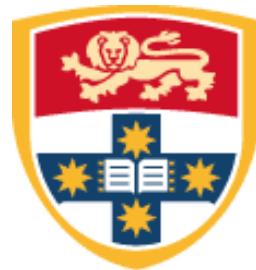


We acknowledge the tradition of
custodianship and law of the Country on which
the University of Sydney campuses stand.
We pay our respects to those who have cared
and continue to care for Country.





MULTIMEDIA RETRIEVAL



THE UNIVERSITY OF
SYDNEY

Week01

Semester 1, 2025

Prof. Zhiyong WANG

School of Computer Science
Rm 349, School of CS Building, J12
The University of Sydney NSW 2006 Australia
T +612 9351 3766 **F** + 612 9351 3838
E zhiyong.wang@sydney.edu.au

Jichao Kan

The University of Sydney NSW 2006 Australia
E jichao.kan@sydney.edu.au



ChatGPT



Explain quantum computing in simple terms



Quantum computers are a type of computer that uses quantum-mechanical phenomena, such as superposition and entanglement, to perform operations on data. In a classical computer, data is stored as bits, which can be either 0 or 1. Quantum computers use quantum bits, or qubits, which can represent a 0, a 1, or both values simultaneously. This property allows quantum computers to perform certain calculations much faster than classical computers.

One of the key differences between classical and quantum computers is that classical computers can only perform one task at a time, while quantum computers can perform multiple calculations at the same time. This makes quantum computers particularly well-suited for tasks that involve a lot of parallel processing.

Quantum computers are still in the early stages of development, and there are many challenges to overcome before they become widely available. However, their potential to revolutionize fields such as medicine, finance, and materials science could solve problems that are currently considered unsolvable.



create a simple website html code with red button and form



To create a simple website with a red button and form, you can use the following HTML code:

```
<html>
<head>
<style>
button {
    background-color: #ff0000;
    color: white;
    padding: 14px 20px;
    margin: 8px 0;
    border: none;
    cursor: pointer;
    width: 100%;
}

form {
    border: 3px solid #ff0000;
    padding: 14px 20px;
}
```



Now generate a test case using Cucumber for the google website



Here is a simple example of a test case using Cucumber for the Google website:

[Copy code](#)

Feature: Google Search

As a user
I want to search for a term on Google
So that I can find information on the topic

Scenario: Search for a term on Google
Given I am on the Google homepage
When I search for "selenium testing"
Then the results page should contain the search term in the title

```
# Implement the steps using Cucumber's step definitions
Given("I am on the Google homepage") do
  # Open the Google homepage
  visit "https://www.google.com"
end
```

```
When("I search for {string}") do |search_term|
  # Find the search input field and enter the search term
  fill_in "q", with: search_term
end
```

```
# Submit the search form
find("input[type=submit]").click()
```

```
do
  # Wait for the results page to load
  expect(page).to have_selector("#result-stats")

  # Verify that the search term appears in the title
  expect(page.title).to include("selenium testing")
end
```

This test case will use Cucumber to define the feature and scenario for searching on Google, and then implement the steps using Cucumber's step definitions. The steps will open the Google homepage, enter a search term, submit the search form, wait for the results page to load, and then verify that the search term appears in the title of the results page.

Note that this example uses Cucumber's built-in methods for interacting with the web page, such as `visit`, `fill_in`, and `find`, which are not available in plain Cucumber. You may need to use Selenium or another tool to perform these actions in your actual test case.

<https://www.cbsnews.com/video/microsoft-google-using-ai-to-revamp-search-engines/>



THE UNIVERSITY OF SYDNEY

School of Computer Science

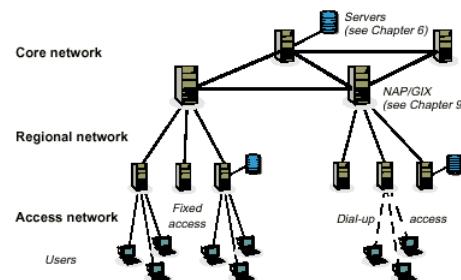
MMI
COMP5423

Week 01 – Introduction

- Overview of Multimedia Retrieval
- Overview of this Unit of Study

History of Computing

- 1940s – Hardware
- 1950s – Computing
- 1960s – Software
- 1970s – Databases
- 1980s – Desktop
- 1990s – Networking
- 2000+ – Multimedia



Media Forms

- Text: books, novels ...
- Audio, Sound,: music, radio ...
- Image, Graphics: photos, drawing, ...
- Video, Clips: TV, movie, ...

Multimedia: any combination of two or more media forms and well integrated to be presented via a single interface, or manipulated by computer programs.

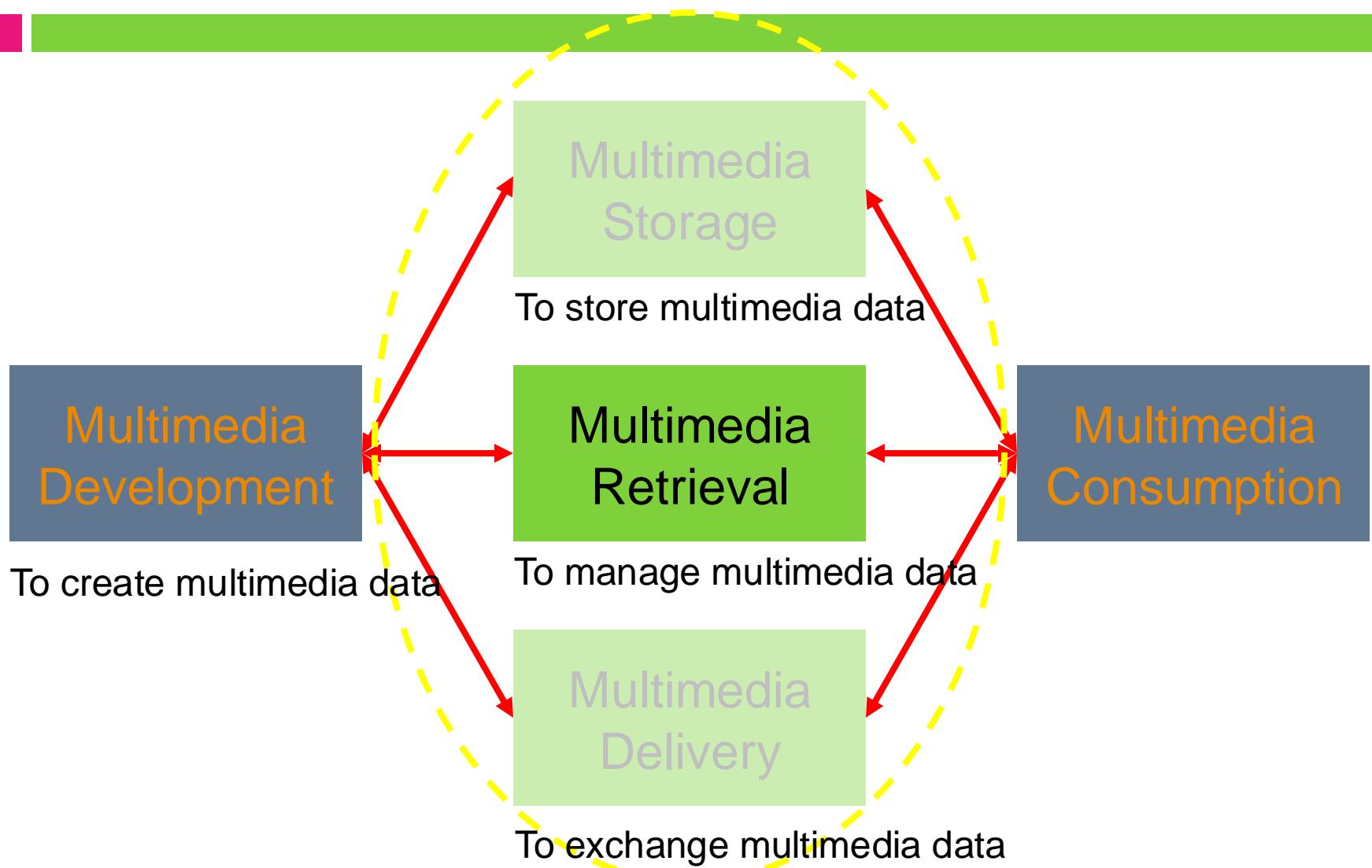
Multimedia

- Multimedia enriches the experiences in obtaining information
 - ▣ Letter → Telephone → Video conferencing

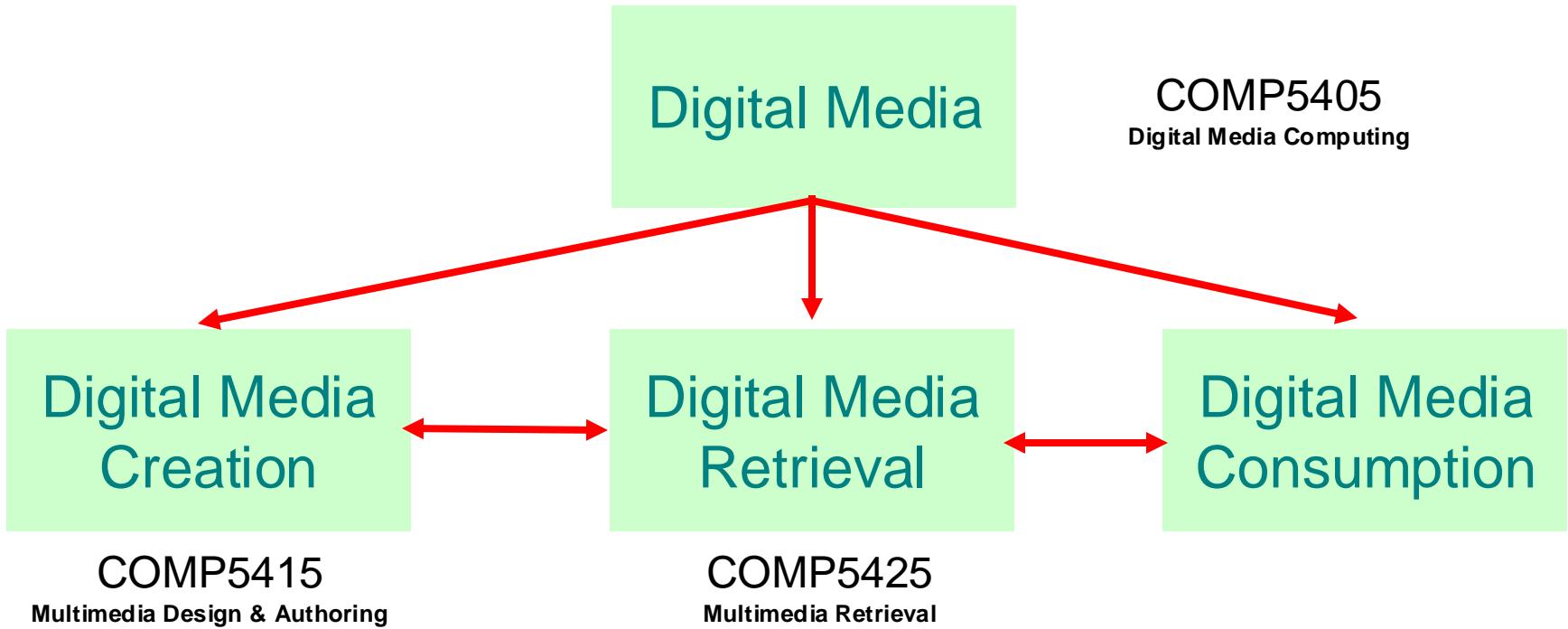


- Multimedia applications are changing our life
 - ▣ Education: Virtual University, eLearning ...
 - ▣ Research: Scientific visualization, Biomedical modeling
 - ▣ Entertainment: Games, ...
 - ▣ Healthcare: Tele-medicine, eHealth, ...
 - ▣ Geography: Remote sensing, ...

