

# Introduction to Media Computing

## 媒体计算基础

### Lecture 1: Introduction

Lecturer: Zechao Li (李泽超)

# 相关学习资料

- ◆ Google、Bing等
- ◆ Wikipedia、百度百科等
- ◆ 国际顶级期刊
  - IEEE T-PAMI、T-NNLS、T-IP、T-MM等
  - ACM T-IST、ACM TOMCCAP、ACM Computer Survey等
- ◆ 国际顶级会议
  - CVPR、ICCV、ACM MM、SIGIR、KDD、AAAI等等



# 课程特点

---

- 名词、概念比较时髦
  - 技术、应用紧跟时代
  - 理论、原理比较乏味
  - 涉及较多基础课程
  - 要紧跟国际技术前沿
- 
- 互动，师生相互交流学习
  - 讲课、自学、讨论等方式结合

# 媒体概念

- 媒体Media：媒介或传播媒体
  - 传播信息的载体
  - 信息传播过程中从传播者到接受者之间携带和传递信息的一切形式的物质工具
  - 1943年美国图书馆协会的《战后公共图书馆的准则》一书中首次使用作为术语
  - 现在已成为各种传播工具的总称，如电影、电视、广播、印刷品(书刊、杂志、报纸)，而自从计算机的普及化，网上媒体在计算机网络中亦成为一种新形式的传播媒体



# 媒体分类

## ■ 载体

- 报纸杂志、广播、电视、互联网、移动网络

## ■ 内容

- 文本、音频、图像、视频等

## ■ 形式

- 流媒体、新媒体、移动媒体、自媒体

文字：书籍、报纸、广告、短信

声音：电话、磁带、语音短信

静止图像：照片、彩信、传真

连续图像：电影、录像、电视、动画



**载体多样性**

# 载体变化

报纸、书籍

电台广播

电视广播

信件

电话传真

网站

音乐服务

视频服务

电子邮件、网上聊天

IP电话、视频电话





# 《理解媒介：论人的延伸》

---

“媒介是人的延伸”

“媒介即是讯息”

“电子媒介是中枢神经系统的延伸”

“媒介使人自恋和麻木”

《Understanding Media: The Extensions of Man by Marshall McLuhan》

# 人类感知感官

---

- 视觉：文字、图像、视频
- 听觉：声音
- 触觉：键盘、鼠标、触摸屏等
- 味觉：？？？
- 嗅觉：？？？



# 计算机处理

---

计算机只能使用二进制“0010101101”表达、存储

文字：每个英文字母需要7比特，汉字16比特

声音：一个比特串

图像：一个比特矩阵

数字化：所有的信息都以二进制形式存储

# 时间依赖性

---

## ■ 时间独立媒体

- 文本
- 图形
- 离散媒体

## ■ 时间依赖媒体

- 音频
- 视频
- 连续媒体

注：“连续”指用户对数据的感觉，并不是数据内部的表示



# 多媒体

- 本课程主要关注多媒体（图像和视频）
- 多媒体从字面上理解就是“多种媒体的综合”，如数字、文字、声音、图像、视频、图形等。
- 多媒体与许多学科密切相关
  - 数字图像处理、信息论、数字信号处理、自然语言处理、数据库、人工智能、认知神经科学、机器学习等等
- 多媒体具有若干种不同的定义
  - “多媒体”（Multimedia）通常是指文字、数据、图像、视像及声音等多种信息形式的综合，它们都是基于人的认识而产生的不同的信息类型。

# 多学科之间区别

- 人工智能、机器学习、统计学、数据挖掘
  - 机器学习：涉及自学习算法，本质上通用，可以应用到众多相关问题
  - 数据挖掘：实用的应用算法（大多是机器学习），利用各个领域产出的数据来解决各个领域相关的问题
  - 统计学：研究怎么收集、组织、分析和解释数据中的数字化信息，包含描述统计学和推断统计学。描述统计学涉及组织、累加和描绘数据中的信息；推断统计学涉及使用抽样数据推断总体。
- 机器学习利用统计学（大多为推断统计学）来发自学习算法；数据挖掘是在从算法得到的结果上应用统计学（大多为描述统计学）解决问题，在各种各样的行业中求解问题，需要用到不同领域的不同技术和实践



# 多学科之间区别

---

## ■ 人工智能、机器学习、统计学、数据挖掘

- 人工智能：目的在于开发一个模拟人类能在某环境下做出反应和行为的系统或软件，领域广泛，定义多个子目标
  - 推理、知识表示、自动规划、机器学习、自然语言处理、计算机视觉、机器人学、通用智能
- 机器学习是从人工智能的一个子目标发展而来

## ■ 起源

- 统计学：1749
- 人工智能：1940
- 机器学习：1946
- 数据挖掘：1980

# 多媒体

---

## ■ 多媒体

➤ Multimedia=Multi+media

➤ Multi: Many

➤ Medium:

- Thing in the middle
- Means to distribute and present information

# Hypermedia and Multimedia

- A **hypertext** system: meant to be read nonlinearly, by following links that point to other parts of the document, or to other documents (Fig. 1.1)
- **HyperMedia**: not constrained to be text-based, can include other media, e.g., graphics, images, and especially the continuous media — sound and video.
  - The World Wide Web (WWW) — the best example of a hypermedia application.
- **Multimedia** means that computer information can be represented through audio, graphics, images, video, and animation in addition to traditional media.

