



缩略语

<http://media.hust.edu.cn>

- ❖ 视频处理与检索
- ❖ Video Processing & Retrieval –VPR
- ❖ 基于内容的视频检索
- ❖ Content-Based Video Retrieval-CBVR

华中科技大学数字媒体实验室

教学要求

<http://media.hust.edu.cn>

- ❖ 系统掌握数字视频处理与检索的主要概念、基本原理和典型方法
- ❖ 熟悉各种数字视频处理和检索的常用方法
- ❖ 结合实践掌握视频处理的常用软件和工具的使用方法，能够设计和开发简单的视频处理和检索程序
- ❖ 了解国际上相关的新成果和新动向

华中科技大学数字媒体实验室

前续课程

<http://media.hust.edu.cn>

- ❖ 多媒体计算机技术
- ❖ 数字信号处理
- ❖ 数字图像处理

华中科技大学数字媒体实验室

课程内容 (32学时, 2学分)

<http://media.hust.edu.cn>

❖ 引 言	(2学时)
❖ 数字视频基础	(2学时)
❖ 数字视频的压缩与编码	(8学时)
❖ 数字视频的特征提取	(6学时)
❖ 数字视频分析与处理	(4学时)
❖ 数字视频的组织与索引	(4学时)
❖ 基于内容的视频检索	(3学时)
❖ 研究方向及研究热点	(1学时)

华中科技大学数字媒体实验室

主要参考书目

<http://media.hust.edu.cn>

- ❖ 周洞汝, 视频数据库管理系统导论, 科学出版社, 2002
- ❖ 章毓晋, 基于内容的视觉信息检索, 科学出版社, 2003
- ❖ 章毓晋, 图象分析 (第2版), 清华大学出版社, 2005



华中科技大学数字媒体实验室

主要参考书目 (续)

<http://media.hust.edu.cn>

- ❖ Yao Wang, Jorn Ssternmann著, 侯正信, 杨喜, 王文全等译, 视频处理与通信, 电子工业出版社, 2003.
- ❖ 冈萨雷斯著, 阮秋琦、阮宇智等译, 数字图像处理 (第二版), 电子工业出版社, 2003



华中科技大学数字媒体实验室

主要参考书目 (续)

<http://media.hust.edu.cn>

- ❖ Richardo Beeza-Yates, 现代信息检索 (英文版), 机械工业出版社, 2005
- ❖ 林福宗编著, 多媒体技术基础, 清华大学出版社, 2002. 2



华中科技大学数字媒体实验室

常用的开发工具

<http://media.hust.edu.cn>

- ❖ DirectShow SDK
- ❖ Intel OpenCV Library
- ❖ Intel IPP (Integrated Performance Primitive)
- ❖ Matlab Video Processing BlockSet

华中科技大学数字媒体实验室

成绩评定办法:

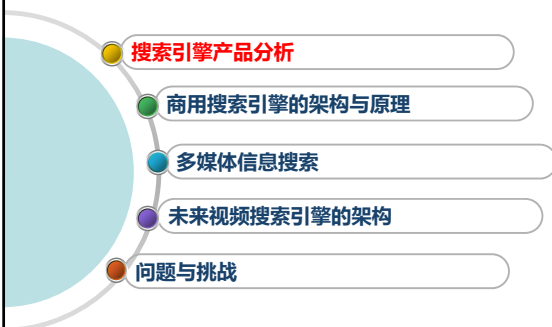
<http://media.hust.edu.cn>

- ❖ Presentation: 20%
- ❖ Project: 40%
- ❖ 笔试: 40%

华中科技大学数字媒体实验室

第一讲 引言

<http://media.hust.edu.cn>



华中科技大学数字媒体实验室

从珍珠翡翠白玉汤说起

<http://media.hust.edu.cn>

- ❖ 珍珠翡翠白玉汤 = 焦糊锅巴+菠菜叶子+白菜帮子+刷锅水
 - ❖ 同一碗汤, 同一个食客, 为什么感受如此不同?
- 破庙中——解除饥渴
金殿上——享受美味