Multimedia Ecosystem



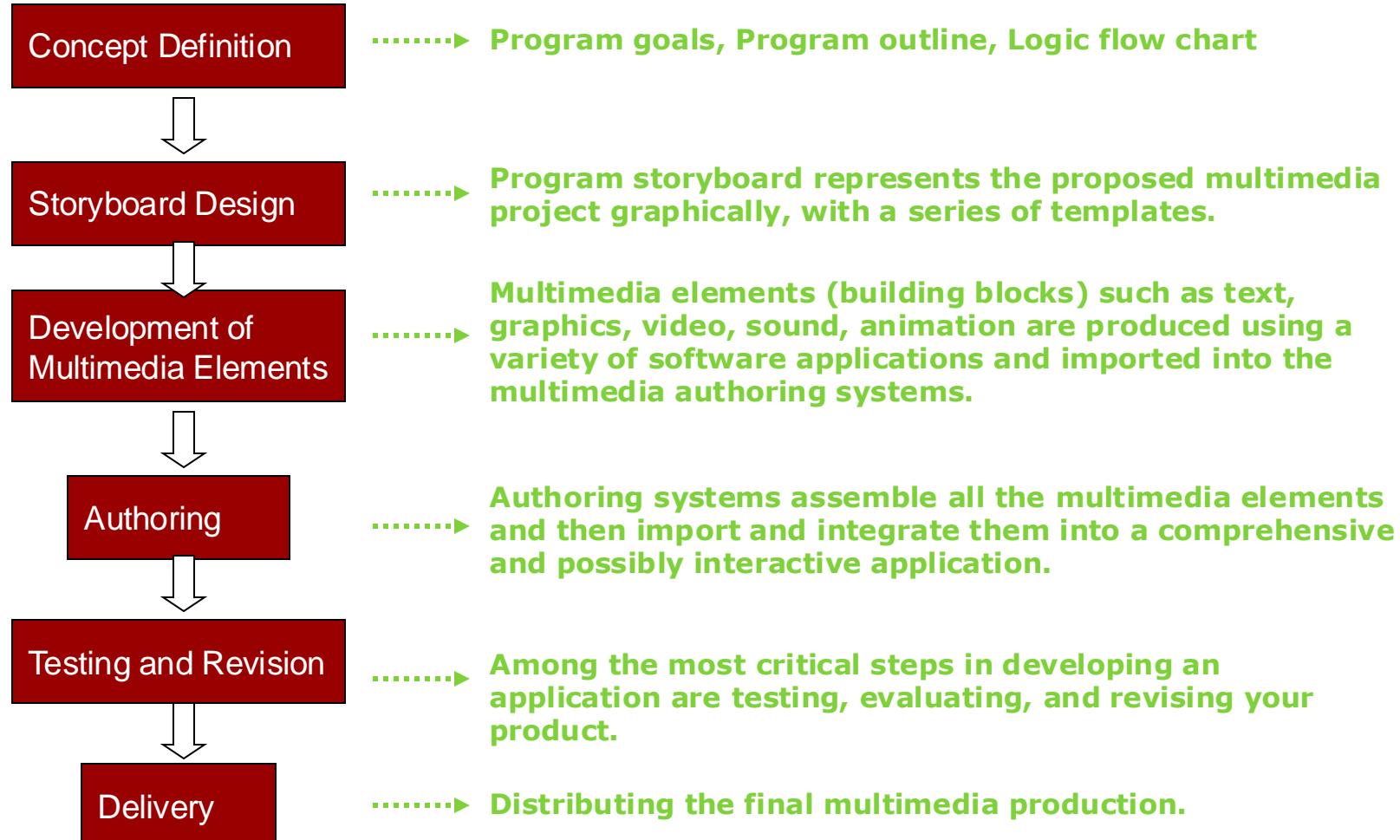
Digital Media Curriculum



COMP5045: Computational Geometry
COMP5048: Visual Analytics
COMP5427: Usability Engineering
IDEA9106: Design Thinking



Multimedia Development



See course COMP5415 Multimedia Design & Authoring

Data Creation --> Data Deluge



Vlog Device



Sports Cam



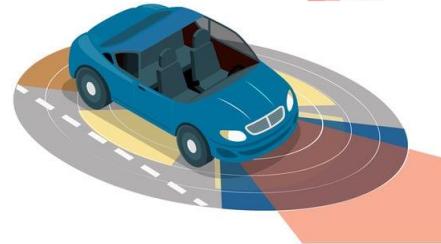
Smart wearables



Drones shooting



Live Stream

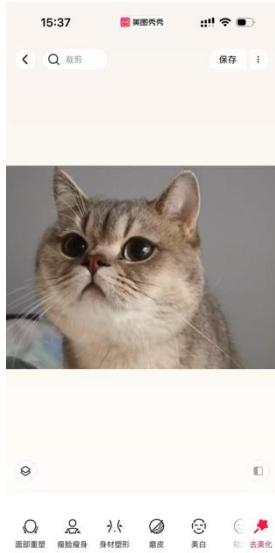


Cars

CAMERAS
RADAR
LIDAR

Data Enhancement

- Once we created the data, we can enhance them



element elimination

Photo Beautification

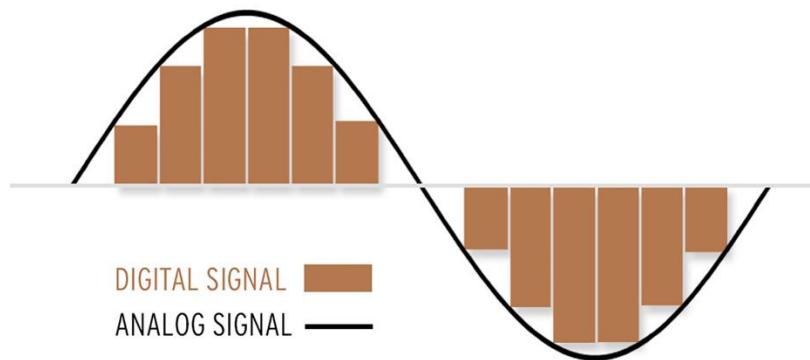


Color Grading

Multimedia Storage

- Digitization: to be manipulated by computer
- Compression: to save space
- Storage media: where to store and how to manage

Digitization

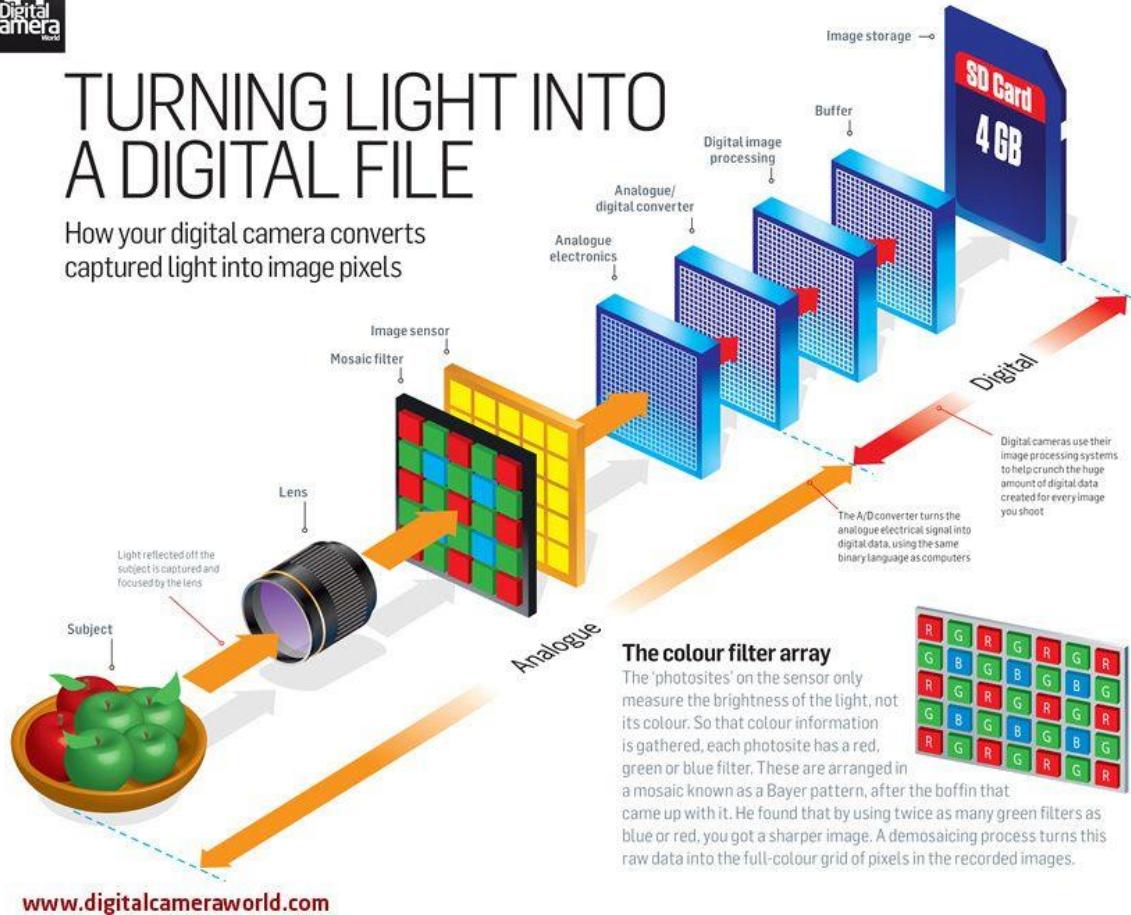


Digitization



TURNING LIGHT INTO A DIGITAL FILE

How your digital camera converts captured light into image pixels



Compression

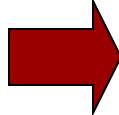
		Resolution/ Length	Color (bits)/ Sampling	Size	Compression
Audio		60s	16X44.1k/s (CD)	10.3MB	MP3
Image		1024 X 768	24	2.25MB	JPEG
VCD video	NTSC	352 X 240	24 / 29.97fps	7.2MBps	MPEG
	PAL	352 X 288	24 / 25fps	7.3MBps	
DVD video	NTSC	720 X 480	24 / 29.97fps	29.6MBps	MPEG
	PAL	720 X 576	24 / 25fps	29.7MBps	
MRI image		256 X 256	12	6-187MB (60-2000)	Lossless

Storage requirement

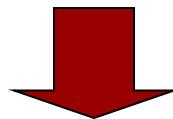


Compression

136	135	134	134	134
134	134	132	134	134
133	137	130	134	134
134	134	134	134	134
134	135	134	134	131



134	134	134	134	134
134	134	134	134	134
134	134	134	134	134
134	134	134	134	134
134	134	134	134	134



RED: (134,0,0)

25	134
----	-----

Compression

JPEG
PSNR =
32.8



JPEG 2000
PSNR =
34.7



Compression Ratio (click to select)

Aware, Inc.