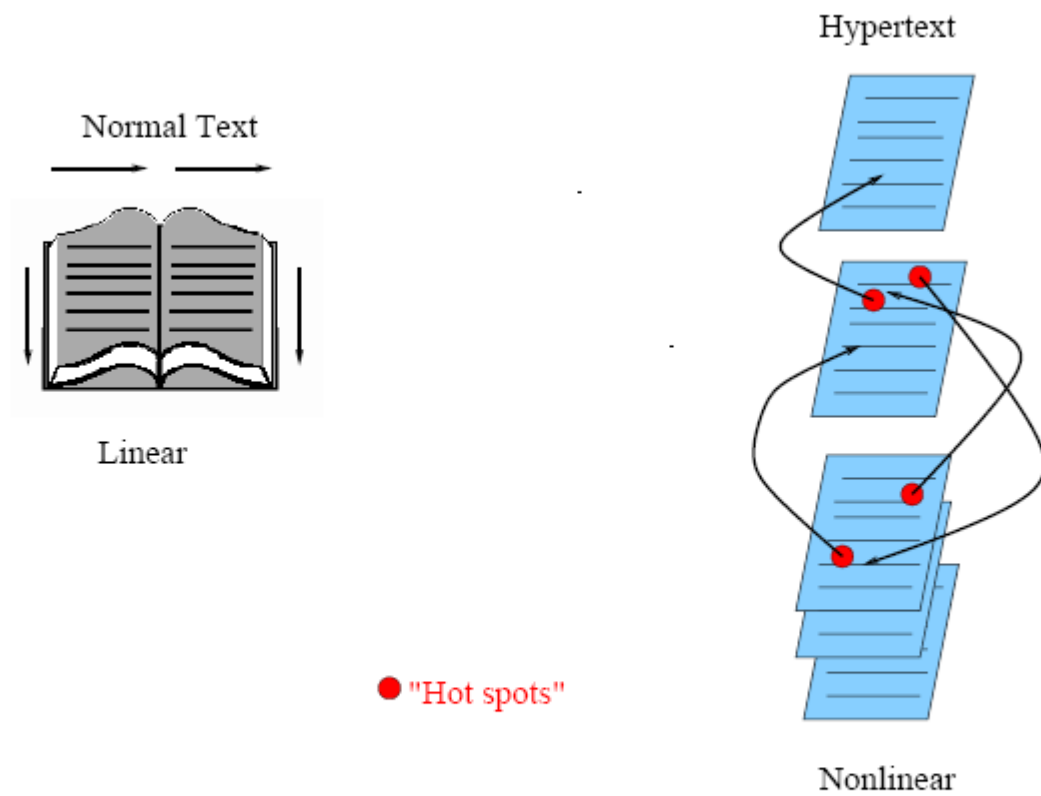


Fig 1.1: Hypertext is nonlinear



# 图像格式

---

## 例子

- GIF (GIF87. GIF89)
- JPEG
- TIFF
- PNG
- PS and PDF

# 多媒体特性

---

## ■ 媒体多样性

- 信息载体的多样化、是人们对计算机更人性化的要求。
- 媒体多样性有特定内涵：包含时基类媒体。

时基媒体(Continuous):如电话、视频、声音。

离散媒体(Discrete):非时基媒体，如文字、图像、数据库数据等。

## ■ 交互性

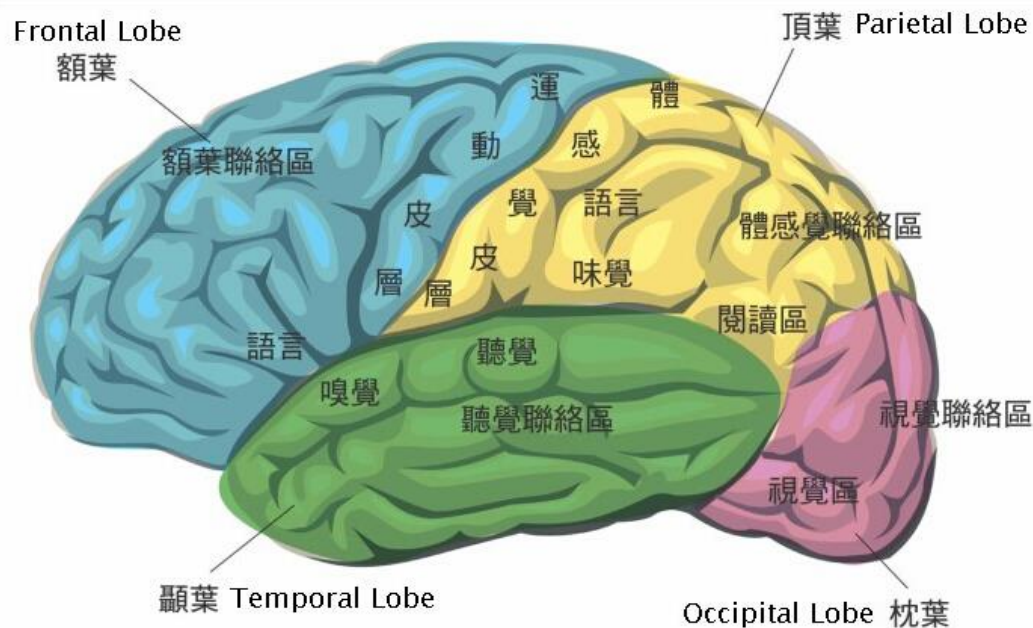
- 交互性是多媒体技术与传统视听技术的根本区别。
- 交互性是计算机人性化的必然要求，向用户提供了有效控制和使用信息的手段。