华中科技大学数字媒体实验室

回力鞋的脱胎换骨

http://media.hust.edu.cn



25人民币 VS 50欧元



魔戒王子—奥兰多·布鲁姆

华中科技大学数字媒体实验室

什么是产品？

http://media.hust.edu.cn

❖ 定义

- 向市场提供的用以满足特定需求的有形或无形利益。

❖ 构成

- 核心产品——**提供给用户的核心利益**
- 形式产品——核心产品的外部表现特征

四大金刚——吃饱

军大衣——保暖

满汉全席——排场

比基尼——性感

华中科技大学数字媒体实验室

什么是搜索引擎？

http://media.hust.edu.cn

❖ 搜索引擎(search engine)

- 指根据一定的策略、运用特定的计算机程序搜集互联网上的信息，在对信息进行组织和处理后，为用户提供检索服务的系统。
 - 摘自百度百科：
<http://baike.baidu.com/view/1154.htm>

❖ 搜索引擎的核心价值

- 让人们最便捷地获取信息，找到所求

华中科技大学数字媒体实验室

Outline

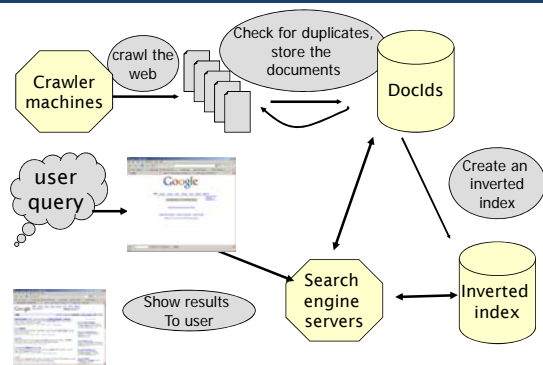
http://media.hust.edu.cn



华中科技大学数字媒体实验室

搜索引擎体系结构

http://media.hust.edu.cn



华中科技大学数字媒体实验室

典型的全文搜索引擎

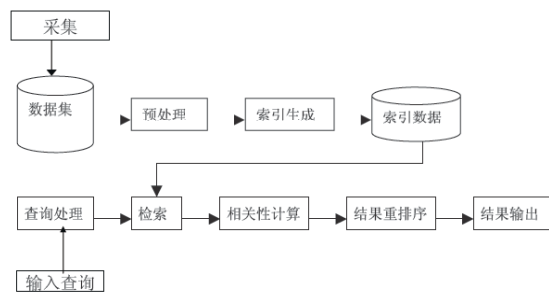
http://media.hust.edu.cn

- ❖ 采集器
- ❖ 分析器
- ❖ 索引器
- ❖ 检索器
- ❖ 人机接口
- ❖ 索引数据库

华中科技大学数字媒体实验室

搜索引擎流程

http://media.hust.edu.cn



华中科技大学数字媒体实验室

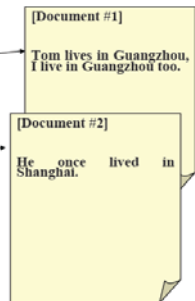
搜索引擎流程

http://media.hust.edu.cn

小型搜索引擎

④ 过程模拟

■ 原始文本



华中科技大学数字媒体实验室

搜索引擎流程

http://media.hust.edu.cn

④ 预处理之后的文本

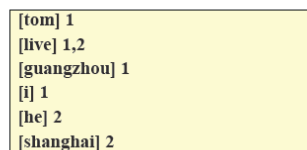


华中科技大学数字媒体实验室

搜索引擎流程

http://media.hust.edu.cn

④ 倒排索引

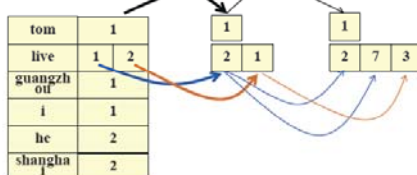


华中科技大学数字媒体实验室

搜索引擎流程

http://media.hust.edu.cn

④ 生成索引文件

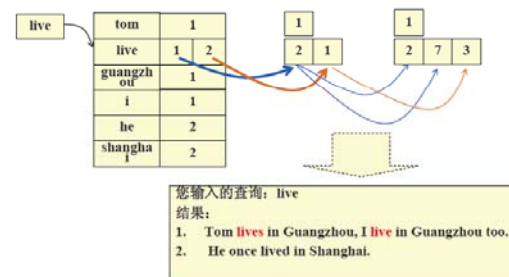


华中科技大学数字媒体实验室

搜索引擎流程

http://media.hust.edu.cn

④ 检索



华中科技大学数字媒体实验室

Outline

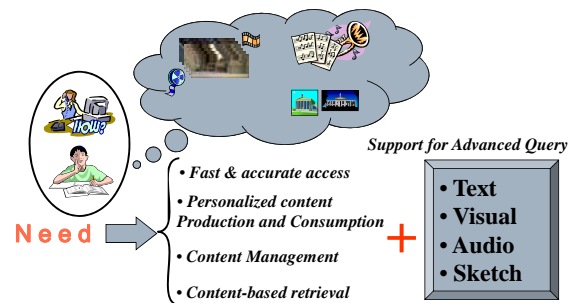
<http://media.hust.edu.cn>

- 搜索引擎产品分析
- 商用搜索引擎的架构与原理
- 多媒体信息搜索**
- 未来视频搜索引擎的架构
- 问题与讨论

华中科技大学数字媒体实验室

Multimedia Search

<http://media.hust.edu.cn>



华中科技大学数字媒体实验室

What is Multimedia?

❖ Definition from the ACM Special Interest Group on Multimedia Retreat in 2003

- More than one media (text, images, audio, video) that are correlated
- Examples:
 - Time correlated: Video with text transcript of the audio
 - Spatially correlated: Images on a page with associated text

❖ A less strict definition:

❖ Not "Just" Text

- Images
- Audio
- Video

Multimedia Search Outline

- ❖ Text Search
 - Keywords
- ❖ Image Search
 - Search based on tags (Flickr, FaceBook)
 - Search based on surrounding text (Google)
 - Content based search
 - Using image features
 - Using faces
- ❖ Audio Search
 - Search based on metadata (iTunes)
 - Content based search
- ❖ Video Search
 - Search based on text (Google/UTube)