Better visual quality can be expected in high compression ration with JPEG2000

Compression



original frame k



original frame k+1



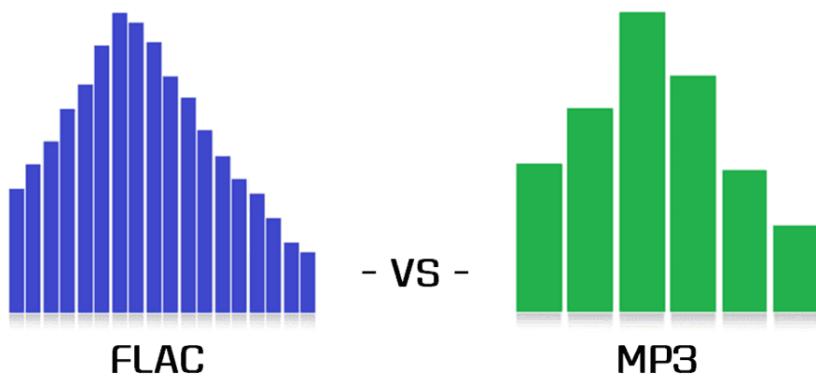
frame difference

- Drop frames
- Encoding frames

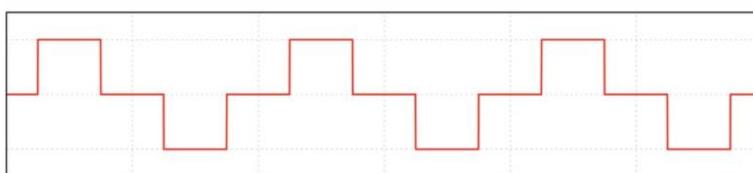
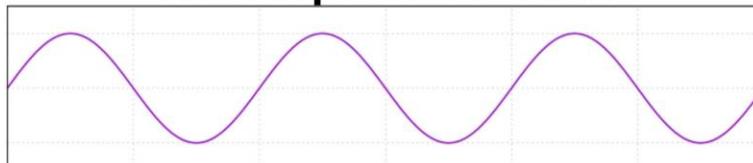
<http://www.it.et.tudelft.nl/~inald/vcdemo/>

Compression

□ WAV vs FLAC vs MP3



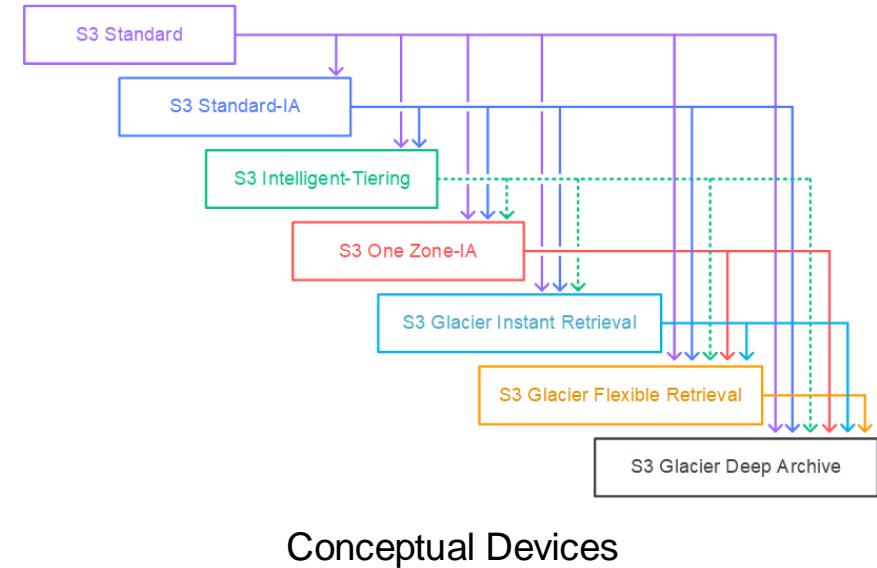
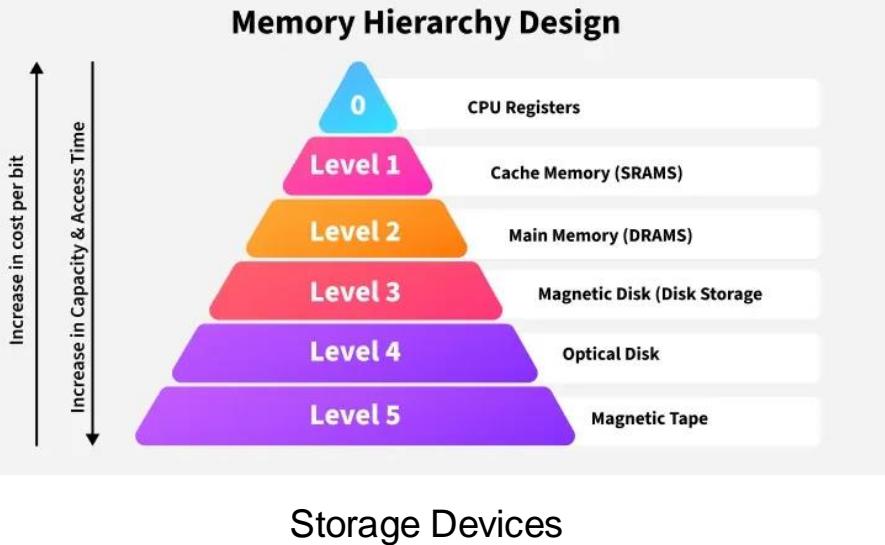
Uncompressed WAV