## ■ 集成性

- 集成性要求体现了人类对信息的完整性要求。媒体之间的关系蕴含着大量的信息。
- 媒体表示、处理的集成化。不是现有技术的综合，需要集成化多媒体信息模型。
- 技术发展的必然。

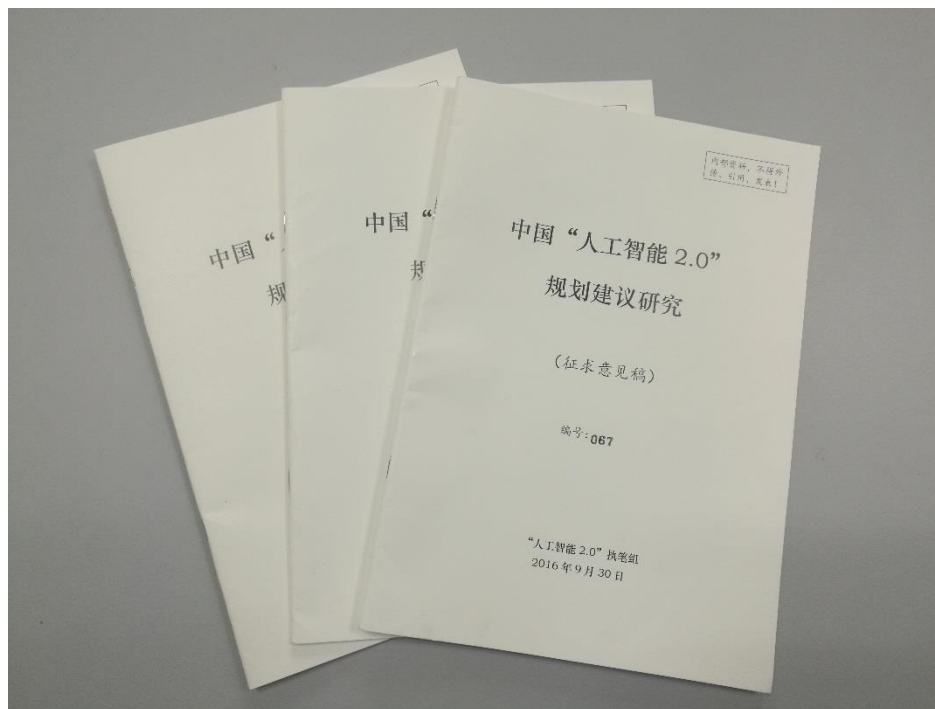
# 为什么研究多媒体

人对客观事物的认识就是通过耳听声音信息、眼观形象信息及手触质地信息等**综合方式**形成的，就是一种多媒体的信息采集、传输和处理过程。因此，在**人脑中基于认识产生的知识信息也是一种多媒体信息**。人脑中的知识信息通常也是以文字、语音、图像等多媒体信息的形式输出到客观世界的。



# 智能媒体计算

## 新一代人工智能发展规划





# 多媒体技术

---

多媒体技术是人们通过计算机或以微处理器为核心的终端设备，交互处理多种媒体（文本、图形、图象、声音、动画和视频等），使多种信息建立逻辑连接，集成为一个系统并具有交互性。

多媒体技术就是对多媒体信息进行综合采集、处理、传输、存储和控制利用的技术。

# 多媒体技术发展

- 能看到的历史138年
  - 照相机发明于1877年
- 能听到的历史176年
  - 留声机发明于1839年
- 能品尝到的历史442年
  - 泸州老窖始建于1573年





# 多媒体技术历史回顾

---

- 1839年，法国达盖尔发明**照相术**，这是图像处理技术的开端。
- 1843年，英国白英发明传真原理。
- 1848年，英国贝开尔发明圆柱扫描原理。
- 1873年，英国梅依发现硒光电效应。
- 1876年，美国古勒哈姆·贝尔发明**电话**。
- 1884年，法国P.G.Nipkov取得了机械扫描盘发明专利。
- 1887，法国赫兹成功产生火花电波。
- 1889，爱尔斯塔发明光电管。



# 多媒体技术历史回顾

---

- 1893，美国爱迪生发明电影。
- 1895，马可尼进行无线电电报实验。
- 1897，德国K.F.Braun发明了阴极射线管
- 1904，英国弗莱明发明二极管
- 1907年，美国福莱斯托发明三极管
- 1925，英国J.L.Baird表演了机械扫描电视
- 1927，美P.Farnsworth取得电子电视的专利



