会有自己的观点看法，因此基于形状特征的图像检索也就成为了基于内容检索中的一个难点。而本文的研究意义，就在于希望找到一种方法，能够更好的对图像的形状特征加以描述和提取，从而使查询的结果更加精确。

1.2 国内外研究状态

现今的科学界应用图像数据库的领域越来越多，图像数据库的发展优势非常的明显，尤其 CBIR 技术更加的成为人们研究的热点问题。随着这种技术在全世界的迅速发展，使得在中国也有许许多多的专业机构对这种技术领域进行不断的探索和研究，值得一提的是在中国国内已经取得一系列的成就。CBIR 的研究过程经过了复杂的演变，在刚开始的研究过程中人们把重点放在了对视觉提取上面，经过长时间的研究后，研究人员逐渐的将研究方向更改为语义特征、多维索引、用户接口、系统设计等方面。近来研究人员又在更多的领域中引进了许多新鲜的知识：运用人工智能领域的推理和神经网络技术提取形状特征并将图像加以分类；开发构造面向对象以及面向 Web 的 CBIR 系统等等，这些都是我们研究过程中需要深入研究的方面，另外，值得一提的是，并行 CBIR 技术也被提上了议程。下面，本文将介绍通用目的的图像搜索引擎中一些比较具有代表性的系统做一简单的介绍。

QIBC: IBM 的 QIBC 是第一个 CBIR 系统，同时也是比较典型的系统。这种系统是正在推向市场的一款系统，它通过一种非常友好的图形界面提供了相当多的检索方法，而且不得不承认的是，它所运用的系统框架和技术对于未来的图像检索系统的发展具有举足轻重的作用。现今有极少数的系统能够想到运营高位特征索引这一问题，但是 QIBC 就是其中之一，这一点就显示出了该系统的优越性。

该系统的结构只有三部分：图像库、特征计算、查询阶段，看起来很简单，却做的极为精妙，首先是索引技术方面，它采用的是基于颜色、纹理、形状和收回草稿的图像索引方法。在对待形状特征的选取上，QIBC 使用了圆周率、主轴方向、一些代数不变量以及形状的面积；第二个：纹路特征的本质就是纹理粗糙度、方向性、对比性的结合；另外，颜色特征的采用是非常重要的，所以，它运用了 YIQ(NTSC 补色 TV 标准)、RGB(标准色彩显示标准)、MTM(孟塞尔数学变换)坐标、lab 以及 k 维颜色直方图等各项标准来表示。QIBC 经过了一次又一次的改进，在改进的众多的版本中，开发人员越来越重视其查询的综合性，于是开发人员成功开发出综合基于文本的关键字查询与基于内容的相似性查询的综合查询。

Virage 系统：Virage 系统实际上是一种多功能的综合型搜索引擎，该系统的开发过程首先将查询引擎开发成该系统的一个插件，这种设计思路不仅能够使得该系统成功应用于通用图像查询，还使得该系统有了进一步的延伸和扩展功能，这种设计理念使得这种基于内容的图像搜索引擎广泛应用到特定的领域中去。该系统是由 Virage 公司开发的，Virage 的技术核心是针对于 Virage Engine 以及在图像对象层上进行操作。Virage Engine 主要有三个方面的功能：对图像进行有效地分析、对图像之间进行合理的比较、对图像进行科学的管理。从分层上来讲，Virage 大致可以分为以下四层：图像表达层、包含图层对象的层、领域对象层、领域事件层。对于那些在色彩、色彩布局、纹路和构架等基础所带来可视频查询，该系统都可以支持，更让人称道的是还可以支持它们的自由查询，为查询带来极大的方便。Virage 中的特征也可以被称为“原语”，根据其各自的特点，可以将其进一步划分为通用型原语和特定领域原语。

VisualSEEK：VisualSEEK 是由哥伦比亚大学研究开发的 EBIR 系统，该系统的姊妹系统是 WebSEEK，这是一个面向 WWW 的文本/图像搜索工具。该系统可以算是我们生活中较为常见的一款系统，它的主要特点就是实现了从压缩领域中抽取视觉特征以及图像区域的空间关系的查询，为我们的生活带来方便的同时，也促进了基于内容的图像检索的发展。该系统同样是由四个部分组成的：图形用户界面、服务器应用、图像检索服务器、图像归档。

WBIS:此系统使用了 WBIS（基于小波的图像索引和查找），一种图像索引检索算法，它有在大型图像库中进行部分略图查找的能力。该系统的特点就是比传

统的算法更快更准确。该系统的准确率主要是因为它一共分为两个步骤，第一步是比较粗糙的选择，是基于变量的一种选择，第二步就是在这些备选的图像中与待查询图像间的特性向量加以匹配，以找到适合的量。两步分选可以大大地提高准确率。