Lossy MP3



Storage Media

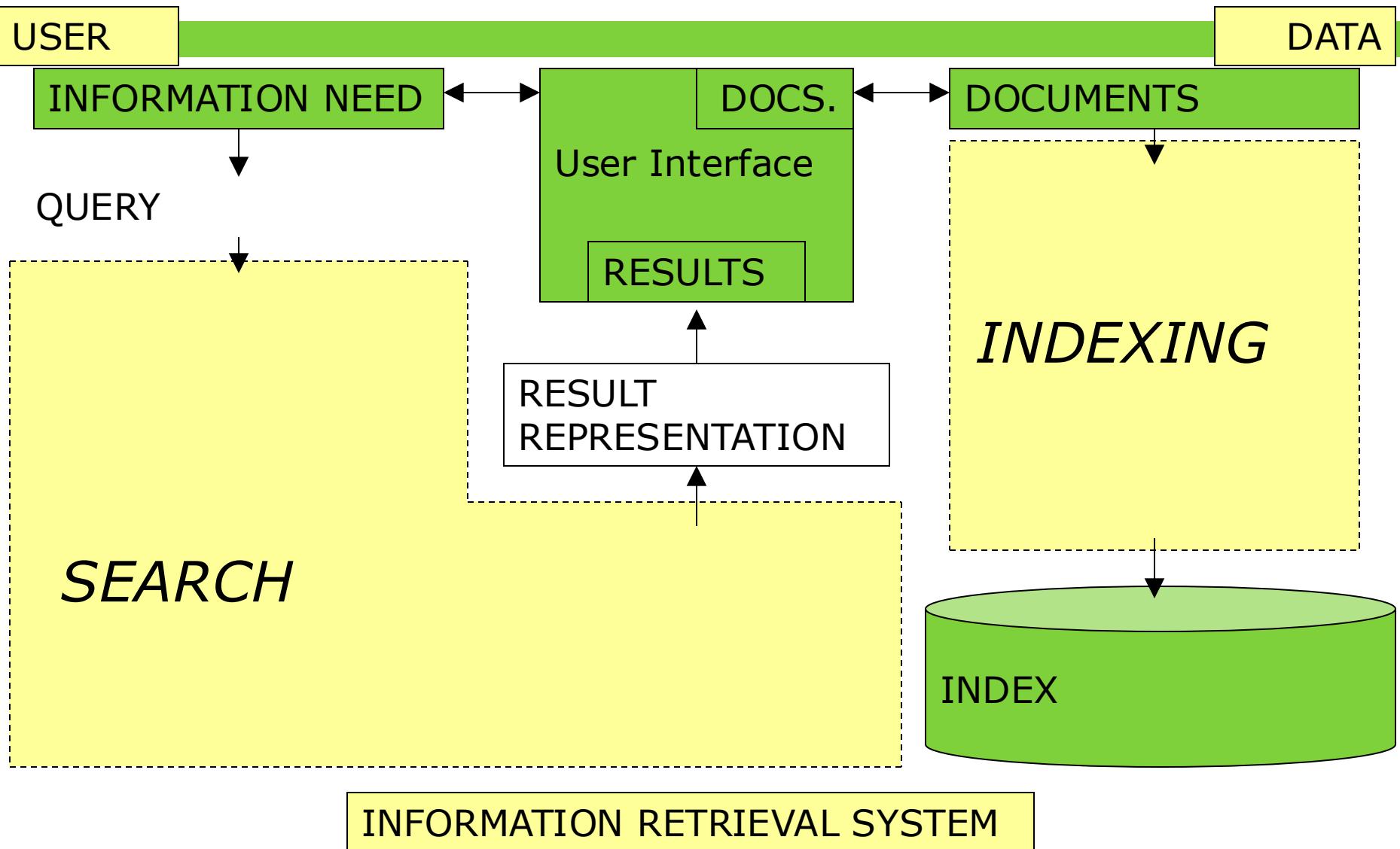


Multimedia Retrieval



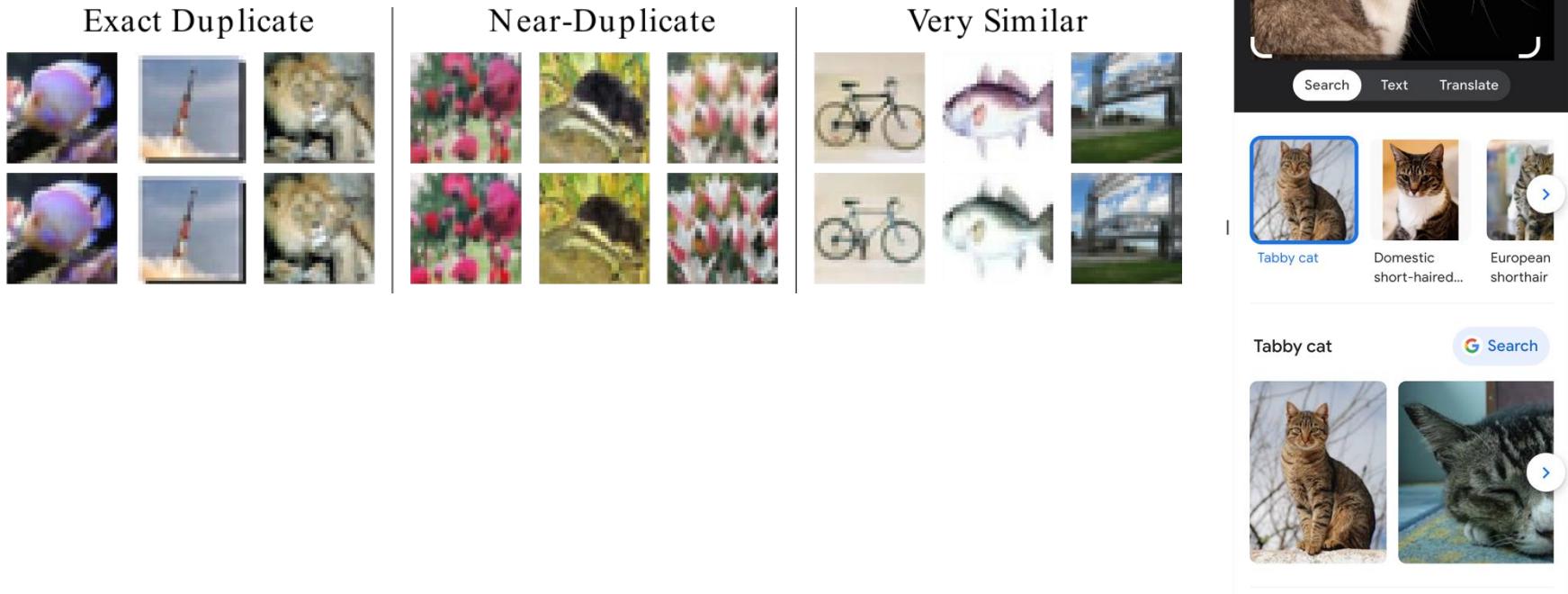
- Information is of no use unless you can actually access it.

Retrieval Process



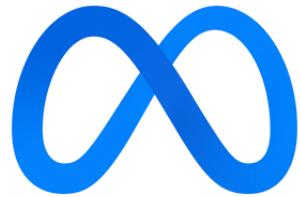
Retrieval

- What do we want?
 - Exact Duplicate? -> easy(X)
 - Near Duplicate



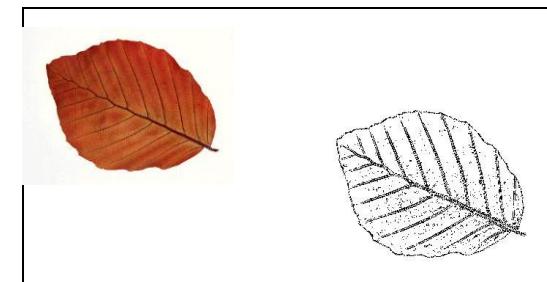
Web + Social

Google  Bing



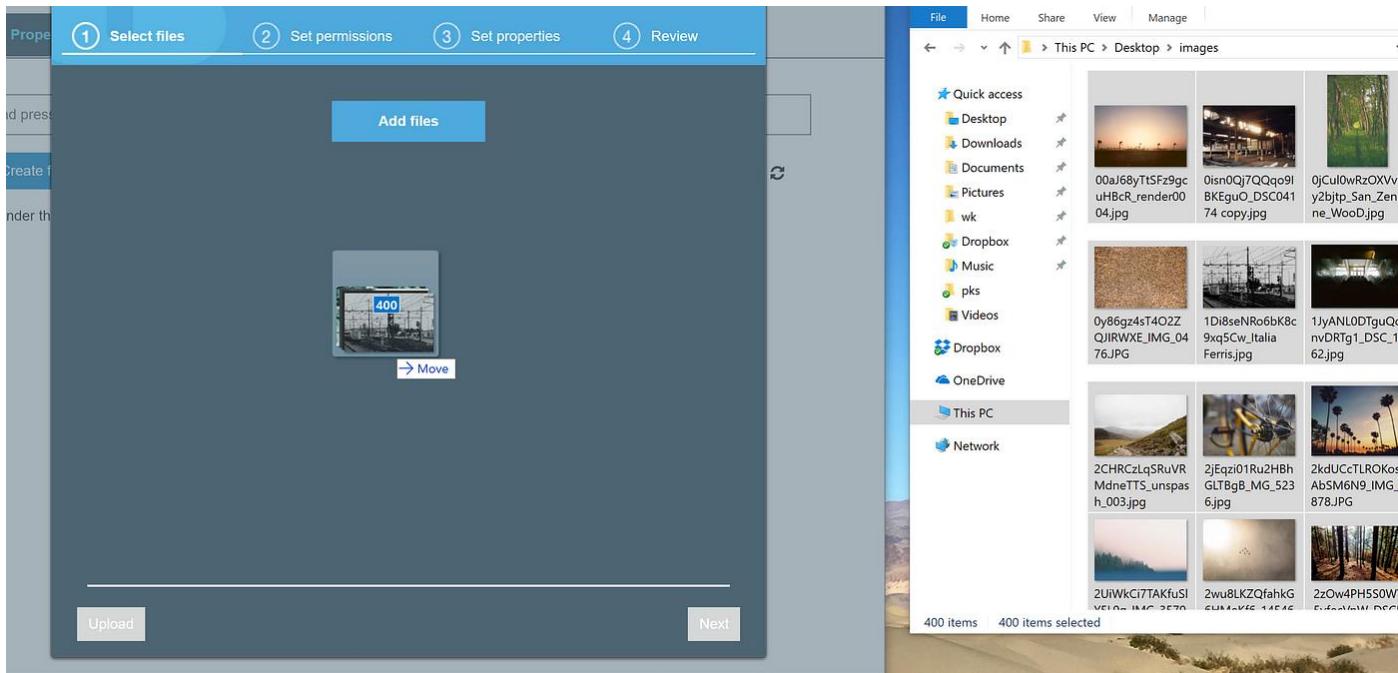
Multimedia

- What is the target: content representation
 - Metadata: name, date, size, ...
 - Visual Features: Color, Shape, Texture, etc...
 - Spatial (multiple objects)
 - Semantics Features(e.g. objects, motions, meaning, etc...)

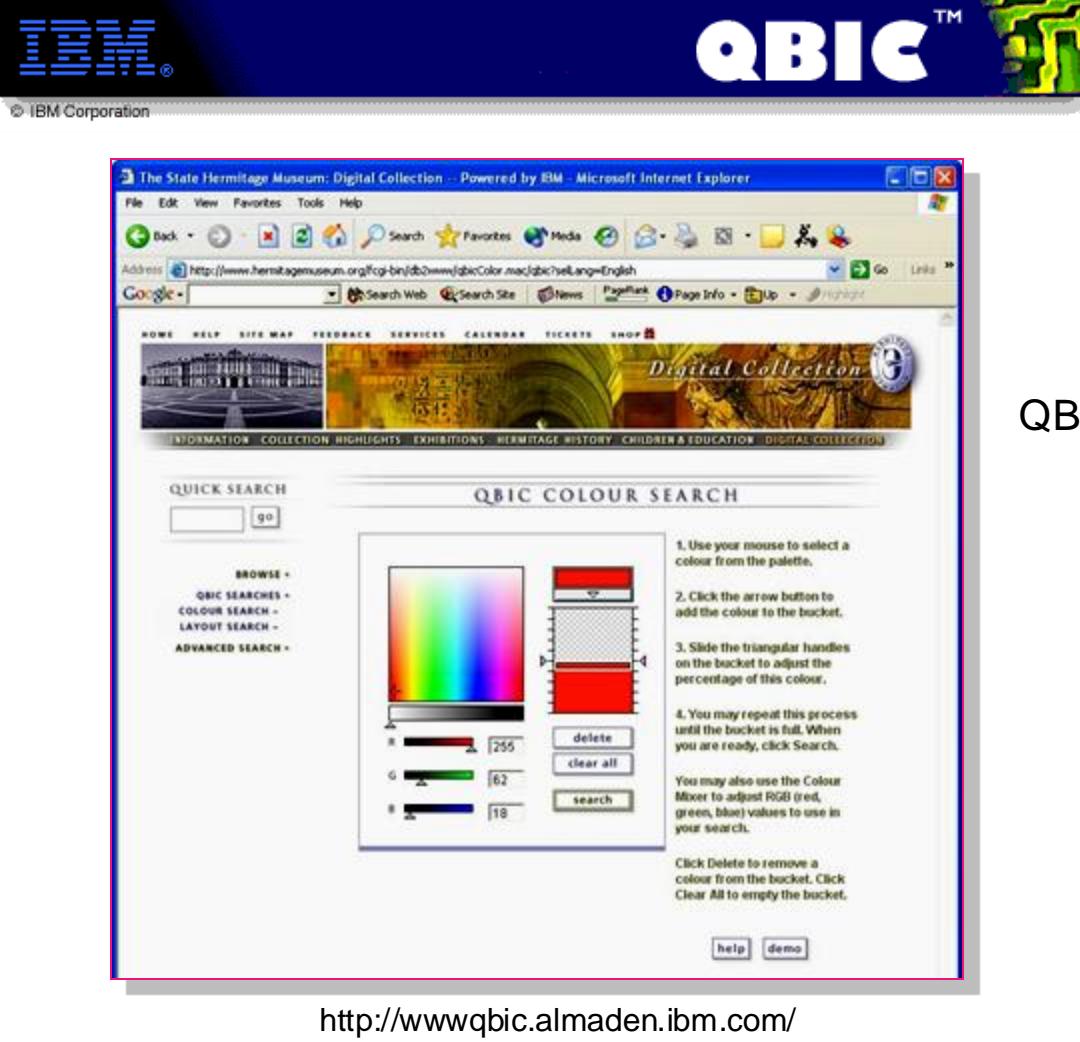


Metadata Based Retrieval: Tagging

□ Amazon MTurk



Visual Feature based Retrieval



The screenshot shows a Microsoft Internet Explorer window displaying the State Hermitage Museum's Digital Collection. The window title is "The State Hermitage Museum: Digital Collection - Powered by IBM - Microsoft Internet Explorer". The address bar shows the URL <http://www.hermitagemuseum.org/cgi-bin/db2/www/qbicColor.mac?qlang=English>. The main content area features a banner with images of the Hermitage building and interior rooms. Below the banner is a navigation menu with links to HOME, HELP, SITE MAP, RESOURCE, SERVICES, CALENDAR, TICKETS, and SHOP. A "Digital Collection" logo is visible. On the left, there is a "QUICK SEARCH" input field and a "go" button. In the center, there is a "QBIC COLOUR SEARCH" section. This section includes a color palette, a bucket for selected colors, and sliders to adjust the percentage of each color. A list of instructions provides guidance on how to use the color search feature. At the bottom, there are "delete", "clear all", and "search" buttons, along with RGB sliders showing current values (R: 255, G: 62, B: 18). A note at the bottom right explains how to use the color mixer or RGB values for searching. At the very bottom of the window are "help" and "demo" buttons.

