

多媒体信息处理 与检索技术

- 语音处理技术
- 图像检索技术
- 版面分析
- 图像分割
- 视频信号处理与检索

朱学芳 著



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

多媒体信息处理与检索技术

朱学芳 著

電子工業出版社

Publishing House of Electronics Industry

北京 · BEIJING

内 容 简 介

本书在介绍多媒体信息处理技术的基础上,重点讲述了多媒体信息的检索技术。涉及的内容有语音信号、图像信号和视频信号的处理与检索,以及文字图像分离处理等技术。

本书从研究和技术实现的角度出发,既包含较深的理论研究,又包含具体的技术实现过程。具有较高的学术水平和实用价值。可作为高等院校信号与信息处理、通信、模式识别与人工智能、信息管理等专业与学科的研究生教学参考书,也可供这些领域的科研及工程技术人员参考。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,侵权必究。

图书在版编目(CIP)数据

多媒体信息处理与检索技术/朱学芳著. —北京:电子工业出版社,2002. 11
ISBN 7-5053-8309-4

I. 多… II. 朱… III. ①多媒体—信息处理 ②多媒体—情报检索 IV. TP391

中国版本图书馆 CIP 数据核字(2002)第 099340 号

责任编辑:赵丽松 李 玮

印 刷:北京市增富印刷有限责任公司

出版发行:电子工业出版社 <http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编 100036

经 销:各地新华书店

开 本:787×1092 1/16 印张:11.25 字数:288千字 彩插:4

版 次:2002年11月第1版 2002年11月第1次印刷

印 数:3000册 定价:25.00元

凡购买电子工业出版社的图书,如有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系。
联系电话:(010)68279077

