

Multimedia Technology

Lecture 1: Overview and Arrangement

Lecturer: *Dr. Wan-Lei Zhao*

Autumn Semester 2022

Outline

- 1 About this Course
- 2 Syllabus
- 3 Course plan
- 4 Brief History about IR and Web
- 5 Brief History about WWW

- Major subjects
 - Deal with information such as text, image and video
 - Text retrieval, content-based image retrieval and video retrieval
 - Focus on how to retrieve above mentioned information
 - Popular machine learning approaches will be covered
 - k-means
 - Popular algorithms in computer vision will be covered
 - BoVW, Fisher Kernel and Minhash encoding
- Cover two central issues
 - **Search**
 - **Recommendation**
 - Direct information (text, image, video, and audio) to the one who wants

- Text Retrieval (4×2 hours)
 - Brief History about IR and Web
 - Pre-processing on Text Information
 - Three Retrieval Models
 - Boolean, vector models and word2vector model
 - Evaluation Measure
 - Web Search
- Machine Learning Approaches (1×2 hours)
 - k -means and its variants
- Image search (3×2 hours)
 - Challenges & Trends
 - Image Features
 - BoVW Framework
 - Fisher Kernel Framework

- Nearest Neighbour Search (3×2 hours)
 - KD-Tree
 - Locality Sensitive Hashing
 - Product Quantizer
 - NN-Descent
- Visual Instance Search (2×2 hours)
 - Challenges & Trends
 - Convolutional Neural Network
 - Instance Search Search
- Recommendation System (1×2 hours)
- Music Retrieval (1×2 hours)
 - Fundamentals about Sound signals
 - Sound Feature Extraction

- Two Course projects
 - One or two course projects
 - Implement after class
 - DIY is encouraged, **do it independently**
 - A hardcopy of the project report is also required

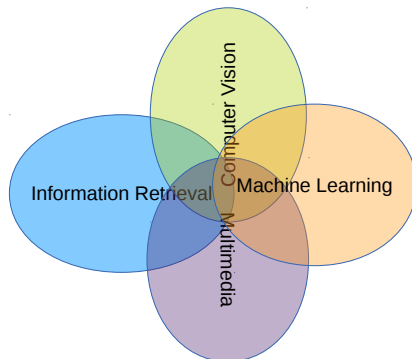
Prerequisites of this course

- Data Structure
 - You have to be familiar with it
 - Otherwise, you are not suggested to take this course
- Good at C/C++
 - It will be used in the lab
 - It is recommended for your course project
- Basic knowledge about Internet
 - Internet protocols
 - Mechanism of WWW
 - HTML and Javascript
- Linear Algebra
 - It will be used in the lab
 - Even you do not know, it does not matter
 - You will learn its basics during this course

Language in the Class

- English or Chinese?
- You might be uncomfortable at the beginning
- Me too:)
- Several advantages:
 - Computer science is defined in English
 - Get you guys used to English





- Intersection of four disciplines
- Related (top-ranked) Conferences:
 - ACM SIGIR, ACM SIGMOID, WWW
 - ACM MM, ACM ICMR & ACM ICME
 - IEEE CVPR & ECCV
 - IEEE ICCV, IEEE ACCV, IEEE ACCV & BMVC
 - ICML & AAAI

- Related (top-ranked) Journals:
 - IEEE Trans. on Knowledge and Data Engineering
 - IEEE Trans. on Pattern Analysis and Machine Intelligence
 - International Journal of Computer Vision
 - IEEE Trans. on Multimedia
 - IEEE Trans. on Image Processing
 - Computer Vision and Image Understanding
- Reference Books
 - R. Baeza-Yates and et al., Modern Information Retrieval: The Concepts and Technology behind Search (2nd edition)
 - Richard Szeliski, Computer Vision: Algorithms and Applications
 - Lecture notes of “Machine Learning” by Dr. Andrew Ng, from Stanford University
 - Related papers will be suggested to read as assignment
- Online Resources:
 - Youku
 - Wikipedia
 - Baidu Baike

Central Issues in Pattern Recognition

- ① Recognition (classification)
- ② Clustering
- ③ Mapping (embedding)
- ④ Regression
- ⑤ **Sort** ← mainly dealing with in this course

Central Issues that Are Covered

- ① Web Document Search
- ② Image & Instance Search
- ③ Content Recommendation
- ④ Music Retrieval
- ⑤ Indexing structure for NN search

Central Issues in Multimedia Information Processing

- How to organize the large-scale data
 - storage
 - retrieval
 - transfer
- How to understand the multimedia content
 - low level processing
 - semantic understanding
- Present to the one who wants it
 - recommendation
 - automatic hyper-linking

- Evaluation: two course projects (coding + report) + Oral Exam.
 $S = 35\% + 35\% + 30\%$
- About course projects
 - Implemented in C, C++/Python
 - If you do not know Python or Matlab, learn it!!
 - Sample codes will be given, you only need to fill “blanks”
 - Team work is NOT encouraged
- Failure is acceptable but no cheating or plagiarism
- If it happens, you are OUT!!
- Any questions?