<http://wwwqbic.almaden.ibm.com/>

QBIC (Query By Image Content)

- The Hermitage Web site uses the QBIC engine for searching archives of world-famous art.



Visual Feature based Retrieval

□ Google Image search



<http://wwwqubic.almaden.ibm.com/>

Semantic + Visual Multimedia Retrieval

Google Microphone Image Search

All Images Videos Shopping Books More Tools
Size ▾ Orange Type ▾ Time ▾ Usage Rights ▾ Clear

night light lamp cat lamp desk lamp cat nightlight girl blanket fruugo tabby ca


shutterstock.com - 2359744343
Shutterstock Red Tabby Cat Lit By Warm Stock P...


shutterstock.com - 2359744361
Shutterstock Red Tabby Cat Lit By Warm Stock ...


alamy.com
Alamy cat grey and white i...


DeviantArt.com
DeviantArt cat in windowsill sweet ...


fotolia.com
Fotolia Fluffy ginger cat on cozy knitted ...


shutterstock.com - 2359744345
Shutterstock Red Tabby Cat Lit By Warm Stock P...

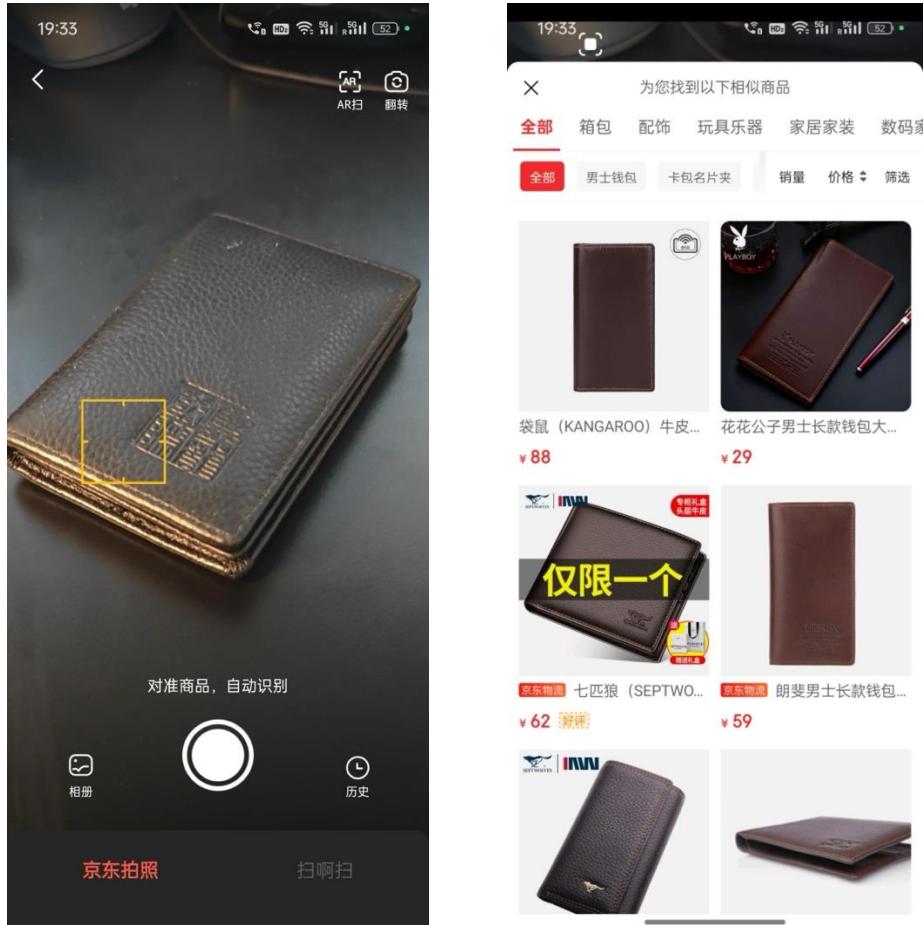

alamy.com
Alamy Orange tabby cat enjoying nap on ...


alamy.com
Alamy Fluffy ginger cat on ...

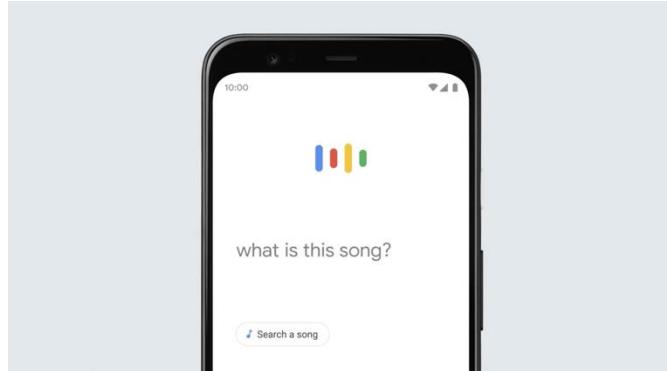

amazon.ae
Amazon.ae Light for Kids Toddler B...


dreamstime.com
Dreamstime The Cat is Brown and Red in Color,...

Semantic + Visual Multimedia Retrieval



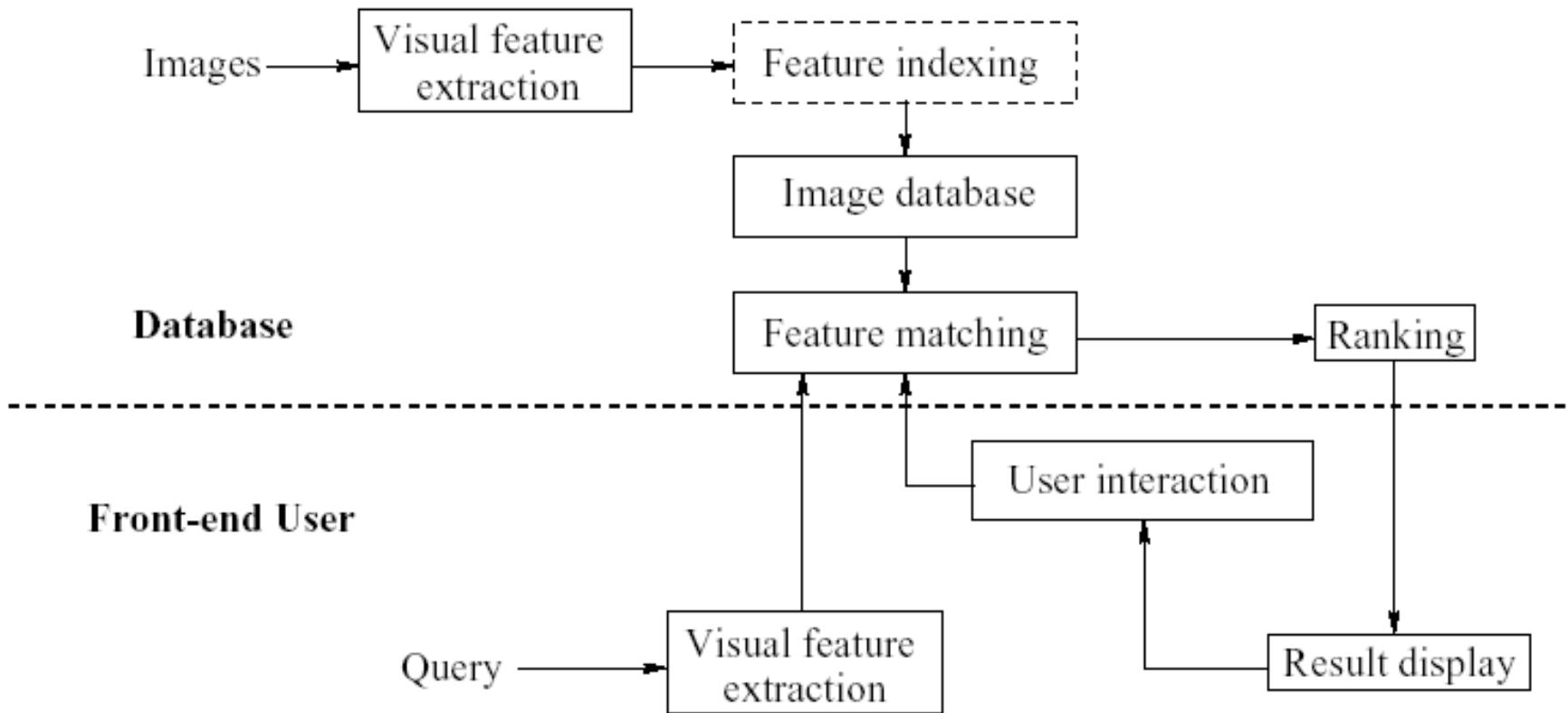
More than image



<https://research.google/blog/recognizing-pose-similarity-in-images-and-videos/>



Multimedia Retrieval



Multimedia Retrieval

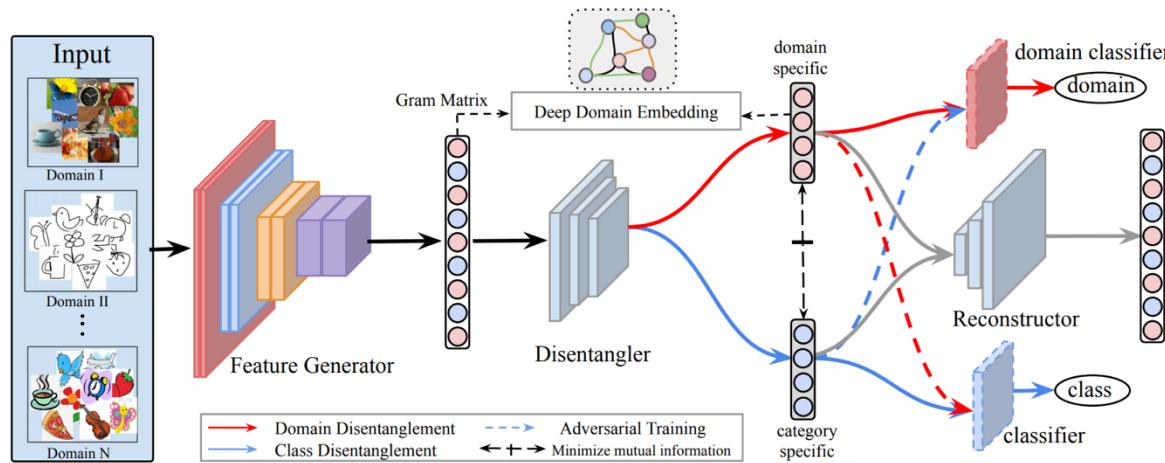
- The features
 - ▣ Carefully crafted features