TX 37
20



图 4-4



图 5-8

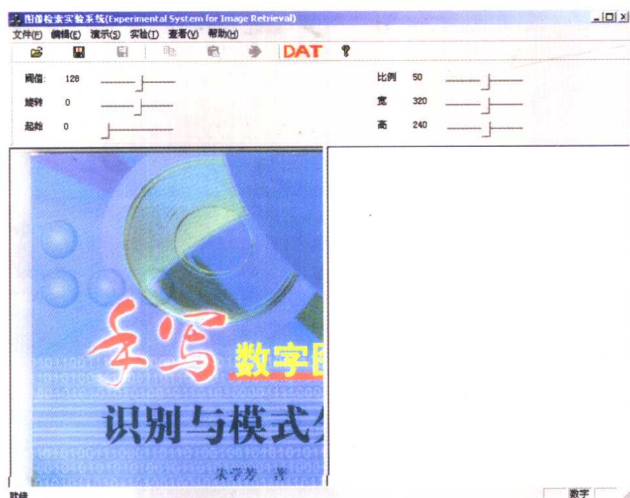


图 5-9



图 5-11

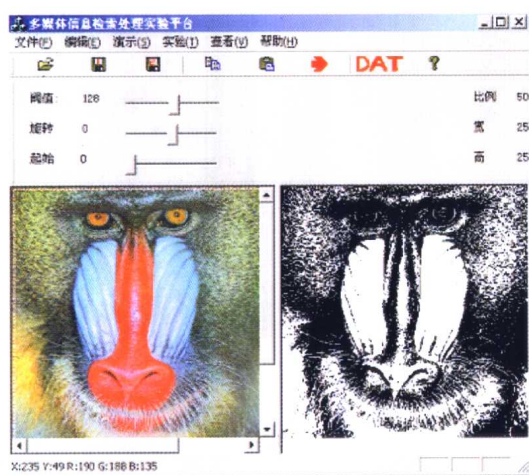


图 5-15

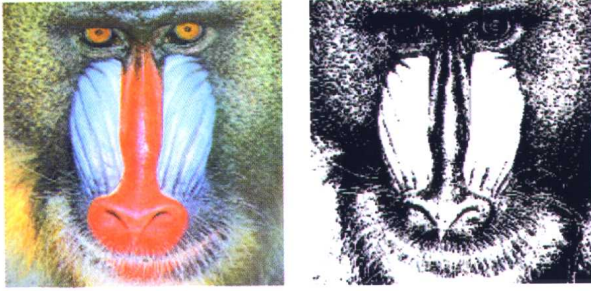


图 5-16

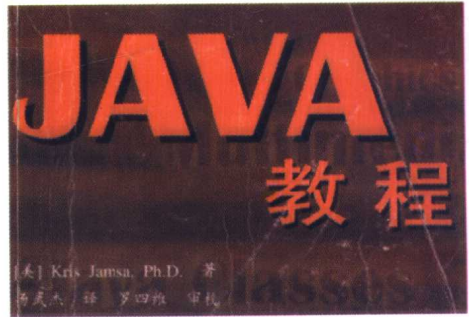


图 6-1