# 多媒体技术历史回顾

---

- 1930，英、苏等国进行机械电视广播。
- 1933，美V.K.Zworykin取得光电摄像管专利。
- 1936，英国BBC电视广播（黑白）
- 1941，NBC、CBC广播开始。
- 1946，美国发明视像管；
- 1946，世界上第一台电子管计算机埃尼阿克（ENIAC）在美国宾西法尼亚大学研制成功。（18000多个电子管，占地170平方米，重达30万千克）；
- 1948，贝尔研究所发明晶体管

# 多媒体技术历史回顾

- 1951，美国试播场顺序制彩色电视，它与黑白电视不兼容，因而未得到推广。
- 1953，美国FCC批准彩色电视NTSC制式国家标准，后日本、加拿大等国相继采用NTSC
- 1956，美国Ampex公司研制成5.08厘米、四磁头横向扫描磁带录象机（VTR）。
- 1958，晶体管计算机诞生。
- 1965，以中小规模IC为主体的计算机问世。
- 1963，荷兰发明光电摄像管。
- 1964，贝尔实验室实现0.5MHZ的可视电话。
- 1967，德国正式开播PAL制式彩色电视，同年，法国和前苏联开播SECAM制彩色电视。

# 多媒体技术历史回顾

- 20世纪70年代，CCD摄像器件、电子新闻采访和电子现场节目制作、二磁头螺旋扫描录像机等发明，以及有线电视（Cable Television）、共用天线电视（Community Antenna Television）、光纤通信、可视数据通信（Videotex）等都取得可喜的发展。
- 1970年，大规模集成电路计算机研制成功，计算机发展到第四代微型计算机。
- 1971年，在美国硅谷第一台微型计算机诞生了，从而开创了微型计算机的新时代，自此以后，几乎每隔二、三年就推出一代新的微处理器，相继的新的更高性能的计算机诞生。直至现在的多媒体计算机。
- 1972年英国独立广播公司（IBA）研制成功第一台数字电视制式转换器，宣告新一代数字电视的诞生。
- 1972年9月，Philips展示播放电视节目的激光视盘。



# 多媒体技术历史回顾

---

- 1978年模拟电视光盘和播放机投入市场。随后，IVD(交互式视频光盘)、CD-DA（数字音频光盘）、WORM（一次写多次读光盘）等陆续推出商品化产品，使大容量数字存储技术获得重要的新进展，这为多媒体时代的到来创造了重要条件。
- 1984年，首展激光CD-ROM(Compact Disk-Read Only Memory)驱动器，每片650MB。
- 为使计算机之间便于信息交换，制定了CD-ROM文件标准，1988年正式通过该项国际标准ISO9660。
- 进入 20 世纪 90 年代，计算机的普及，标志互联网时代的到来，人们运用计算机技术及网络通信技术创造着这个时代 的辉煌。

# 多媒体技术历史回顾

- 20世纪90年代，世界向着信息化社会发展的速度明显加快，而多媒体技术的应用在这一发展过程中发挥了极其重要的作用。多媒体改善了人类信息的交流，缩短了人类传递信息的路径。应用是20世纪90年代计算机应用的时代特征，也是计算机的又一次革命。
- 国际标准化协会、国际电子学委员会、国际电信协会等国际组织，于20世纪90年代领导制定了三个重要的有关视频图像压缩编码的国际标准，JPEG标准；H.261标准；MPEG标准。  
2、音频信息处理的应用在多媒体技术中，存储声音信息的文件格式主要有：WAV文件、VOC文件、MIDI文件、AIF文件、SON文件及RMI文件等。
- 20世纪90年代末，互联网检索系统开始起步，google，百度



# 多媒体技术历史回顾

---

- 21世纪以来，多媒体技术得到了快速发展
- 3G、4G无线网络
- 智能移动终端
- 社交平台
- 物联网
- 大数据
- 。 。 。



# 技术发展回顾结论

---

- 视听技术、电子技术、通信技术和计算机技术是在相互交替中得到发展的，其中一个技术的进步往往建立在其它技术的进步上。
- 而由这些技术的综合集成，产生的多媒体系统，其功能和应用前景都是其它任何一种技术无法比拟的。



# 多媒体技术不断扩张

---

- 计算机处理多种媒体
- 各种微处理器对多种媒体的处理
- 因特网上多媒体技术的浪潮
- 无线通信世界多媒体的涌现
- 多媒体技术走入家庭



# 多媒体信息系统

---

- 信息系统的发展目标：产生出新一代高度集成、功能强大、智能化的多媒体信息系统。
- 多媒体信息系统是计算机信息系统在本质上的一次飞跃。
  - 多媒体信息系统集成的范围更加广泛。
  - 使用的信息真正实现多媒体化。
  - 多媒体信息系统的应用范围更加广泛。
  - 系统的使用更加符合人们地习惯、更加人性化。



# 多媒体技术与多媒体系统的关系

---

- 多媒体技术是建立多媒体系统的基础；
- 多媒体系统的建立与应用又反过来促进多媒体技术的不断完善、发展；
- 多媒体技术与多媒体系统都会随着与它们相关的其它技术的进步而不断向前发展。