Be an Active Learner

- Level 1
 - Catch the concept
- Level 2
 - Understand the idea
 - Know how to use it
- Level 3
 - Able to re-implement the algorithms
 - Knows where it works
 - Knows where it fails

Outline

- 1 About this Course
- 2 Syllabus
- 3 Course plan
- 4 Brief History about IR and Web**
- 5 Brief History about WWW

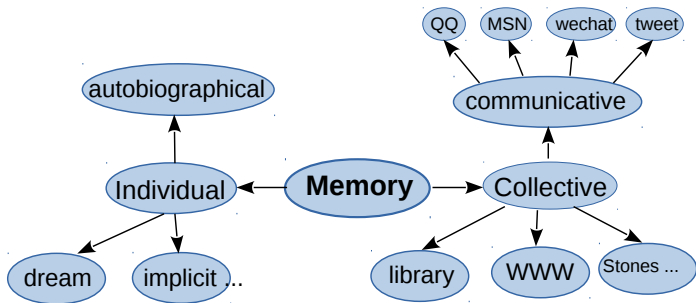
Why Multimedia Information Retrieval

- ① Multimedia Information are everywhere due to the **proliferation** of digital devices
 - Video, Audio and text
- ② Applications needs from the indurstry
 - kuaishou
 - a local company

Evolution of Storage Media in information age



About Memory



- No creatures in this world are created with knowledge
- Compared to other creatures, knowledge are kept apart from brain
- Knowledge are shared across different generations and geologies

IR in two different eras

	before WWW	WWW era
Media	text document, TV, film & CD	in electronic forms
Publishing	months or years	hours
Storage	books & papers	disc, DVD and etc & web
Indexing	title, author, keywords and date	and contents
Interface	library	browser

- According to IBM, 90% of the knowledge in the world are created in last two years
- Powerful IR system is required to coordinate the distribution of information/knowledge

The Birth of WWW

- 1981-1991: the invention of the Web
 - In 1980, Tim Berners-Lee worked in CERN (European Organization for Nuclear Research)
 - Manage information for physicists such that they can share
 - In 1984, he returned to CERN
 - In 1989, he wrote a proposal about “large hypertext database”
 - By Christmas 1990, he built all necessary elements for web
 - HTTP, HTML, web browser and httpd



The growth of World Wide Web

- Early times of growth (1991-1995)
 - Microsoft has its first browser: Cello
 - Mosaic (from UIUC) is the first successful browser
 - W3C was founded by Berners-Lee in 1994 at MIT
- Commercialize (1996-1998)
 - More and more dot-coms appeared
- Boom and Bust (1999-2001)
 - More and more dot-coms appeared
 - Internet becomes popular in China
 - Many currently well-known companies were established: Baidu, Alibaba
 - **Search Engines** were born

The growth of World Wide Web

- Early times of growth (1991-2001)
 - First version of Java was released in 1995
 - First version of PHP was released in 1995
 - JavaScript was invented by Netscape in 1995
 - Static web to dynamic web
 - Strong support for multimedia



Gosling



Lerdorf



Gutmans



Suraski

WWW is everywhere

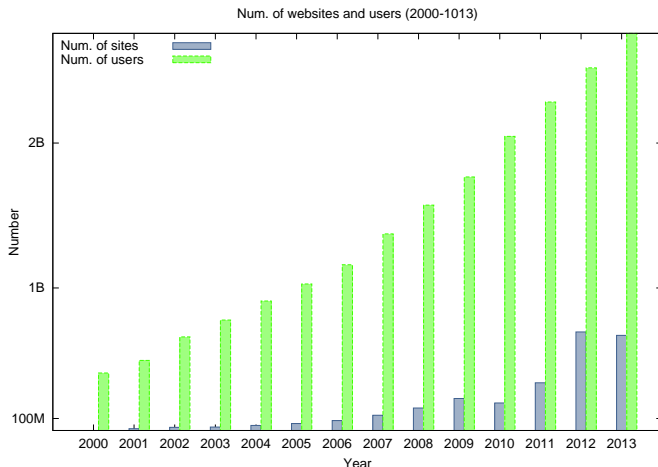
- Ubiquitous web (2002-present)
 - Introduction of Web 2.0 is the milestone
 - Wikipedia was born in 2001
 - Flickr was born in 2004
 - Facebook was born in 2004
 - Youtube was born in 2006
 - Twitter was born in 2006
 - Smartphone was released in 2007
- All technologies and media are intertwined to reshape the world
- Impact on our daily life of many aspects
- IR becomes the main interface to them all

Semantic Web

- Web 3.0 (20??)
 - Proposed by Berners-Lee¹
 - Websites are linked by semantic meta data
 - Machine builds the link automatically
 - Requires technology of natural language understanding
 - Still a vague concept
- Automatic documenting, e.g. books and recipes

¹Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web, in *American Scientific*, 2000