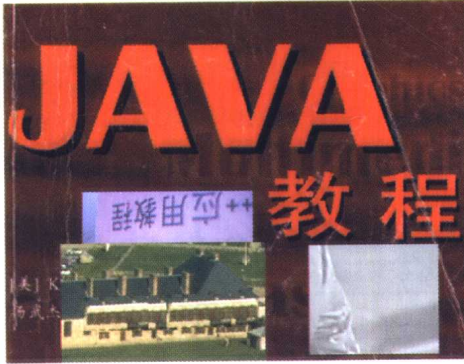


图 6-4



图 6-6



(a)

(b)

图 6-8

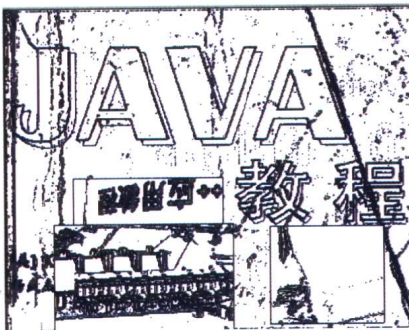


图 6-9



图 6-15

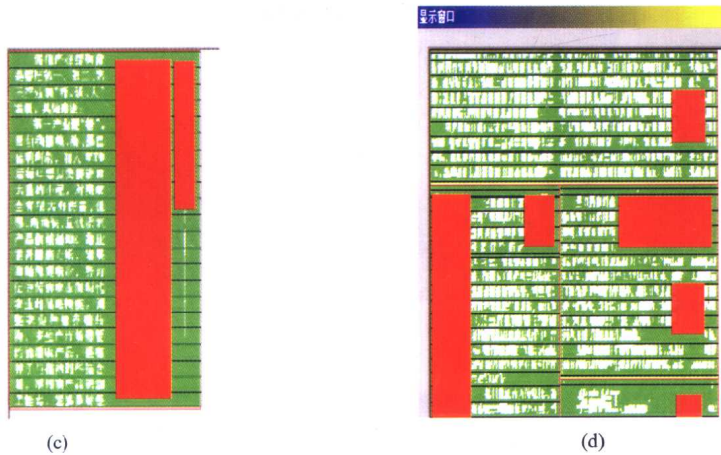
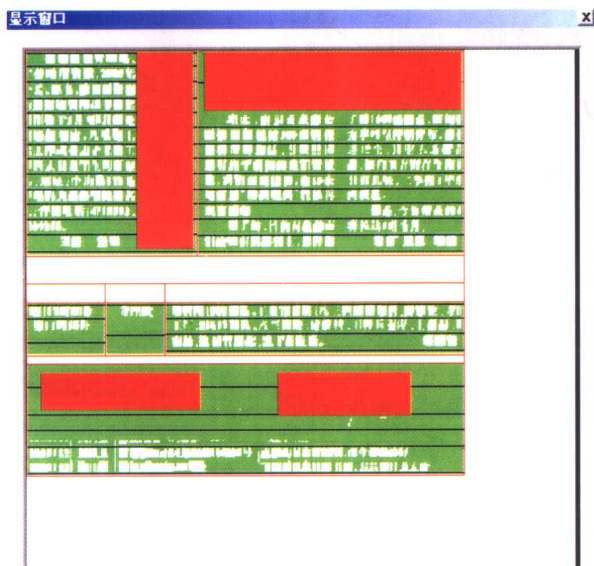
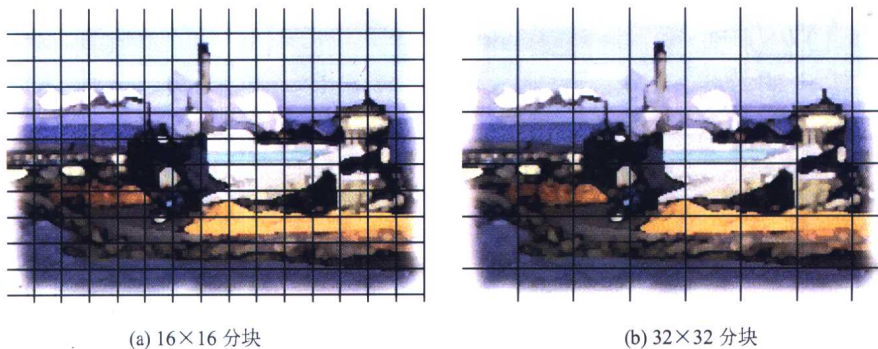