# 做什么？

141.8 亿



10.6 亿

facebook

观看量1000万小时/天



1 亿

开心网

1.78 亿

renren

北京百万摄像头 TB级/天



智能



5亿 用户

10亿条/天



GPS

1亿用户



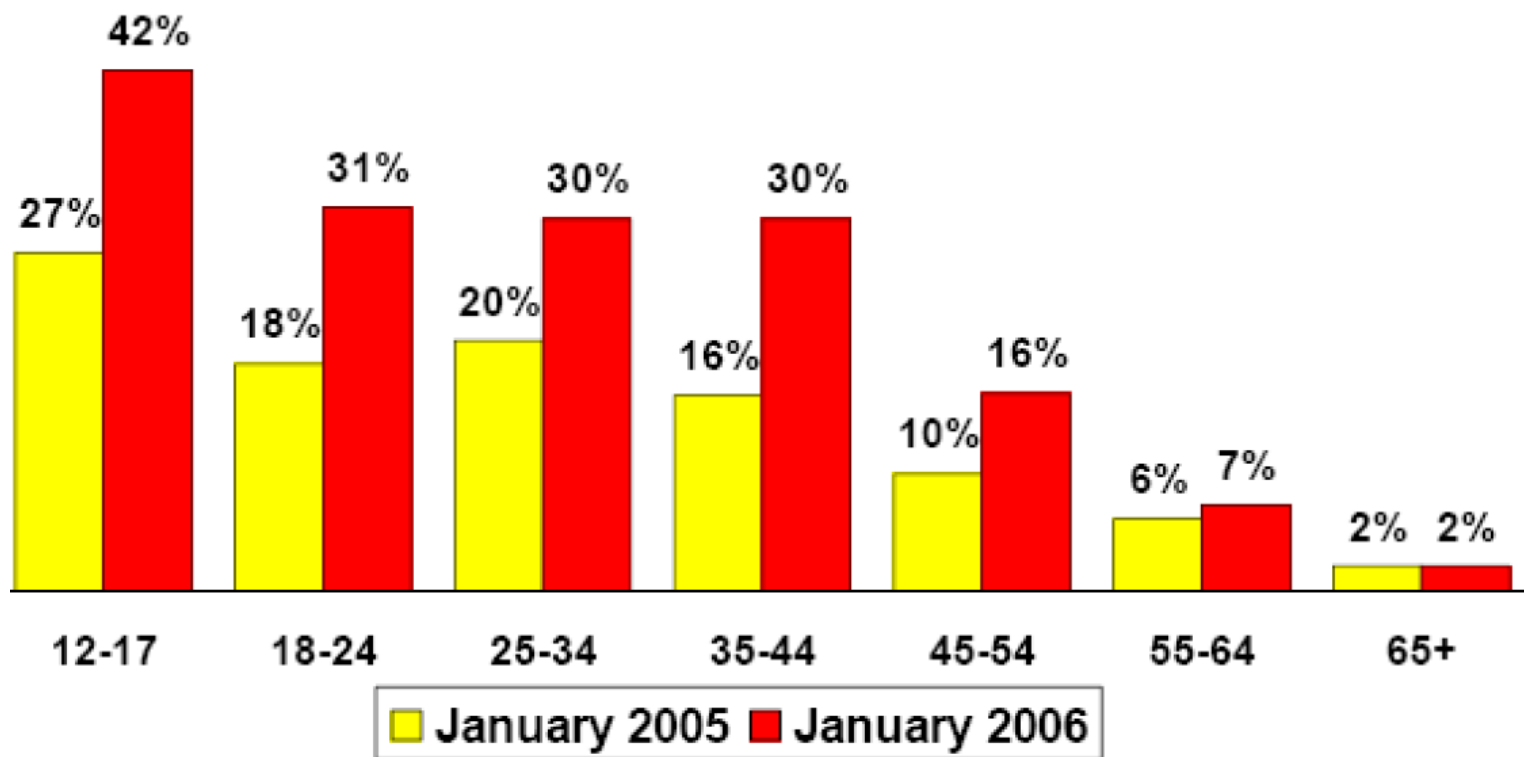
3亿 用户

# 美国消费电子类产品比例

<b>Percent Of U.S. Population Who Own/Use On-Demand Devices/Media</b> <small>Base: Total Population 12+</small>		
	<b>2005</b>	<b>2006</b>
Satellite TV	<b>25%</b>	<b>25%</b>
Portable DVD Player	<b>19%</b>	<b>24%</b>
High-definition TV (HDTV)	<b>10%</b>	<b>19%</b>
TiVo/DVR (2005 = ownership of TiVo/DVR; 2006 = any usage of TiVo/DVR, including cable/satellite)	<b>6%</b>	<b>18%</b>
Video On Demand (watched last month)	<b>10%</b>	<b>14%</b>
Portable MP3 player other than iPod	<b>10%</b>	<b>14%</b>
iPod	<b>6%</b>	<b>11%</b>
Handheld wireless e-mail device such as a BlackBerry	<b>3%</b>	<b>4%</b>

Source <http://www.edisonresearch.com/>

# 美国青少年拥有iPod





# 典型应用

---

- 各种电脑游戏（在某种程度上，游戏的发展推动多媒体技术的迅速发展）
- 多媒体教室（目前我校的多媒体教室应该说还处于初级阶段）
- Mp3播放机
- 数码相机和数码摄像机
- 多功能的手机