 - Search based on associated media (Lectures with slides)
 - Search based on content (TrecVid News Search)



MediaMagic

Text Search

- ❖ Documents (Web pages) represented by words
- ❖ Inverted index links keywords to the documents that contain them
- ❖ Keyword query retrieves documents containing that word

Inverted Index Example



Image Search - Tags

❖ Search over tags associated with images

- Users manually add Tags to images
- Find images with tags that match the query key

❖ Limitations

- Tags require human effort to create
- Tags may be wrong



Image Search - Text

❖ Use text associated with images for search

- Search web for images
- Use surrounding text
- Text in URL for image filename
- Text in HTML on page
- Same as text search



Sunset at Rocky Point

❖ Example: Google Image Search for "Sunset" gives

- Sunset at Rocky Point in Australia
- Sunset Beach, Oahu
- Frank Smiles at Sunset



Frank Smiles at Sunset



Sunset Beach

❖ Because the keyword "Sunset" was in the title of all these images

Image Similarity Search

❖ Query is an image

❖ Search finds similar images

❖ Similarity is defined by features of the image

- Color Content
 - Color Histogram
 - Color Correlation
- Image descriptors
 - Gradients at image keypoints
 - Quantize for "Visual words"
- Faces
 - Detection
 - Recognition



Query Image



Search Results

Image Search – Similar Color Content

❖ Color Histogram

- Distribution of pixel colors in image
- No spatial information
- Similarity based on histogram distance



❖ Color Correlogram

- Color histogram as a function of distance between pixels
- Distribution of pixel color plus spatial information
- Similarity based on correlogram difference

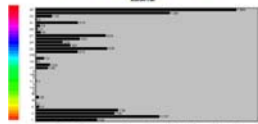


Image Search Results – Color Content

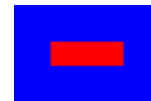
❖ Comparison of Image Retrieval

- Color Histogram
- Color Correlogram

❖ Correlograms are better for image retrieval



Images with different color histograms but similar correlograms



Images with identical color histograms but different correlograms

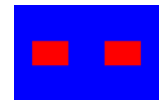


Image Search –Faces

❖ Face Detection

- Find faces in images
- Search for all images with faces
- Ex: Google advance search for images with faces
- Good results!