图 6-16



(b)

图 6-17



(a) 16×16 分块

(b) 32×32 分块

图 7-2

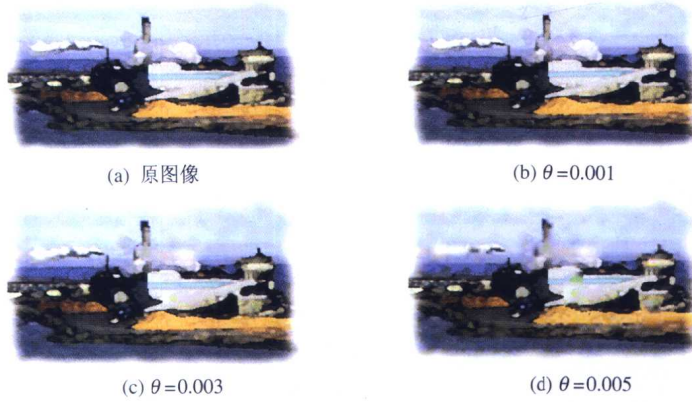


图 7-3

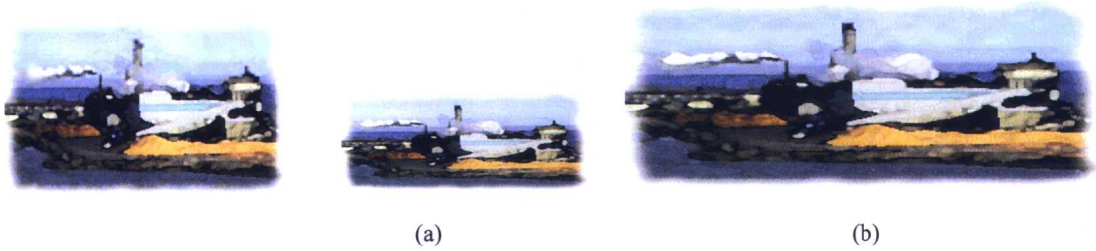


图 7-6

图 7-7

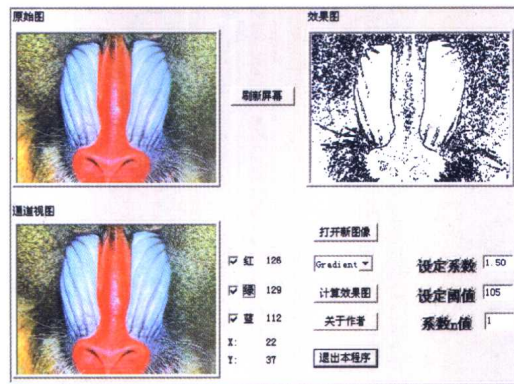


图 8-1



图 8-2



图 8-3

AJS185 / 04

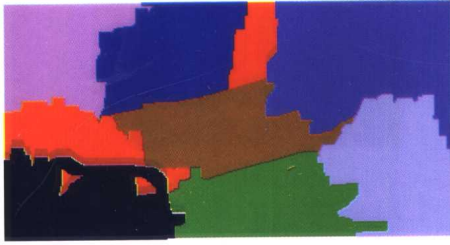


图 8-16

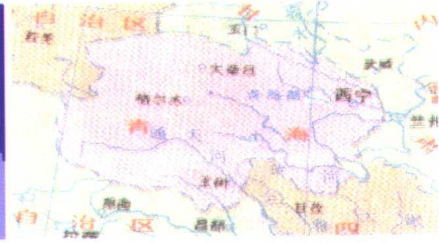
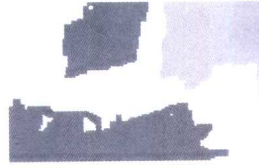