# 与人方便

---

- 可以不用粉笔写字的多媒体教室
- 远程教育可以在家里答疑
- 可以躺在被窝里用手机聊天
- 可以坐在火车上用笔记本上网
- 可以说声“关机”就让计算机自动睡觉
- 一个回车键就可以把照片发给所有同学
- 一个网页就能传播总书记的讲话



# 省钱的技术

---

- 电话-电话的IP电话
- 用msn、qq、Skype打可视电话
- PPLive、PPStream看免费电影、电视
- 不需要耗材的数码相机
- 可以让学校不需要多盖教室而多收学费的网络教学



# 双刃剑

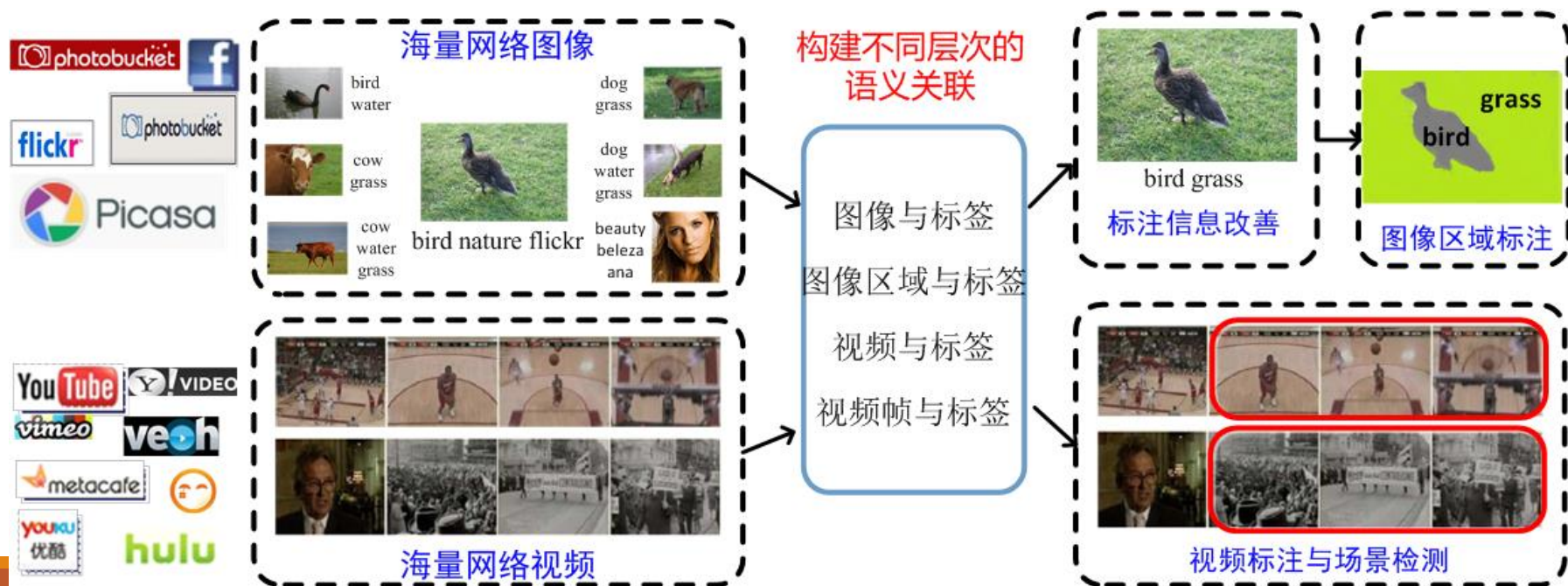
---

- 很多老教授需要熬夜准备多媒体讲稿
- 洗衣机总是发出各种声音以显示自己的工作状态
- 如果不升级计算机就不能玩最新的游戏
- 如果不知道多媒体就被视为菜鸟
- 隐私