Statistics on WWW

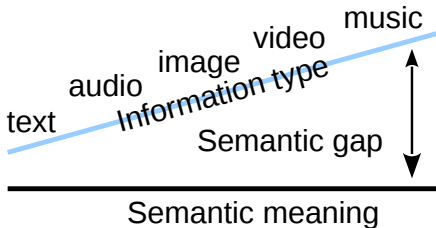


- The growth rate of user is much higher than that of websites
- The growth rate of clicks would be even much higher

Challenges in Modern Information Retrieval



fishes



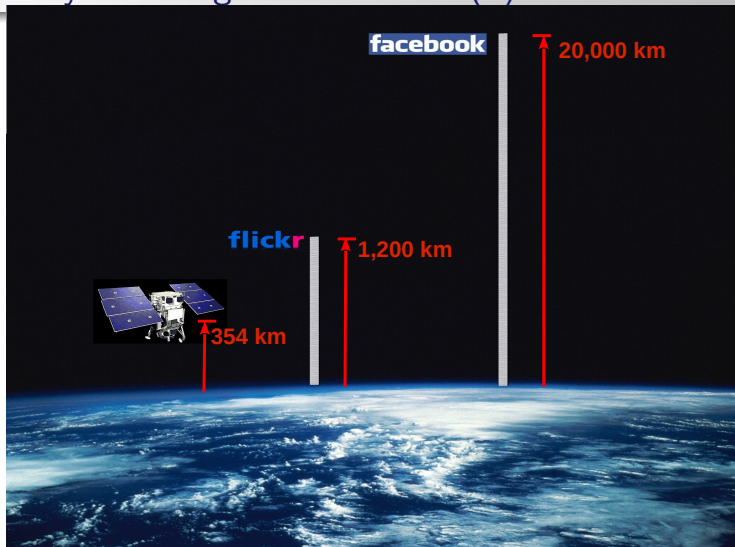
- How to bridge such a “semantic gap”
- A word is worth a thousand pictures
- A picture is worth a thousand of words

Scalability in the age of **BIG** data (1)

- A glance at big data today
 - **1.1** billion websites until Nov. 2014
 - **>3,000** images uploaded to Flickr in every minute²
 - **>200,000** videos uploaded per day to YouTube (**>1,000** years)
 - TV News: thousands hours of programs broadcasted each day
 - **>100** billion photos in Facebook till Jun. 2011
- Challenges: facilitate fast browsing and sharing
 - How to store?
 - How to organize?
 - How to retrieve?







²Statistics was collected on Apr. 28th 2010.

Scalability in the age of **BIG** data (2)



- Given the thickness of one photo: 0.2 mm

Top Rank Search Engines

	Num. of monthly unique user clicks
	1.1 billion
	350 million
	300 million
	245 million
	125 million
	100 million

- Google takes lion's share of the market
- Baidu is not in the rank (unfortunately)³

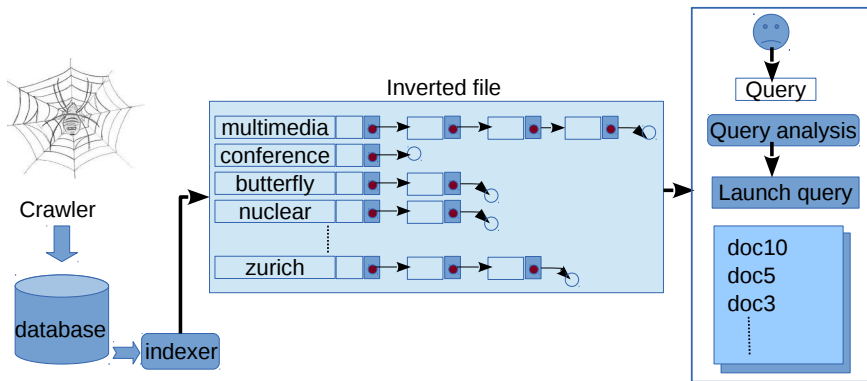
³Cited from: <http://www.ebizmba.com/articles/search-engines>

Sketch the framework of a search engine



- Draw a framework about a search engine in **5 minutes**
- Put all elements you could figure out, do it now ...

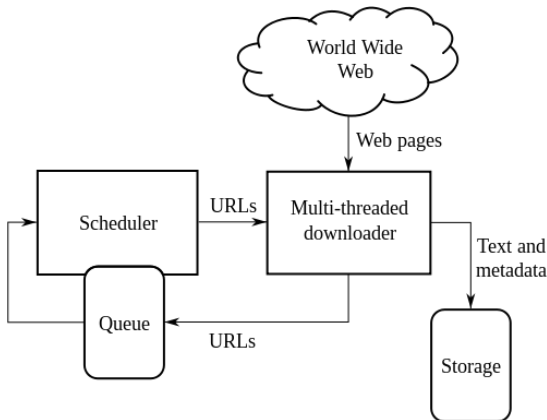
Framework of a search engine



- Observations

- Information are highly distributed in Internet
- The indexer (search engine) keeps information in a centralized manner

Structure of a crawler



- Observations

- Crawler plays very important role
- Experiences of using Baidu and Google

Q & A

Thanks for your attention!