❖ Example:

- FXPAL Photo Application (2004: Girgensohn et al.)



Faces in Photo Collection



Photo Collection



Face Detection

Image Search –Faces

❖ Face Recognition

- Search for all images of a particular person
- Bad results!

❖ Face Similarity

- Similarity search based on face features
- Use face similarity to help manually label faces
- Good results!

❖ User Interface for Labeling Faces

- Drag face to label



Audio (Music) Search – Text

❖ Search text fields

- Title
- Artist
- Album
- Genre

❖ Example

- iTunes



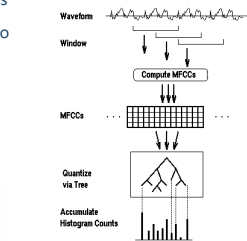
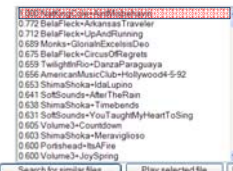
Audio (Music) Search - Sound

❖ Find similar sounding music

- Compute spectral feature vectors (MFCC)
- Quantize features to create audio histogram
- Audio histogram describes sounds
- Order of sounds is lost

❖ Example

- 1997: Jon Foote, FXPAL
- Similarity of Nat King Cole and Gregorian Chant
- Music Retrieval Demo



<http://www.rotorbrain.com/foote/music/>

Video Search – Whole Video

❖ Search for an entire video

- Search using surrounding text

❖ Example: Google/YouTube

- Sunset - Nitin Sawhney**
4 min · Nov 17, 2000 · 4.5 (28 ratings)
Nitin Sawhney... Sunset Very best video of Nitin Sawhney... Sunset
<http://www.youtube.com/watch?v=5jgZr7f>
[Watch video here](#) · [Related videos](#)
- Sunset at Cafe Del Mar 2004**
4 min · Oct 17, 2006 · 4.5 (22 ratings)
Mar 2004. A really Cool Sunset at Cafe Del Mar Ibiza I overlaid two songs on this video, the first is Water in Motion by sonic
<http://www.youtube.com/watch?v=1y2MhKqG>
[Watch video here](#) · [Related videos](#)
- The Kinks - Waterloo Sunset**
1 min · Jul 21, 2006 · 4.5 (429 ratings)
Waterloo Sunset... video of the kinks performing live version of waterloo sunset... The Kinks
<http://www.youtube.com/watch?v=MdDzYvY>
[Watch video here](#) · [Related videos](#)

Video Search: Segments of Video

❖ Video

- Sequence of frames (images), typically with audio
- 30 frames/second

❖ Text Transcript of Audio

- Time-correlated with video

❖ Segments of Video

- Shot: Unbroken segment of video from a single camera
- Story (news): Sequence of shots from the same news story

❖ Keyframe

- Representative image from a video segment



Slide 40

Video Search – News Programs

❖ Find segments of news on a topic of interest

- Find news story
- Find shots within story

❖ TRECVID

- Sponsored by NIST (National Institute of Standards)
- Data base of 60 hours of news video (ABC, NBC) in 2004 – similar content other years
- Task – user has 15 minutes to find shots relevant to a topic

❖ Example Topics

- "Find shots of a hockey rink with at least one of the nets fully visible from some point of view"
- "Find shots zooming in on the US Capitol dome"
- "Find shots of Saddam Hussein"



Slide 41

Video Search – News Retrieval

❖ TRECVID task is to find shots relevant to the query

- Use keyword search and image search

❖ Keyword Search

- Retrieve stories relevant to keyword

❖ Image Search

- Retrieve stories with shots relevant to keyword

❖ Merge results of image and keyword search

- Examine shots within the retrieved stories

❖ TRECVID Search

- User enters keywords and/or images for query
- System returns relevant stories
- User explores stories for relevant shots