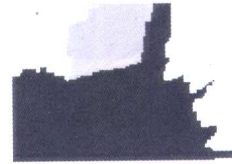
图 8-17



(a) R 基色图



(b)G 基色图

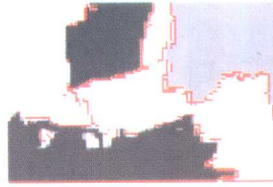


(c)B 基色图

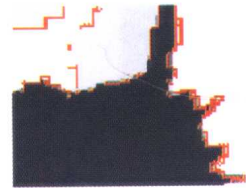
图 8-19



(a) 图 8-19(a)的分割结果



(b) 图 8-19 (b)的分割结果

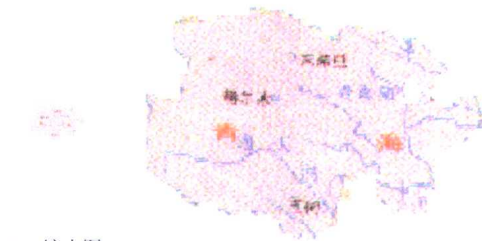


(c) 图 8-19 (c)的分割结果

图 8-20



图 8-24



(a) 缩小图

(b) 放大图

图 8-25



图 9-1



图 9-2

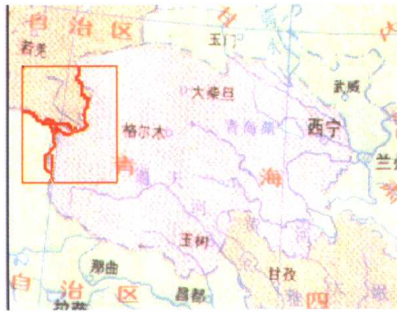


图 8-26

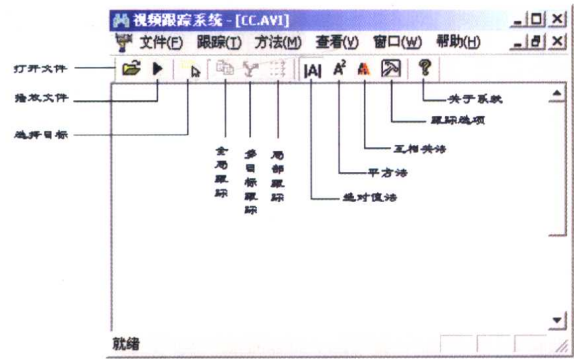


图 10-4

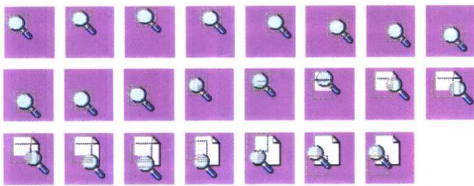


图 10-6

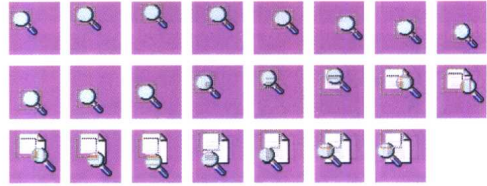


图 10-7

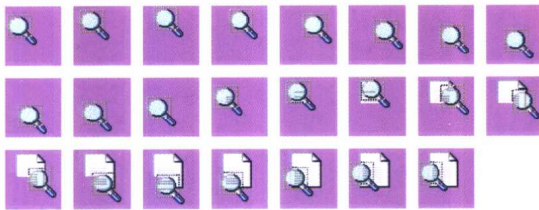


图 10-8

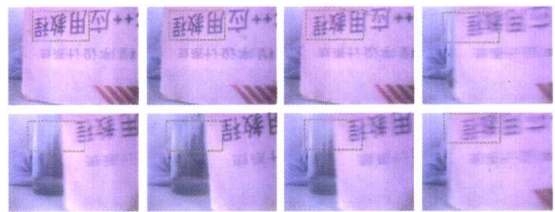


图 10-9



图 10-10

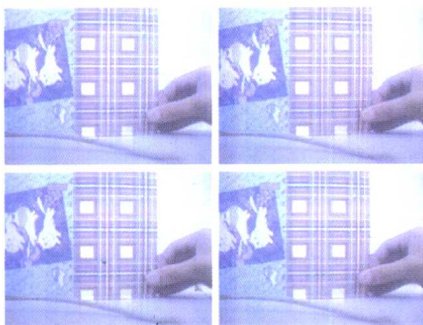
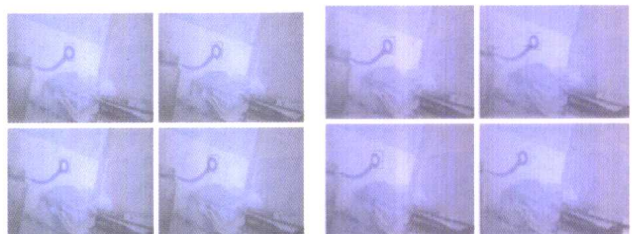


