

目 录

第 1 篇 绪论

第 1 章	引言	1
1.1	为什么需要基于内容的视觉信息检索	1
1.1.1	视觉信息的快速增长	1
1.1.2	视觉信息检索	2
1.1.3	传统视觉信息检索的局限	3
1.2	什么是基于内容的视觉信息检索	4
1.2.1	基于内容视觉信息检索的特点	4
1.2.2	基于内容的视觉信息检索示例	4
1.3	如何实现基于内容的视觉信息检索	6
1.3.1	归档和检索流程图	6
1.3.2	查询和检索需解决的问题	6
1.3.3	所需的基础知识	7
1.4	基于内容视觉信息检索的发展	7
1.4.1	第一代视觉信息检索系统	7
1.4.2	新一代视觉信息检索系统	8
1.4.3	目前状况和回顾	9
1.5	本书概要	9
1.5.1	整体框架安排	9
1.5.2	各章概要	10
	参考文献	12
第 2 章	基于内容图象检索研究的基本技术	15
2.1	基本工作框架和功能模块	15
2.1.1	基本框架	15
2.1.2	模块功能简介	15
2.2	图象检索技术	16
2.2.1	查询模块和技术	16
2.2.2	描述模块和技术	18
2.2.3	匹配模块和技术	20
2.2.4	提取模块和技术	22
2.2.5	验证模块和技术	22
2.3	视频检索技术	23
2.3.1	视频查询和访问	23
2.3.2	视频组织和描述	24
2.3.3	视频匹配要点	25
2.3.4	视频提取和特点	25
2.3.5	对返回视频的验证	26
2.4	检索涉及的领域和技术	27
	参考文献	27
第 3 章	检索系统及其应用领域	29
3.1	典型的图象检索系统	29
3.1.1	图象检索系统功能和概况	29
3.1.2	QBIC	30
3.1.3	计算机辅助人面像推断和检索系统	31

3.2	典型的视频检索系统	32
3.2.1	视频检索系统特点和概况	32
3.2.2	面向对象的视频信息数据库	32
3.2.3	视频索引和检索原型系统	33
3.2.4	代数视频系统	34
3.3	系统性能指标和评价准则	35
3.4	基于内容视觉信息检索的应用领域	37
	参考文献	38
第 2 篇 基于内容的图象检索		
第 4 章	基于颜色特征的图象检索	41
4.1	颜色模型	41
4.1.1	面向硬设备的颜色模型	41
4.1.2	面向视觉感知的颜色模型	43
4.1.3	均匀颜色空间模型	46
4.2	颜色特征的表达	48
4.2.1	统计直方图	48
4.2.2	累积直方图	49
4.2.3	颜色布局	49
4.3	颜色匹配算法	50
4.3.1	简单直方图匹配方法	50
4.3.2	累积直方图及优越性	52
4.4	对 10 种基于颜色特征检索方法的比较	53
4.4.1	10 种检索算法及编号	53
4.4.2	实验结果	54
4.4.3	结果分析和讨论	56
4.4.4	HVC 和 HSI 彩色空间的比较	56
4.5	局部累加直方图方法	57
4.5.1	局部累加直方图	57
4.5.2	相似色区间的划分	58
4.5.3	基于局部累加直方图的 4 种检索算法	61
4.5.4	算法性能检验及结果	61
	参考文献	62
第 5 章	基于纹理特征的图象检索	64
5.1	纹理描述模型	64
5.1.1	纹理概述	64
5.1.2	基于空间性质的纹理模型	65
5.1.3	基于频率性质的纹理模型	66
5.1.4	基于结构感知性质的纹理模型	67
5.2	统计法纹理描述	67
5.2.1	基本纹理描述符	67
5.2.2	分形模型	70
5.3	频谱法纹理描述	71
5.3.1	基于傅里叶变换的纹理描述	71
5.3.2	基于小波变换的纹理描述	72
5.3.3	国际标准 MPEG-7 推荐的纹理描述符	72
5.4	结构法纹理描述	73
5.5	基于纹理的检索方案	74
5.6	利用纹理特征的检索算法示例	76
	参考文献	77