Slide 42

Story Summary Quads

- ❖ Query-dependent story summary
 - Use 4 highest scoring shots
 - Allocate space proportional to score



Slide 43



Outline

<http://media.hust.edu.cn>

- 搜索引擎产品分析
- 商用搜索引擎的架构与原理
- 多媒体信息搜索
- 未来视频搜索引擎的架构**
- 问题与挑战

华中科技大学数字媒体实验室

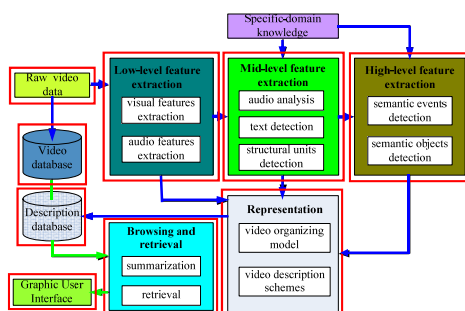
视频数据的特点

<http://media.hust.edu.cn>

- ❖ Data-rich & computing-intensive
 - $25 \times 352 \times 240 \times 3 = 6.34 \text{ MB/s}$ (uncompressed)
 - Terabyte video data processing (about 46hs)
- ❖ Complex algorithms
 - Non-structured data
 - inter-discipline knowledge
- ❖ Challenges for programming
 - Real-time processing
 - Parallelization