图 10-12



(a) ceA.avi

(b) ceC.avi

图 10-13



图 11-2

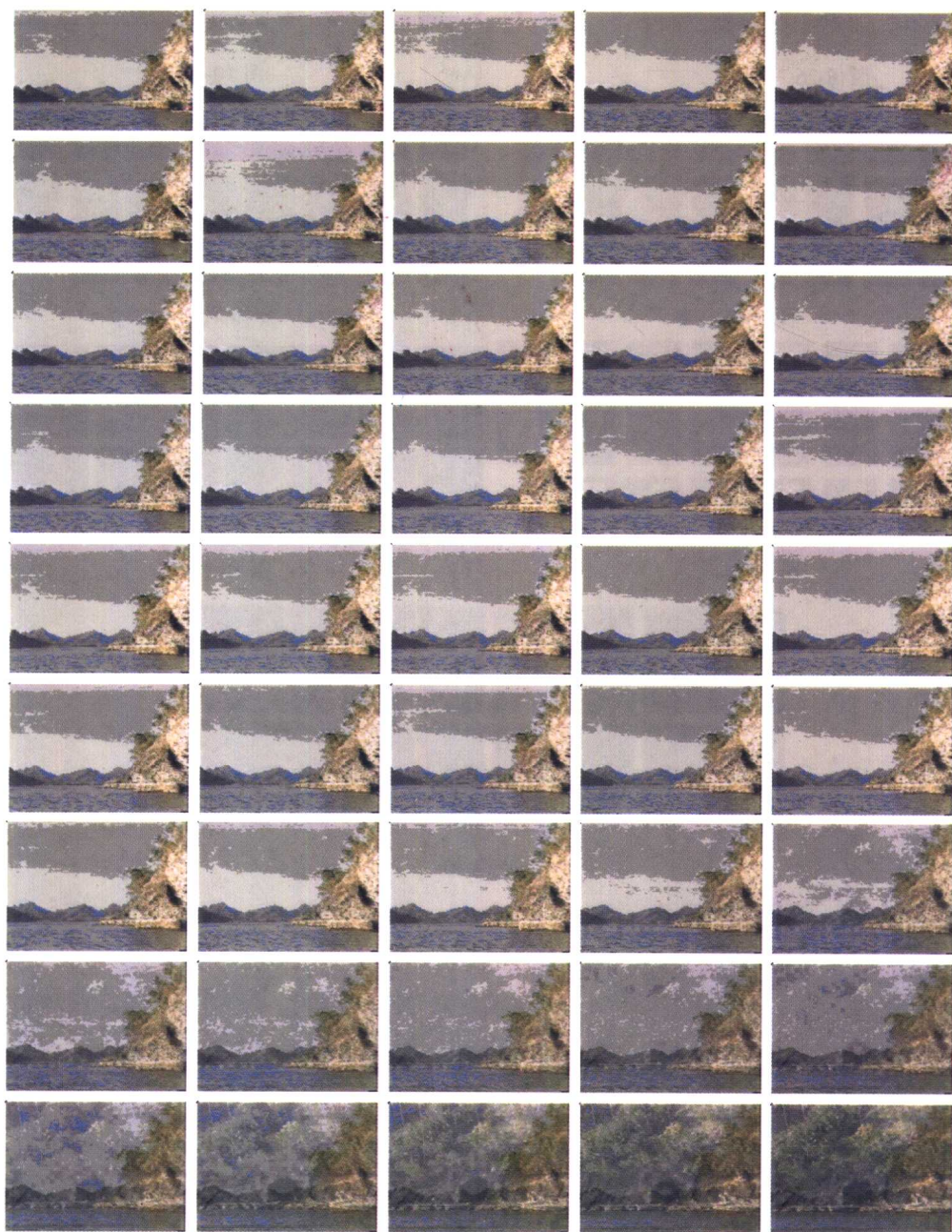


图 11-3

序 言

多媒体信息是指以文字、图像、影像、声音和动画等为表现形式的媒体信息,其含义一般认为是指运用存储与再取技术得到的相关信息,尤其是计算机中的数字信息。由此可见,多媒体信息处理就是指对文字、声音、图形、静态影像、活动影像等多媒体信息在计算机运算下的综合处理。从狭义上讲,就是计算机系统对数字化声音文件、动画视频文件、图形、图像文件等进行的综合处理。

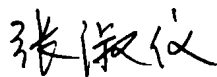
随着大容量数字存储技术、大规模集成电路技术、各种编码技术的发展和国际标准的制定及其实用化产品的涌现,多媒体技术正向多媒体计算机、多媒体通信、多媒体信息处理三大分支迅速发展,其中基于 Internet 的多媒体信息处理被认为是最为重要的发展方向。各种多媒体信息处理及检索技术在诸如远程教育、远程会议、远程医疗、电子商务、文献情报、预定系统、交通管理等多媒体信息系统中有着广阔的应用前景。

从 1984 年美国率先研究多媒体技术以来,该领域一直与其相应的支撑技术及基础研究和应用互相渗透、互相促进。一方面,以语音、图像、视频、动画等为核心对象的基础技术研究和应用有力地推动了现代信息技术的发展,也促进了多媒体技术与系统向实用化方向发展。另一方面,随着对多媒体信息及其技术的应用,人们也向数字语音处理技术、数字图像处理技术、数字视频处理技术等方面提出一些新的研究课题,如语音数据库及其检索、图像数据库及其检索、视频及动画的检索等。尤其是基于内容的检索突破了传统的基于表达式检索的限制,直接对图像、视频、音频内容进行分析,抽取特征和语义,利用这些内容特征建立索引,并进行检索。

由于多媒体信息处理技术的发展,特别是应用于检索技术的迫切需要,有必要对多媒体信息处理与检索技术进行总结。《多媒体信息处理与检索技术》一书的出版正是反映当前国内外相关方面的最新发展及作者近年来的研究成果。

本书在实际研究多媒体信息处理技术的基础上,研究基于内容的多媒体信息的检索技术,包括语音信号的处理与检索、图像信号的处理与检索、视频信号的处理与检索,以及文字图像分离处理等。作者较系统地方法研究及技术实现角度对多媒体信息处理与检索技术加以探讨,侧重于技术实现过程,具有实际应用意义。因此,本书对研究和开发文字、声音、图形、静态影像、活动影像等多媒体信息处理与检索技术有极大的帮助。预计本书的出版将对多媒体信息处理及检索技术的研究与发展起积极推动作用。