第 6 章	基于形状特征的图象检索	79
6.1	形状描述	79
6.1.1	形状描述概述	79
6.1.2	基于外部参数的形状描述	80
6.1.3	基于内部参数的形状描述	83
6.1.4	基于变形的形状描述	84
6.2	基于形状的检索	85
6.2.1	基于特征的方法	86
6.2.2	基于变换域的方法	88
6.2.3	基于变形的的方法	89
6.2.4	草图查询	90
6.3	小波轮廓描述符	91
6.3.1	小波轮廓描述符的定义	91
6.3.2	小波轮廓描述符的性质	93
6.3.3	相对傅里叶轮廓描述符的一些优越性	94
6.4	小波模极大值和多尺度不变矩方法	96
6.4.1	小波模极大值	96
6.4.2	算法要点	98
6.4.3	实验及结果	99
	参考文献	102
第 7 章	基于空间关系的图象检索	105
7.1	空间关系概述	105
7.2	基于目标结构的表达	107
7.2.1	网格	107
7.2.2	四叉树	108
7.2.3	二叉树	108
7.2.4	K - d 树	109
7.2.5	R -树	109
7.3	基于关系的表达	110
7.3.1	符号投影和 2-D 串表达	110
7.3.2	2-D 串族	112
7.3.3	变换下不变的表达	114
7.4	度量关系	116
7.4.1	空间距离测量	117
7.4.2	空间朝向测量	117
7.5	利用空间关系查询和检索	119
7.5.1	使用文字查询和检索	119
7.5.2	利用范例的图标查询和检索	120
7.5.3	借助草图查询和检索	122
7.6	利用子区域向量匹配的空间关系检索	123
7.6.1	算法原理和步骤	123
7.6.2	算法应用示例——篮球站位分析系统	124
7.6.3	算法性能实验及结果分析	125
	参考文献	127
第 8 章	综合特征检索	129
8.1	不同特征的特点	129
8.1.1	不同特征的自身特点	129
8.1.2	不同特征的相对特点	130
8.2	综合特征	131

8.2.1	特征的综合使用	131
8.2.2	特征归一化	131
8.3	结合颜色和纹理的检索	133
8.3.1	颜色的空间布局	133
8.3.2	综合利用颜色和纹理特征进行检索	133
8.4	结合颜色和形状的检索	134
8.4.1	方法介绍	134
8.4.2	试验结果	135
8.5	结合纹理和形状的检索	136
8.5.1	边缘点数直方图	136
8.5.2	边缘直方图描述符	137
8.5.3	使用纹理和形状的组合测度	138
8.6	结合空间关系的检索	139
	参考文献	142
第 9 章	图象检索实验平台和系统	143
9.1	基于特征的图象查询和检索系统	143
9.1.1	系统模块和功能	143
9.1.2	系统查询实验结果	145
9.2	基于特征的算法试验平台	146
9.2.1	设计原则和功能要求	146
9.2.2	平台主要模块	147
9.2.3	平台有特色的检索功能	149
9.3	基于服务器和客户端的网上图象查询系统	151
9.3.1	系统设计	151
9.3.2	系统流程和分析	153
9.3.3	系统界面	155
9.4	WWW 上图象浏览检索系统	156
9.4.1	系统组成及功能	156
9.4.2	系统实现	156
9.4.3	系统用户界面特点	157
	参考文献	158
第 3 篇 基于内容的视频检索		
第 10 章	视频数据库模型和管理	160
10.1	视频和视频数据库特点	160
10.2	数据模型	161
10.2.1	实体-联系模型	162
10.2.2	语义对象模型	162
10.3	视频数据模型	163
10.3.1	时间线模型	163
10.3.2	时间层次模型	164
10.3.3	代数模型	165
10.3.4	视频对象数据模型	166
10.4	数据库结构和模型	167
10.4.1	数据库结构	167
10.4.2	数据库模型	167
10.5	对视频库的管理和检索	169
10.5.1	视频数据库管理	169
10.5.2	视频库检索的一种结构化框架	170
10.5.3	关于视频库检索的讨论	171