华中科技大学数字媒体实验室

Architecture

<http://media.hust.edu.cn>


华中科技大学数字媒体实验室

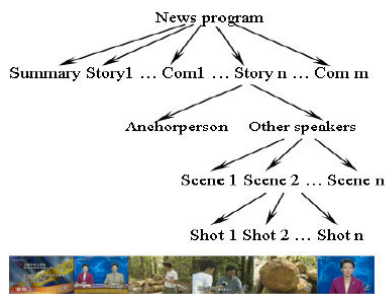
Video Organization

<http://media.hust.edu.cn>


华中科技大学数字媒体实验室

Video Organization-for News

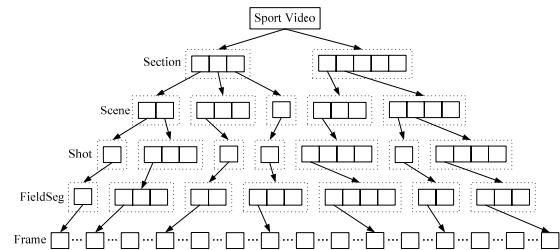
<http://media.hust.edu.cn>



华中科技大学数字媒体实验室

Video Organization for Sports

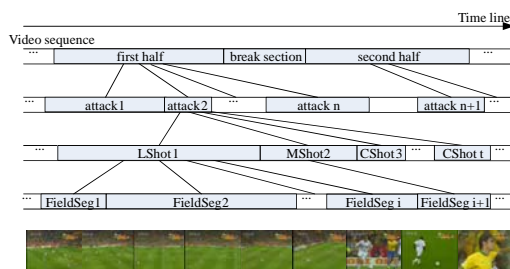
<http://media.hust.edu.cn>



华中科技大学数字媒体实验室

Video Organization for Soccer

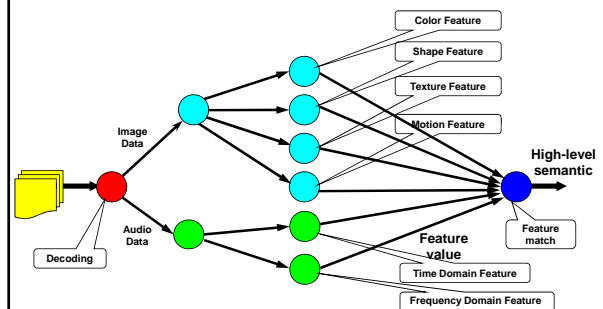
<http://media.hust.edu.cn>



华中科技大学数字媒体实验室

如何提取特征？

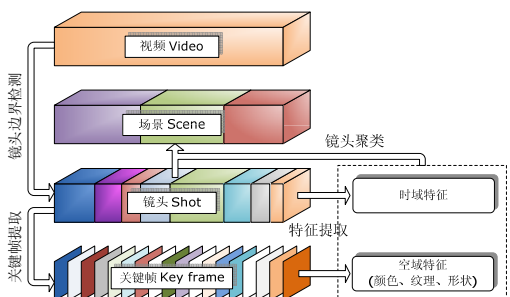
<http://media.hust.edu.cn>



华中科技大学数字媒体实验室

数字视频的结构分析

<http://media.hust.edu.cn>



华中科技大学数字媒体实验室

MPEG-7与视频数据检索系统

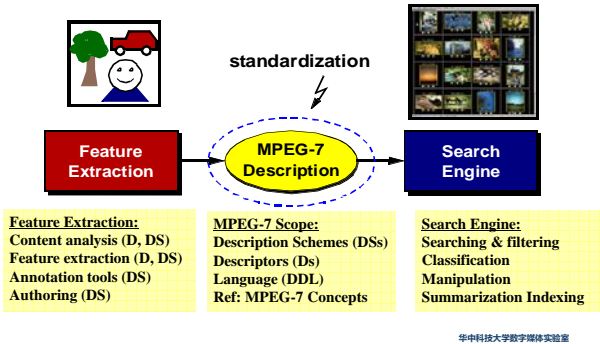
<http://media.hust.edu.cn>

- ❖ MPEG - 7标准是关于视频描述数据的标准，它为视频描述提供了标准的模式
- ❖ MPEG - 7标准使得所有的描述数据可以与原始视频数据分离，使得视频描述数据库的建立成为可能
- ❖ 符合MPEG - 7标准的视频数据描述文档主要以XML文档的形式存在

华中科技大学数字媒体实验室

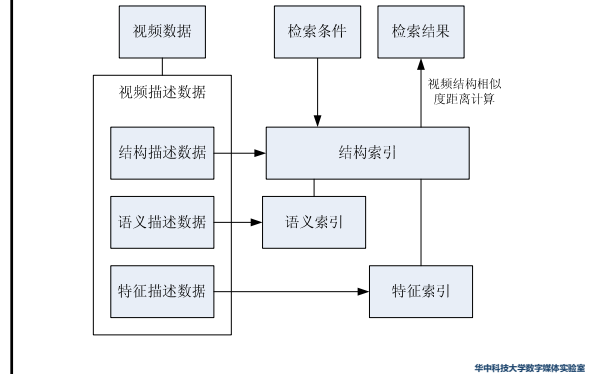
数字视频的描述-MPEG-7

http://media.hust.edu.cn



视频数据索引

http://media.hust.edu.cn



视频索引的分类

http://media.hust.edu.cn

- ❖ 从选取的索引内容出发，还可以把视频索引分成三类：
 - 基于注释的索引(Annotation-based Indexing)
 - 基于特征的索引(Feature-based Indexing)
 - 基于特定领域的索引(Domain-specific Indexing)

华中科技大学数字媒体实验室

Outline

http://media.hust.edu.cn



问题与挑战

http://media.hust.edu.cn

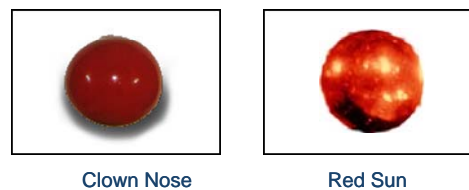
- ❖ 语义内容的提取
- ❖ “语义鸿沟” - Semantic Gap
- ❖ 基于多模特征的检索

华中科技大学数字媒体实验室

Perceptual/Conceptual Issue

http://media.hust.edu.cn

Similar Percepts / Dissimilar Concepts



华中科技大学数字媒体实验室

Perceptual/Conceptual Issue

<http://media.hust.edu.cn>

Dissimilar Percepts / Similar Concepts



John's Car



Mike's Car

华中科技大学数字媒体实验室

结论

<http://media.hust.edu.cn>

如何有效地填补“语义鸿沟”，
从而提高语义级视频检索的效率和服务
质量已经成为基于内容视频检索领域的
研究难点和最为关键的问题！

华中科技大学数字媒体实验室