MARS: MARS 是由伊利诺斯大学 Urbana-ChamPaign 分校开发的一款系统，系统的名称来源于多媒体分析和检索系统的英文缩写。MARS 这个系统打破了传统模式的特有构架，从而有了自己的创新和独特。在科研方面，它呈现出来的主要特征有很多，例如索引与检索技术的结合(检测算法如何能够发挥出底层索引结构的优势)、计算机和人的融合(相关反馈技术)。

不得不说，无论是从应用领域还是调研角度来看，MARS 都能够遥遥领先于其他的图像检测软件。而 MARS 最大的特征则是因为它本身就是一个不同学科之间相互交叉融合的特定产物，这些学科包括：传统的信息检测技术、计算机视觉、数据库管理系统等。我们有理由相信，这种混合形态的 MARS，必然能够让其他的系统自愧不如。

在国内，基于内容的图像检索系统也有不少，其中最为主要的有中国科学院计算技术研究所智能信息处理重点实验室开发的基于内容的图像检索演示系统 (Mires)，该系统是一个综合利用高层语义特征和低层可视特征的图像检索系统。清华大学开发的用于检索 INTERNET 上静态图像的图像检索原型系统，该系统是图像检索中的一个突破。

当前，基于内容图像检索领域的研究热点是基于区域的小波检索和基于区域的图像检索。需要重点说一说的就是基于区域的方法，它是把图像看成是一个区域的集合，之后就对各个区域进行科学的相似度搭配，然后再通过那些相似度，准确地筛选出与图像相匹配的区域。在这样的处理方法中，首先需要对图像进行有效地分割区域处理，之后再根据相似度，把最为相近的局域集合在一块。最后，再根据颜色形态上的距离，从而准确衡量出图像到区域之间的距离。其中，最具代表的系统有：来自 Standford 大学的 SIMPUcity、以及来自 Califomia 大学的 Blobworkd 等。当然，这两种方法也有一定的相似之处。就是它们都要用到大量的图像索引技术。

CBIR 技术的应用过程中，主要根据图像的主题对象特征以及画面内容特征这

两种图像特征进行搜索。从目前的研究形式来看，基于颜色、纹理和形状的图像检索是比较常用的方式。在对颜色的相关特征进行描述的过程中，我们可以使用诸如颜色相关图、颜色一致性矢量、颜色直方图等常用方法对颜色特征进行表述。另一方面，我们还可以通过颜色的颜色对、主色调等方式对颜色进行表述，可以说，描述颜色的方法多种多样。纹理分析方法主要有以下四类：统计法，结构法，模型法，频谱法。由于图像的形状相关信息不会随着图像的颜色改变而改变，图像的形状是稳定的，对于人类来说，图像形状特征也是人们区别物体的最关键的因素，因此可以说，形状特征是图像的核心特征之一。所以，本文主要从形状特征来研究初唐楷书字体的检索技术。

由于物体形状的相对稳定性，物体形状是物体的本质特征之一。研究人员就是根据这些核心特征的综合进而实现对图形库的先进检索技术，这些物体的核心特征保证了这种技术的实用性，是这种技术的设计思路的来源，研究人员在对图像进行检索过程中就是根据物体的形状特征区别物体，进而检索出所需要的物体的。同时还可以针对专用的领域，它是利用形状特征，来进行各色图像库之间的检索。在目前的这个专业领域里，建筑的设计、商标库的注册与检测、外观设计的专利管理，机械零件库的管理和利用，服装设计库、医学图像库等都是相当重要的内容。

在目前的科技发展中，对于图像检索的特征描述方式，主要有两种。它们分别是：基于区域，基于边界。基于区域这个方法，主要是利用对区域内灰色发布信息的研究，而后者则利用图像的边缘信息。基于区域的形状特征检索这种方法主要是通过图像分割技术提取出图像中感兴趣的物体，依靠区域内图像的颜色分布信息提取特征，因而能够提取出颜色较为相近的图像。具体的描述方法主要有区域面积，形状参数，不变矩等等。其中最为常用的一种方法还属于不变矩，它主要是使用对变换不敏感的基于区域的几个矩作为形状特征，Hu 提出了 7 个这样的矩，基于他的发现，后来的研究出现了很多新的版本。黄勇，王崇骏，王亮等提出了一种基于形状不变矩的图像全局形状，并且依靠这些形象上的特征来提取其算法。

在基于边界的形状特征提取方面，巧妙地利用了各种技术的结合，它的基础是边缘检测，并且搭配着小波轮廓描述子、傅立叶描述子，根据形态上下文的特

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