中国科学院院士



2002 年 8 月于南京大学

目 录

第 1 章 多媒体检索技术概述	(1)
1.1 基于内容的多媒体数据检索	(1)
1.2 基于内容的检索特点	(2)
1.3 系统构成	(3)
1.4 多媒体数据基于内容查询	(4)
1.4.1 提交查询的方式	(4)
1.4.2 查询类型	(5)
第 2 章 语音处理技术	(7)
2.1 语音信号的时域特性显示	(7)
2.1.1 语音的 WAV 文件格式	(7)
2.1.2 时域特性及表示	(8)
2.1.3 时域特性显示程序	(9)
2.2 自动分段	(11)
2.2.1 自动分段过程	(11)
2.2.2 寂静段、有声段子程序、自动分段结果	(11)
2.3 语音信号的频域显示及声谱差特征	(12)
2.3.1 傅氏变换及语谱图	(12)
2.3.2 语谱图显示程序	(13)
2.3.3 语谱图显示	(13)
2.3.4 声谱差特征	(14)
2.4 语音信号的非均匀采样的实现	(14)
2.4.1 分段信号频率的获取	(15)
2.4.2 重采样的实现	(16)
2.4.3 实验结果及分析	(17)
2.5 语音非均匀重采样值的重构实现技术	(18)
2.5.1 重构的实现技术	(18)
2.5.2 重构实验结果	(19)
第 3 章 音频检索的技术实现	(22)
3.1 音频检索	(22)
3.2 系统设计框架	(23)
3.2.1 基于语音的特征相似度的算法	(23)
3.2.2 基于语音数据相关性的算法	(24)
3.3 数据预处理	(25)
3.3.1 数据预处理的实现	(25)
3.3.2 特征抽取模块	(25)

3.3.3	检索匹配	(25)
3.3.4	波形显示程序	(26)
3.3.5	匹配结果显示	(27)
3.4	基于相关检索的程序设计	(28)
3.4.1	均值和均方值的计算	(28)
3.4.2	检索和匹配	(28)
3.4.3	语音检索系统界面	(30)
第 4 章	图像检索技术	(32)
4.1	图像检索技术概述	(32)
4.1.1	基于图像外部特征的检索方法	(32)
4.1.2	基于文本描述的方法	(32)
4.1.3	传统图像检索方法的局限性	(33)
4.2	新型图像检索技术——基于内容的图像检索技术	(33)
4.2.1	基于内容的图像检索技术简介	(33)
4.2.2	基于内容的图像检索技术概念模型	(33)
4.3	静态图像检索技术	(34)
4.3.1	基于颜色特征的检索	(34)
4.3.2	基于形状特征的检索	(35)
4.3.3	基于纹理特征的检索	(35)
4.3.4	基于空间位置关系特征的检索	(36)
4.3.5	基于对象特征的检索	(36)
4.4	基于内容的图像检索系统	(36)
4.4.1	基于内容的图像检索系统效果评价	(36)
4.4.2	基于内容的图像检索典型系统介绍	(37)
4.5	目前图像检索技术研究的热点问题	(38)
4.5.1	两类图像检索技术的结合	(38)
4.5.2	对基于内容编码技术的研究	(38)
4.5.3	对用户查询接口的研究	(38)
4.5.4	对图像数据库的研究	(38)
4.6	图像检索实例	(39)
4.6.1	基于色彩直方图的检索实例	(39)
4.6.2	基于区域形状的检索实例	(41)
第 5 章	多媒体信息检索平台	(43)
5.1	Windows 环境下的图像处理	(43)
5.1.1	位图操作的函数简介	(43)
5.1.2	BMP 图像的读取	(45)
5.1.3	位图的显示	(45)
5.1.4	图像数据的数组化获取	(46)
5.2	平台界面及功能介绍	(46)
5.2.1	总体功能介绍	(46)