# 媒体对象的多粒度语义分析

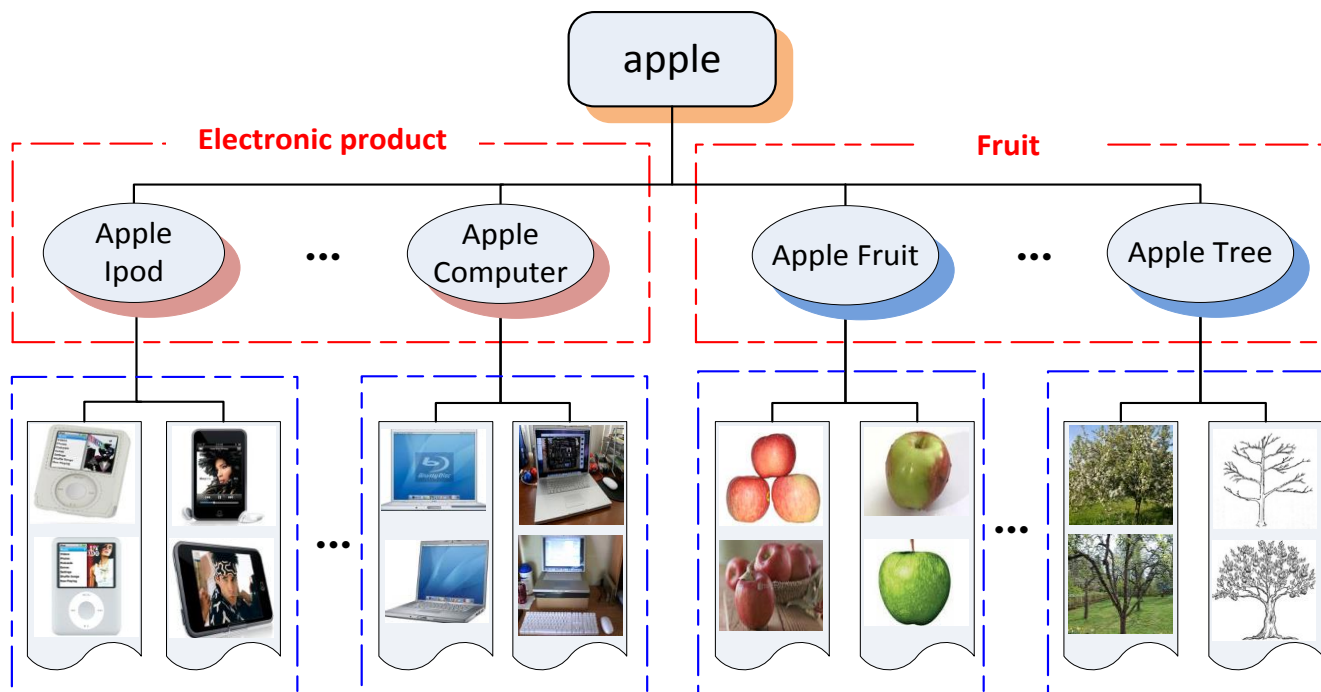
## 基本任务

- 面对海量的网络图像与视频数据，结合社会标签数据，建立不同层次的媒体对象与标签之间的语义关联
- 面临困难与挑战：大规模数据，社会标签的主观性、有噪音与不完整



# 图像检索

基于关键词的检索：考虑关键词的歧义性



# 图像检索



Semantic Clusters

Item-based search result

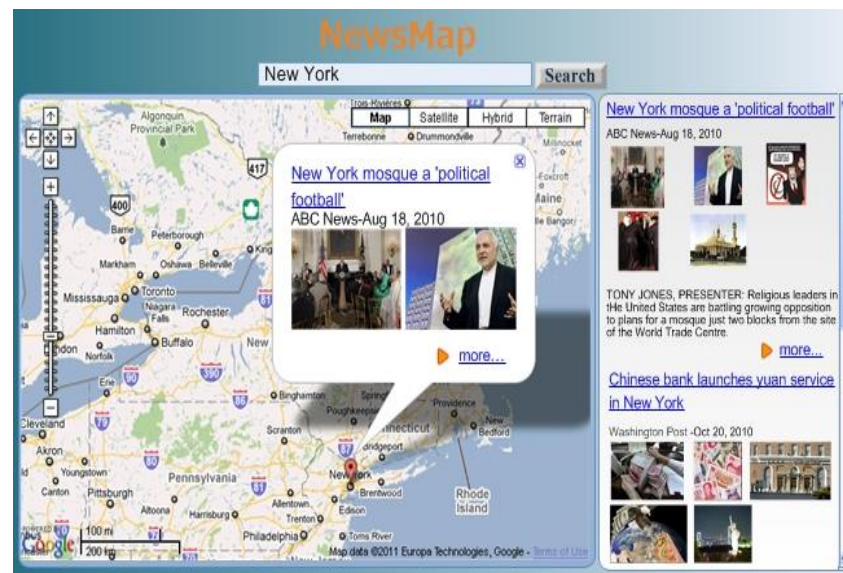
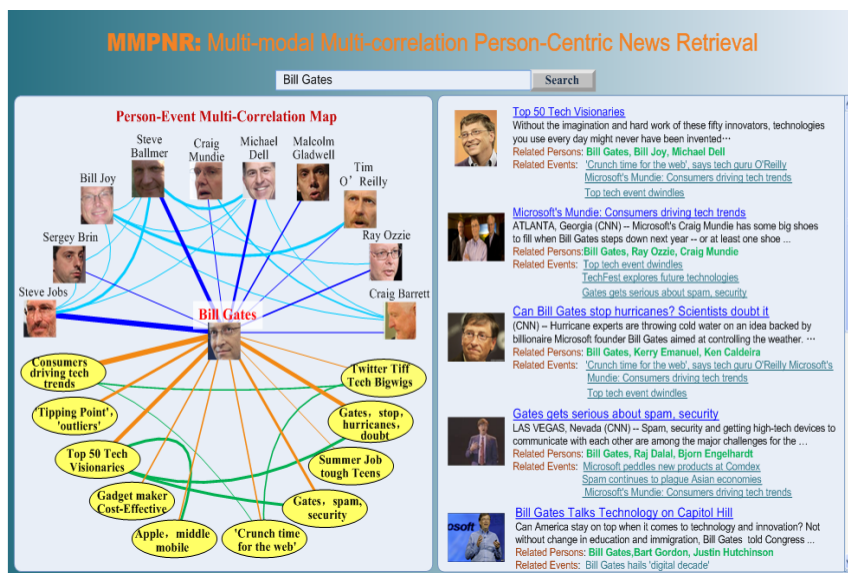