参考文献	172
第 11 章 镜头检测和表达	174
11.1 视频序列的时域分割——镜头检测	174
11.2 切变检测	176
11.2.1 基于全局特征的变化检测	176
11.2.2 基于局部特征的变化检测	179
11.2.3 根据变化和不相似性确定切变	179
11.2.4 基于双重窗口的切变检测算法	181
11.3 渐变检测	185
11.3.1 渐变类型和检测次序	185
11.3.2 淡入淡出和叠化的检测	186
11.3.3 擦除和糙化的检测	189
11.3.4 基于模型的淡入淡出和叠化检测	190
11.4 不同镜头检测算法的性能比较和讨论	193
11.4.1 性能比较测度	193
11.4.2 对检测方法的讨论	194
11.5 镜头关键帧	195
11.5.1 关键帧	195
11.5.2 一种简单有效的关键帧提取算法	196
参考文献	198
第 12 章 镜头聚类 and 镜头集合描述	200
12.1 镜头聚类	200
12.1.1 镜头集合体	200
12.1.2 不随视频类型变化的镜头聚类方法	201
12.1.3 针对特定视频类型的镜头聚类方法	202
12.2 场景转换图	205
12.3 逻辑故事单元	206
12.3.1 情节和逻辑故事单元	206
12.3.2 确定逻辑故事单元的边界	208
12.4 情节代表帧的选取	208
12.4.1 基本步骤	208
12.4.2 讨论和改进	209
12.4.3 效果和情节层描述	210
参考文献	212
第 13 章 视频数据组织和索引	214
13.1 视频数据组织	214
13.1.1 组织概述	214
13.1.2 视频分段	215
13.1.3 组织实例	216
13.2 数据库索引概述	217
13.3 基本的数据库索引方法	218
13.3.1 Hash 索引	218
13.3.2 B -树索引和 B^+ -树索引	219
13.3.3 R -树索引	220
13.3.4 借助三角不等式的索引	220
13.4 基于视觉特征的索引	221
13.4.1 场景目标运动特征	221
13.4.2 摄像机运动特征	222
13.4.3 元数据特征	223
13.5 基于语义特征的索引和注释	223

13.5.1	语义特征和摘要	223
13.5.2	领域知识	224
13.5.3	视频注释	225
	参考文献	226
第 14 章	基于运动特征的视频检索	228
14.1	运动轨迹	228
14.1.1	运动轨迹检测	228
14.1.2	运动轨迹描述符	229
14.2	运动特征的提取和描述	230
14.2.1	短时运动特征和描述	230
14.2.2	全局摄像机运动模型	231
14.2.3	局部运动矢量场	233
14.2.4	运动活力描述符	234
14.3	基于双线性模型的全局运动矢量检测	235
14.4	局部运动矢量检测和视频目标分割	236
14.4.1	计算局部运动矢量场的 Horn-Schunck 算法	236
14.4.2	带全局运动补偿的改进 Horn-Schunck 算法	237
14.4.3	局部运动模型和视频目标分割	238
14.5	基于运动特征的视频检索	239
14.5.1	两类检索方案	239
14.5.2	基于全局运动特征的检索方案	240
14.5.3	基于局部运动特征的检索方案	242
	参考文献	244
第 15 章	视频节目查询应用	247
15.1	视频查询	247
15.1.1	查询层次及特征	247
15.1.2	查询工具	248
15.1.3	查询语言	249
15.2	视频描述	250
15.2.1	基于手工标记的文字描述	250
15.2.2	基于自动提取语义的内容描述	250
15.3	视频节目摘要	252
15.3.1	视频摘要	252
15.3.2	家庭录象摘要	253
15.3.3	新闻摘要示例	254
15.4	相似测度和评价	256
15.4.1	对称的相似测度	256
15.4.2	非对称的相似测度	257
15.4.3	检索效果和效率	258
15.5	各种视频节目的查询检索特点	259
15.5.1	基于应用生成的模型	260
15.5.2	基于领域知识的索引	260
15.5.3	典型的使用群体及他们的查询要求	263
	参考文献	265
第 16 章	用户界面和浏览显示	267
16.1	界面功能和布局	267
16.1.1	界面功能	267
16.1.2	界面布局示例	268
16.2	视频片段浏览和显示	269

16.2.1	可视化关键帧	269
16.2.2	可视化视频片段	270
16.2.3	可视化视频结构	272
16.2.4	基于场景的显示	273
16.3	运动边缘层叠图	275
16.3.1	算法流程	275
16.3.2	全局运动估计	275
16.3.3	提取和叠加运动边缘	280
	参考文献	282