5.2.2	程序界面及功能介绍	(48)
5.3	程序的实现代码	(52)
5.3.1	BMP 图像文件和位图的处理	(52)
5.3.2	平台界面的部分类介绍	(54)
5.3.3	界面相关功能的实现	(63)
第 6 章	版面分析	(74)
6.1	彩色图像特性统计分析	(74)
6.2	编程实现过程	(76)
6.2.1	程序框架	(76)
6.2.2	图像的处理过程及中间结果	(76)
6.3	文字分离实验结果	(79)
6.3.1	无表格情况时的分离结果	(79)
6.3.2	有表格情况时的分离结果	(80)
6.4	流程图及部分源程序	(81)
6.4.1	主流程图	(81)
6.4.2	主流程图对应的源程序	(81)
6.4.3	设置输出矩形框的流程图	(85)
6.4.4	基于直方图的分析流程图	(86)
6.4.5	基于直方图的分析源程序	(86)
6.4.6	后处理流程图	(89)
第 7 章	彩色图像放大重构	(91)
7.1	彩色图像重采样方法	(91)
7.2	彩色图像的重构	(92)
7.3	彩色图像重采样、重构实验	(93)
7.4	Hermite 插值技术实现图像放大	(95)
7.4.1	两点三次 Hermite 插值	(96)
7.4.2	图像信号的重构与放大显示	(97)
第 8 章	图像分割	(99)
8.1	图像分割的一般概念	(99)
8.1.1	图像分割的一般模型	(99)
8.1.2	图像分割的一般算法	(100)
8.1.3	图像分割的方法特点	(100)
8.2	图像边缘提取	(101)
8.2.1	微分算子法	(102)
8.2.2	梯度算子	(103)
8.2.3	Sobel 算子	(103)
8.2.4	Kirsh 算子	(104)
8.2.5	Laplacian 算子	(104)
8.2.6	Prewitt 算子	(104)
8.2.7	Marr 算子	(105)

8.3	分割系统的程序设计	(105)
8.3.1	采用区域中灰度范围分割	(106)
8.3.2	基于子区域的区域增长算法	(106)
8.3.3	待分割图像的预处理	(107)
8.3.4	程序结果后处理	(108)
8.4	软件系统的编程环境和框架	(108)
8.4.1	CDIB 类	(109)
8.4.2	CMainFrame 类	(111)
8.4.3	CSplitmapDoc 类	(112)
8.4.4	CSplitMapView 类	(114)
8.4.5	Process.cpp 文件	(116)
8.4.6	主要函数流程图	(117)
8.5	分割结果	(119)
8.5.1	系统主界面	(119)
8.5.2	直方图	(119)
8.5.3	模拟彩色图分割	(120)
8.5.4	实际彩色图分割	(121)
8.5.5	区域图的处理	(121)
8.5.6	抽样分割	(123)
第 9 章	视频信号处理基础	(124)
9.1	视频信号格式基础	(124)
9.1.1	RIFF 文件	(124)
9.1.2	视频交互文件 AVI 格式	(125)
9.1.3	AVI 文件头	(126)
9.1.4	对 AVI 格式的说明补充	(131)
9.2	播放 AVI 的程序实现	(131)
9.2.1	MCIWnd	(131)
9.2.2	程序实现	(132)
9.3	接入摄像头的程序实现	(133)
9.4	缩小 AVI 的程序实现	(134)
9.5	视频放大的程序实现	(136)
第 10 章	目标图像跟踪	(147)
10.1	简介	(147)
10.1.1	数字图像的处理和匹配技术	(147)
10.1.2	彩色模型	(147)
10.2	现有算法	(148)
10.2.1	模板匹配法	(148)
10.2.2	其他匹配算法	(149)
10.2.3	快速匹配算法	(149)
10.2.4	关于彩色图像	(149)

10.3	系统介绍	(150)
10.3.1	系统的主要过程框图	(150)
10.3.2	设计思路	(152)
10.3.3	系统界面	(154)
10.4	实验及结果比较	(155)
10.4.1	动画	(155)
10.4.2	自然摄制图像	(156)
10.4.3	多目标跟踪	(157)
10.4.4	压缩图像	(158)
10.4.5	算法的执行时间	(159)
10.4.6	跟踪中的漂移现象	(159)
第 11 章	视频流浏览与查询	(160)
11.1	视频特征分析	(160)
11.1.1	视频基础知识	(160)
11.1.2	视频特征	(160)
11.2	镜头分割	(161)
11.3	代表帧	(162)
11.4	视频信息的查询	(165)
11.4.1	基于内容检索的特点	(165)
11.4.2	基于内容的视频检索模块	(166)
11.4.3	视频检索类型	(166)
参考文献		(169)
后记		(170)