Visual clusters under the same semantic topic

- ❑ 共索引近千万网络图像
- ❑ 可方便用户更加快速的找到所需要的图像

# 多媒体新闻检索

关注新闻四要素：4W – who, where, what, and when



- 可提供基于新闻人物、新闻事件和地理位置的新闻检索服务
- 可自动生成针对某新闻事件的多媒体新闻摘要

# 跨场景的服装检索

**应用场景：**用户在街道上用手机拍摄或在网络相册中发现一件喜欢的衣服，系统可以从网上商店寻找类似的服装



日常生活照




























购物网站图

**跨场景检索：**是指以日常生活的自然图片为查询，从在线购物网站的标准图片中找到相似的集合返回给用户。二者在人的拍照姿势和背景复杂度上有较大差别。

# 跨场景的服装检索

## 检索结果实例：可单独提供上衣与下衣的检索结果

	Sleeve					
	Collar	★★★★	★★✖★	★★✖★	★★★★	★★★★
	Color	★★★★	★★✖★	★★✖★	★★★★	★★★★
	Length	★★★★	★★★★	★★★★	★★✖★	★★✖★
	Sleeve					
	Collar	★★★★	★★★★	★★✖★	★★✖★	★★✖★
	Pattern	★★★★	★★★★	★★✖★	★★✖★	★★✖★
	Length	★★✖★	★★✖★	★★✖★	★★★★	★★★★
	Plain					
	Front	★★✖★	★★★★	★★★★	★★★★	★★★★
	Color	★★✖★	★★★★	★★★★	★★★★	★★★★
	Length	★★★★	★★✖★	★★★★	★★✖★	★★✖★
	Sleeve					
	Button	★★★★	★★✖★	★★✖★	★★✖★	★★✖★
	Material	★★★★	★★✖★	★★✖★	★★✖★	★★✖★
	Length	★★✖★	★★✖★	★★★★	★★★★	★★★★
	Shape	★★✖★	★★✖★	★★★★	★★★★	★★★★
	Curling	★★✖★	★★✖★	★★★★	★★★★	★★★★
	Color	★★✖★	★★✖★	★★★★	★★★★	★★★★
	Length	★★✖★	★★✖★	★★★★	★★★★	★★★★

[Demo](#)

# 视觉辅助的实时聊天系统

应用场景：在用户进行文本聊天的同时，自动为其配以图示；同时还添加了中英对照翻译，方便不同母语用户之间的交流。



# 自动问答系统

MMAnswering

What happened on September 11, 2001 ?



## Original QA

### Questions:

What happened on September 11, 2001 ?

### Description:

I Need it for a school project. I need to know what happened.

### Best Answer(From Y!A):

Three buildings in the World Trade Centre were destroyed by controlled demolition, and a hole was blown by a missile, and the whole thing was blamed by the government on 19 Arabs whom were said to have hijacked four passenger jets.

## Images:



## Videos:





# 多媒体研究中面临的问题

---

- 数字音频技术;
- 数字视频技术;
- 多媒体系统平台技术;
- 多媒体数据转换和同步技术;
- 多媒体数据的压缩与解压缩技术;
- 高性能存储系统设计与制造技术;
- 高分辨率图象显示技术;
- 多媒体网络技术;
- 音频/视频信号识别、合成和转换的智能处理技术;
- 多媒体技术标准的研究。

# 怎么做



用户希望输入文本搜索图像和  
视频本身包含的语义内容



多媒体分析  
与理解

指通过机器学习或其它  
方法自动预测图像或视  
频中包含的语义内容

物尽其用，为我所用

Use everything you can utilize



# 工具

---

## ■ 数学

- 最优化
- 矩阵分解
- 稀疏编码
- ...

## ■ 工具包

- **CVX: Matlab Software for Disciplined Convex Programming:**  
<http://cvxr.com/cvx/>
- **SPArse Modeling :** <http://spams-devel.gforge.inria.fr/>

# 相关研究人员

---

- Thomas S. Huang
- Ramesh Jain
- Shih-Fu Chang
- Alex Hauptmann
- Nicu Sebe
- Chong-Wah Ngo
- Tat-seng Chua
- Shuicheng Yan
- Jiebo Luo
- Anil K. Jain
- Wen Gao
- Songde Ma
- Yong Rui
- Shiqiang Yang
- Hanqing Lu
- Qingming Huang
- Yueting Zhuang
- Xiangyang Xue
- Changsheng Xu
- Tao Mei



# 多媒体技术

---

- 了解多媒体的设备，防止日后遇到
- 了解多媒体关键技术，免为广告词困惑
- 了解多媒体发展方向，跟上时代步伐
- 了解一些常见的多媒体软件，便于日后在工作中使用
- 通过课程、拿到学分
- 掌握原理性的东西，考试
- 掌握技术前沿，发表论文，毕业

---

# Attitude changes everything

# 态度决定一切

谢谢大家