第 4 篇 热点和趋势

第 17 章	压缩域检索	284
17.1	压缩域检索的特点和方法分类	284
17.1.1	压缩域检索的特点	284
17.1.2	压缩域检索方法分类	285
17.2	对变换域压缩图象的检索	186
17.2.1	对离散傅里叶变换图象的检索	286
17.2.2	对离散余弦变换图象的检索	287
17.2.3	对小波变换图象的检索	287
17.2.4	对小波域子带编码图象的检索	288
17.3	对空域压缩图象的检索	289
17.3.1	对矢量量化图象的检索	289
17.3.2	对分形编码图象的检索	290
17.3.3	对预测编码图象的检索	290
17.4	压缩域的视频检索	292
17.4.1	压缩域的切变检测	292
17.4.2	压缩域的渐变检测	293
	参考文献	294
第 18 章	网上搜索引擎和分类检索	297
18.1	网上图象搜索引擎	297
18.1.1	搜索引擎的分类	297
18.1.2	基于“巡查员”的搜索引擎的工作原理	298
18.1.3	图象搜索引擎框架和模块	298
18.2	图片分类和筛选	299
18.2.1	图片分类	299
18.2.1	图片筛选	300
18.3	网上 GIF 格式图片的筛选	302
18.3.1	图片预筛选	302
18.3.2	基于数据量比例的筛选算法	303
18.3.3	基于颜色统计的筛选算法	305
18.3.4	两种筛选算法的结合使用	306
18.4	图象分类	306
18.4.1	最近特征直线分类算法	306
18.4.2	基于关联规则的图象分类	308
18.5	基于特征元素和关联规则的图象分类算法	310
18.5.1	特征元素的提取	310
18.5.2	关联规则的挖掘算法	312
18.5.3	分类实验结果比较	313
18.6	基于特征元素的图象检索	314
	参考文献	315

第 19 章 基于语义的检索	317
19.1 语义模型	317
19.1.1 语义表达和索引模型	317
19.1.2 图象语义层次模型	319
19.2 语义提取和描述	320
19.2.1 语义视觉模板	320
19.2.2 渐进式框架	322
19.3 图象有意义区域提取	324
19.3.1 流程框图	324
19.3.2 有意义区域提取方法	325
19.3.3 实验结果与讨论	329
19.4 多级图象语义内容描述方案	330
19.4.1 多级描述模型	330
19.4.2 目标层描述中的目标识别	332
19.4.3 目标层描述中的关系表述	334
19.5 基于目标层描述的检索系统	335
19.5.1 系统概述	335
19.5.2 基于目标的查询	337
19.5.3 基于空间关系的查询	337
参考文献	338
第 20 章 检索中的交互反馈	340
20.1 反馈的必要性	340
20.2 相关反馈	341
20.2.1 相关反馈特点	341
20.2.2 正例和反例	342
20.2.3 基于权重调整的相关反馈方法	343
20.3 基于目标的自适应反馈	344
20.3.1 自适应反馈	345
20.3.2 无记忆反馈	346
20.3.3 有记忆反馈	346
20.4 关联反馈	347
20.4.1 背景基础	347
20.4.2 关联反馈	349
参考文献	350
第 21 章 检索性能评价	352
21.1 相似性测度	352
21.1.1 测度公理	352
21.1.2 相似测度	353
21.1.3 测度的比较	354
21.1.4 距离函数	355
21.2 系统性能评价	356
21.2.1 系统性能指标	356
21.2.2 评价方法	359
21.2.3 系统的评价	360
21.3 评价基于内容检索的系统	362
参考文献	364
第 22 章 研究展望和应用前景	365
22.1 研究趋势	365
22.1.1 研究方向讨论	365
22.1.2 若干有新概念的研究思路	368

	22.1.3 情感层次的检索	370
	22.2 应用前景	371
	22.2.1 应用领域的拓展	371
	22.2.2 一些有特色的应用系统示例	371
	参考文献	374
附录 A	国际标准 MPEG-7 介绍	376
	A.1 国际标准 MPEG-7	376
	A.1.1 MPEG-7 是什么	376
	A.1.2 MPEG-7 与其它 MPEG 标准有什么联系	376
	A.1.3 MPEG-7 与基于内容的多媒体信息检索有什么联系	377
	A.2 MPEG-7 概况和内容	378
	A.2.1 MPEG-7 标准制订的方式是怎样的	378
	A.2.2 MPEG-7 的工作进展情况如何	378
	A.2.3 MPEG-7 主要用到哪些概念	379
	A.2.4 MPEG-7 主要标准化哪些内容	379
	A.2.5 MPEG-7 有哪些可能的应用和应用方式	380
	A.3 MPEG-7 描述工具及特点	381
	A.3.1 MPEG-7 标准分几个部分	381
	A.3.2 MPEG-7 标准中有哪些描述工具	381
	A.3.3 MPEG-7 中的多媒体描述方案是什么样的	382
	A.4 MPEG-7 描述符和描述方案	383
	A.4.1 MPEG-7 考虑哪些特征	383
	A.4.2 MPEG-7 共有哪些描述符	383
	A.4.3 MPEG-7 中 D, DS, DDL 的关系是什么样的	384
	参考文献	385
附录 B	有关国际刊物, 会议, 网站	386
	B.1 有关国际刊物	386
	B.2 有关国际会议	387
	B.3 有关网站和网页	387