https://docs.opencv.org/3.4/da/df5/tutorial_py_sift_intro.html

Multimedia Retrieval

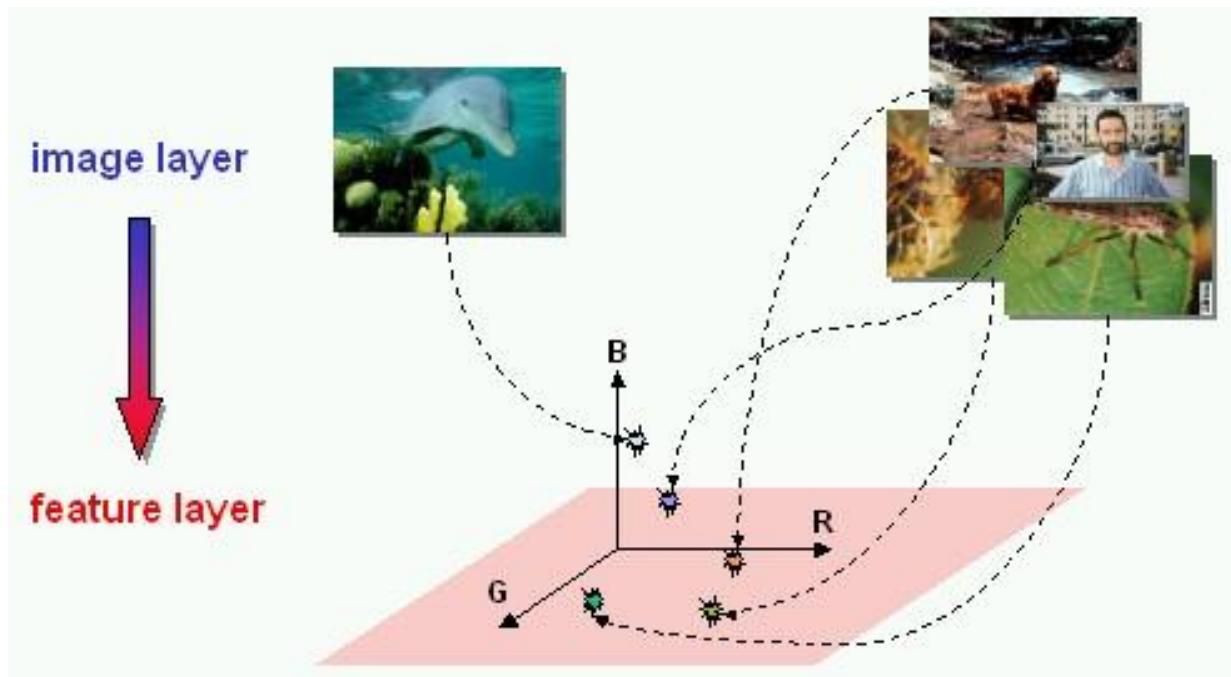
- The features
 - Deep Learning



<https://cs.stanford.edu/~liyichen/projects/domain2vec/>

Multimedia Retrieval

- How to find: similarity metric
 - ▣ Euclidean distance, Hausdorff distance, ...



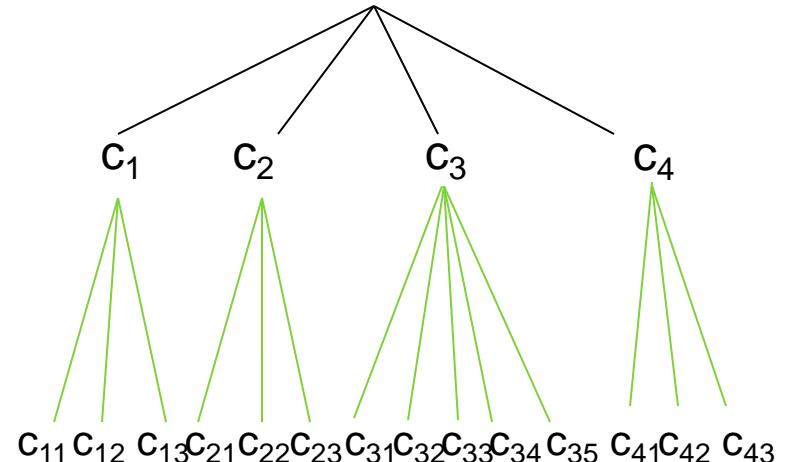
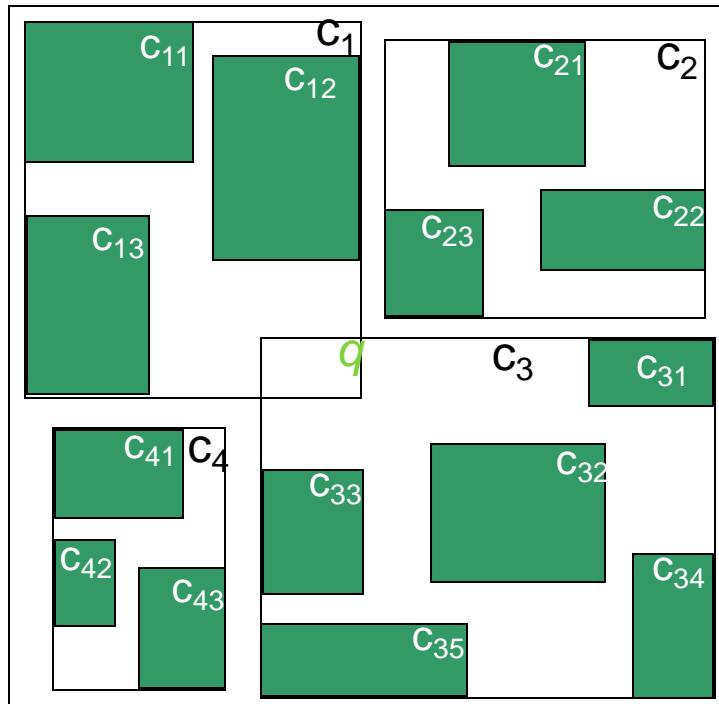
Multimedia Retrieval

- How to interact between users and systems
 - ▣ Introduce users' subjective into systems
 - ▣ Make systems more flexible and intelligent



Multimedia Retrieval

- How to organize databases: indexing



R-Tree

Multimedia Retrieval

- Feature extraction is performed to obtain multi-dimensional feature vectors characterizing multimedia contents.
- Appropriate similarity measurement is employed to measure the similarity between query item and database item.
- Feedback provides the interactions between users and systems.
- Efficient indexing techniques are employed to organize databases.



Auto Annotator



File

Image Directory

- campus_09.ppm
- campus_10.ppm
- campus_100.ppm
- campus_101.ppm
- campus_102.ppm
- campus_103.ppm
- campus_104.ppm
- campus_105.ppm
- campus_106.ppm
- campus_107.ppm
- campus_108.ppm
- campus_109.ppm
- campus_11.ppm
- campus_110.ppm
- campus_111.ppm
- campus_112.ppm
- campus_12.ppm
- campus_13.ppm
- campus_14.ppm
- campus_15.ppm
- campus_16.ppm
- campus_17.ppm
- campus_18.ppm
- campus_19.ppm
- campus_20.ppm
- campus_21.ppm
- campus_22.ppm
- campus_23.ppm
- campus_24.ppm
- campus_25.ppm
- campus_26.ppm



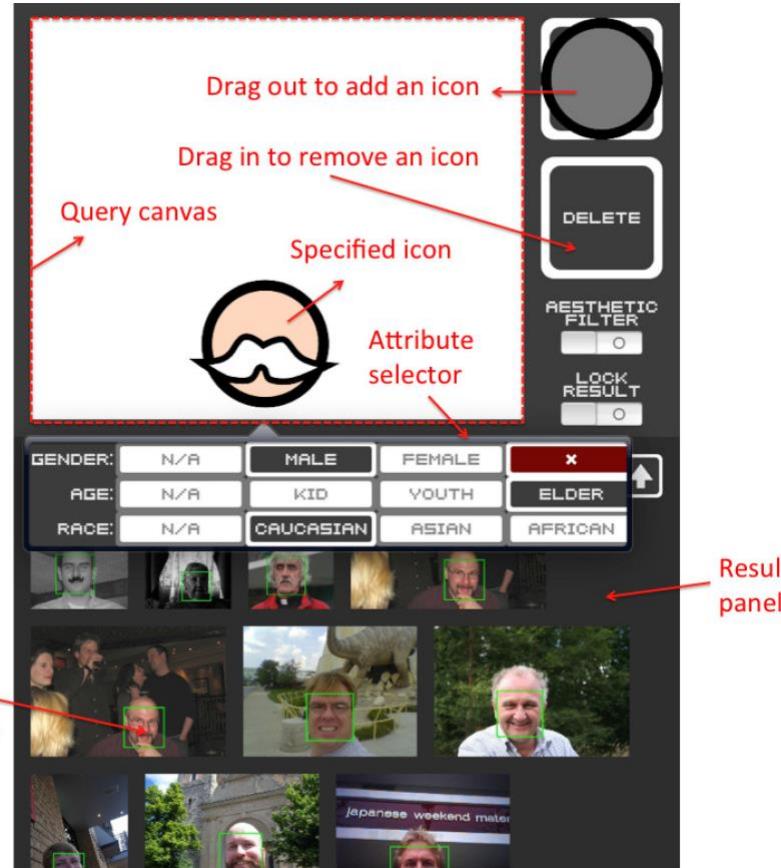
Region Annotation

tree

R	Category	Percentage
R 1	building	47%
R 2	tree	56%
R 3	building	43%
R 4	tree	51%
R 5	building	53%
R 6	tree	51%
R 7	building	63%
R 8	tree	88%
R 9	tree	60%
R 10	tree	61%
R 11	building	69%
R 12	grass	46%
R 13	sand	56%
R 14	sand	97%
R 15	sand	90%
R 16	sand	53%
R 17	tree	58%

Annotate

Face search



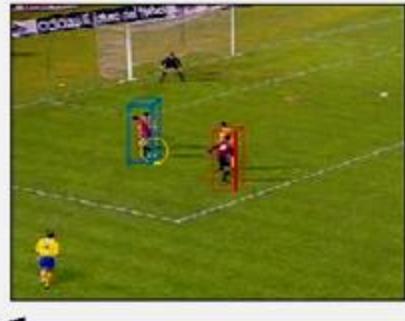
- Photo search by face positions and facial attributes on touch devices, ACM MM 2011

<http://www.youtube.com/watch?v=Mf3RlxpRVDC>
<http://www.youtube.com/watch?v=VC-t3UztT5g>

Photo Search by Face Positions and Facial Attributes on Touch Devices

Submitted to ACM Multimedia
Grand Challenge 2011

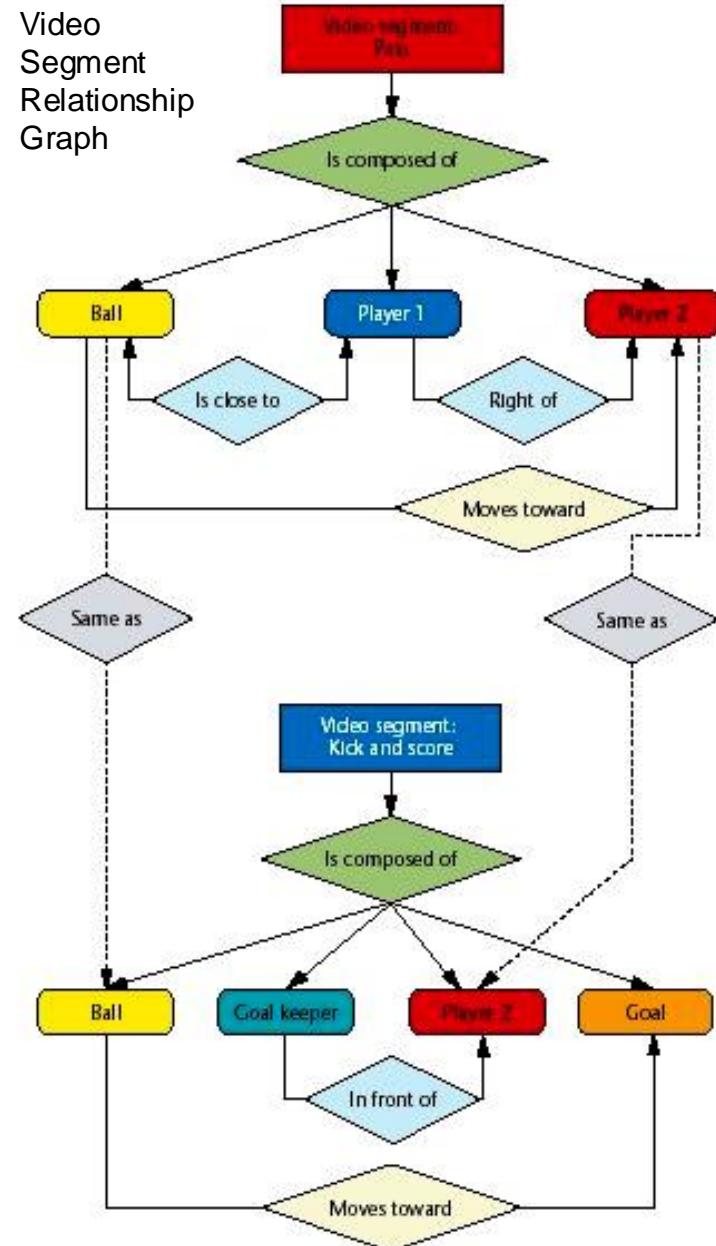
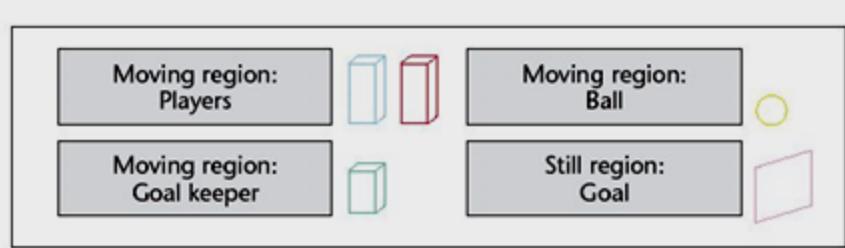
Event Analysis for Soccer Video



Video segment 1: Pass



Video segment 2: Kick and score



Content Augmentation

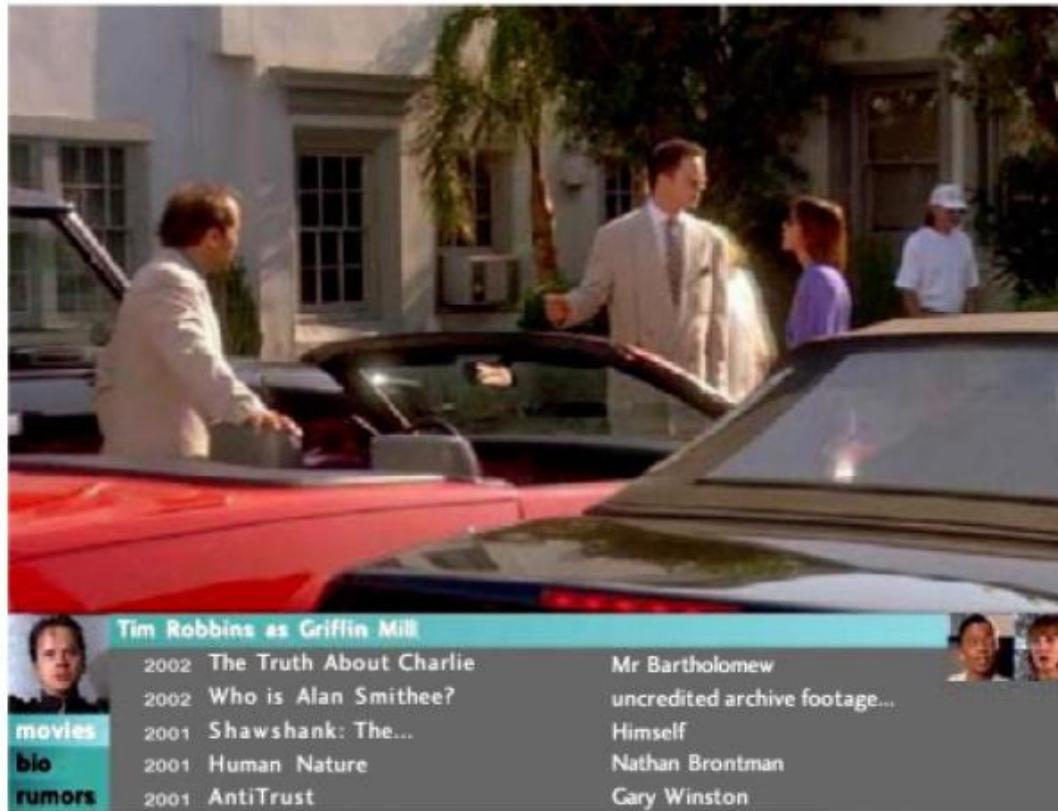


Figure 8.5. InfoSip screen.

Knowledge + QA

□ Knowledge graph



Google search results for "dubai":

Web Images Maps News Videos More Search tools

About 91,100,000 results (0.33 seconds)

Dubai - TripAdvisor.com.au
www.tripadvisor.com.au/Dubai
Find Dubai Hotels and Dubai Accommodations!
Best Travel Review Website, 2013 - British Travel Awards
TripAdvisor has 1,944,690 followers on Google+
TripAdvisor HomePage - Dubai Restaurants - 453 Hotels in Dubai

Dubai - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Dubai
Dubai (/duːˈbaɪ/; Arabic: دُبَيّ; IPA: [dub'bejj]) is an emirate in the United Arab Emirates (UAE) federation. The main city of the emirate is also ...
United Arab Emirates - Mohammed bin Rashid Al ... - Abu Dhabi - The Dubai Mall

Definitely Dubai - The Official Tourism Portal Of The City Of ...
www.definitelydubai.com
Dubai is a unique destination that is both a dynamic business centre and a tourist paradise, offering more attractions, shopping, fine dining and quality hotels.
Things To Do - Definitely Dubai - Things to See - Why Dubai?

News for dubai

Too good to not be a bubble? Dubai's DAMAC real estate developer triples profits in 2013
Al-Bawaba - 2 hours ago
ArabianBusiness...
Dubai luxury real estate developer DAMAC tripled its 2013 net profit and plans more new projects this year, it said, dismissing growing ...

Dubai to have sustainable residential city for Emiratis
Khaleej Times - 1 day ago

Dubai food festival: A flavor for every taste
Arab News - 7 hours ago

More news for dubai

Dubai
City in United Arab Emirates
Dubai is an emirate in the United Arab Emirates federation. The main city of the emirate is also called Dubai. The emirate is located on the southeast coast of the Persian Gulf and is one of the seven emirates that make up the country. [Wikipedia](#)
Population: 2.106 million (2013)
Weather: 23°C, Wind N at 10 km/h, 73% Humidity
Local time: Wednesday 10:38 am
Gross domestic product: 82.11 billion USD (2008)
Colleges and Universities: American University in Dubai, SPJCM, more

Points of interest

Burj Khalifa The World The Dubai Mall Jumeirah Beach Hotel Dubai Creek

Feedback

Google voice search: <http://www.youtube.com/watch?v=-cZ4mdrlWYM>

Connecting the media

StoryImaging

New Story Go

Processed Story

Java is a programming language originally developed by James Gosling at Sun Microsystems (which is now a subsidiary of Oracle Corporation) and released in 1995 as a core component of Sun Microsystems' Java platform. The language derives much of its syntax from C and C++ but has a simpler object model and fewer low-level facilities. Java applications are typically compiled to bytecode (.class file) that can run on any Java Virtual Machine (JVM) regardless of computer architecture. Java is a general-purpose, concurrent, class-based, object-oriented language that is specifically designed to have as few implementation dependencies as possible. It is intended to let application developers "write once, run anywhere." Java is currently one of the most popular programming languages in use, and is widely used from application software to web applications.

Final Imaging

Video Summarization



<http://sydney.edu.au/engineering/it/~gguia5470/keyframe-demo/>



THE UNIVERSITY OF SYDNEY

School of Computer Science

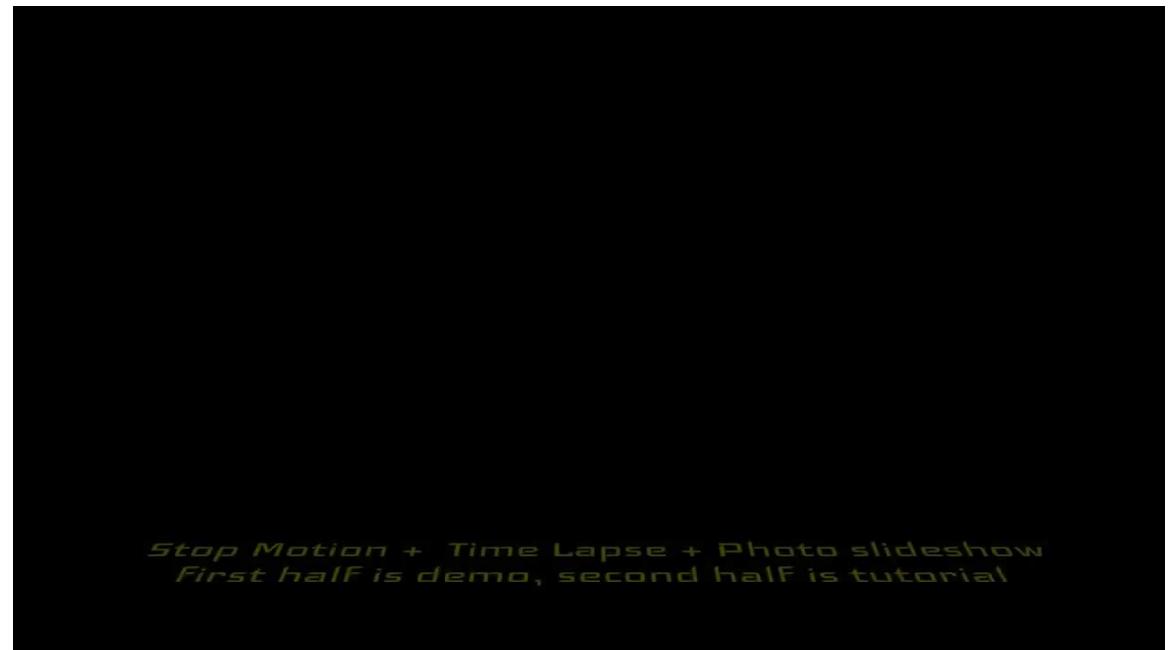




PicPac Stop Motion



<http://picpac.tv/>



INCUBATE

<http://incubate.org.au/>



THE UNIVERSITY OF SYDNEY

School of Computer Science



Video Synopsis



<http://www.vision.huji.ac.il/video-synopsis/>

Large Language Model Search

ChatGPT 4o > Find usyd building information

The University of Sydney's Camperdown and Darlington campuses are home to a rich array of architecturally significant buildings, each contributing to the university's storied history and vibrant academic environment. Here's an overview of some notable structures:

Main Quadrangle and Great Hall

Designed by architect Edmund Blacket, the Main Quadrangle and Great Hall are iconic examples of Victorian Academic Gothic Revival architecture. Constructed between 1855 and 1862, these sandstone buildings feature intricate carvings, stained-glass windows, and a hammerbeam roof in the Great Hall. The Quadrangle serves as a central hub for university activities and ceremonies. [EN.WIKIPEDIA.ORG](#)

New Law School Building

Opened in 2009, the New Law School Building was designed by the architectural firm Francis-Jones Morehen Thorp (FJMT). Located on Eastern Avenue, this modern facility includes a moot court, law library, and flexible teaching spaces. Its design emphasizes transparency and openness, fostering a collaborative learning environment. [EN.WIKIPEDIA.ORG](#)

Charles Perkins Centre

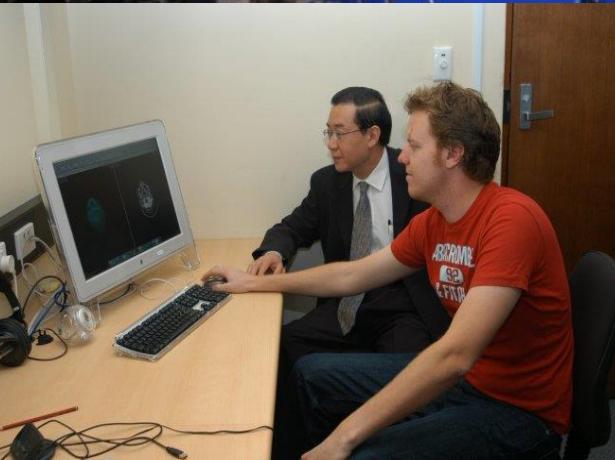
Completed in 2014, the Charles Perkins Centre is a state-of-the-art research and education hub focusing on health-related fields. The building spans six storeys with four basement levels, housing advanced laboratories, teaching spaces, and clinical facilities. Its distinctive full-height atrium and modern design make it a standout on campus. [EN.WIKIPEDIA.ORG](#)

Abercrombie Building

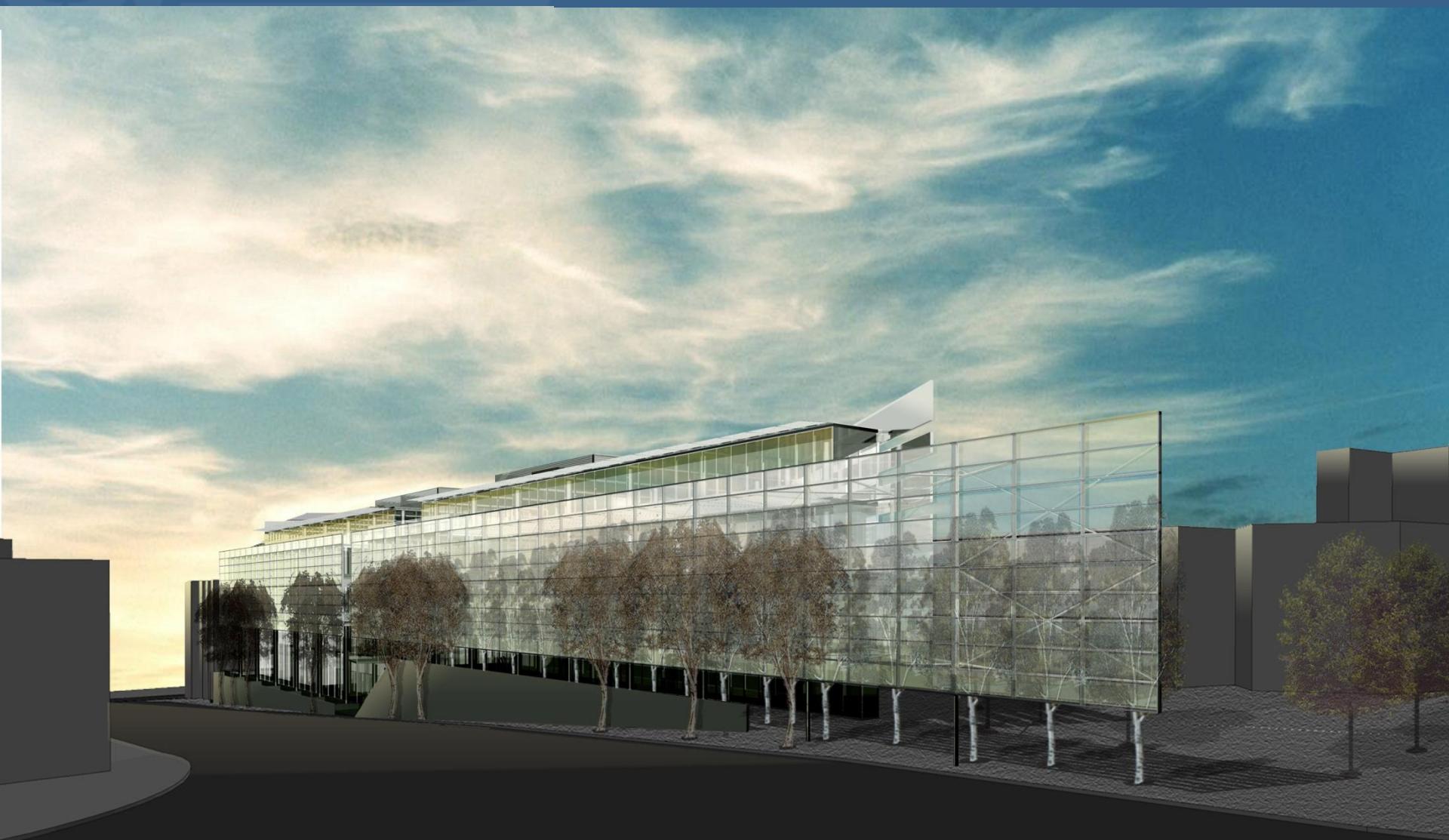
Message ChatGPT

+ 4o 0 ↑

Multimedia Laboratory



School of Computer Science Building



School of Computer Science Building



Course Description

- The explosive growth of **multimedia data**, including text, audio, images and video has imposed unprecedented challenges for search engines to meet various information needs of users. This unit provides students with the necessary and updated knowledge of this field in the context of **big data**, from the information retrieval basics of a search engine, to many advanced techniques towards **next generation search engines**, such as content based image and video retrieval, large scale visual information retrieval, and social media.



Multimedia Retrieval

Week	Topics
01	Introduction
02	Textual Information Retrieval
03	Web Information Retrieval
04	Multimedia Basics
05	Multimedia Information Retrieval (I)
06	Multimedia Information Retrieval (II)
07	Social Media
08	Large Scale Retrieval
09	Recommender Systems
10	Adv 1: Large Models
11	Adv 2: Information Summarisation
12	Project Presentation
13	Course Review

break



Administrative Issues

- Lectures / Labs
- Assessment
- Resources

Canvas: <https://canvas.sydney.edu.au/>

Double check, if anything is unclear or inconsistent!



Lectures & Labs

- Lectures
 - ▣ **Wed 18:00-20:00**
 - ▣ **J02.03.302.PNR Building.PNR Lecture Theatre (2) 302**
- Labs
 - ▣ **Week 02 – Week 12**
 - ▣ **Find your session (e.g., Lab 115, School of CS Building)**
 - ▣ Python
 - Any other languages (e.g. C/C++ and Java) are also acceptable, though no support for these languages
- Consultation
 - ▣ **Tue 17:00 – 18:00**
 - Online through zoom (see Canvas)
 - ▣ By appointment



Assessment

Tasks	Start	Due	Marks
Project (Group)	Week 02 Wed, 05/03	Proposal: Week 06 Tue, 01/04	10%
		Final: Week 12 Tue, 20/05	20%
Homework (Individual)	Week 02 Wed, 05/03	Week 08 Tue, 15/04	15%
Final Exam	TBA	N/A	55%

**! Please let me know if you notice any discrepancy of various information, such as assessment dates.
It would be your responsibility if you make a wrong assumption.**



Expectations

- Students attend scheduled classes, and devote an *extra* 6-9 hrs per week
 - ▣ doing assessments
 - ▣ preparing and reviewing for classes
 - ▣ revising and integrating the ideas
 - ▣ practice and self-assess
- Students are responsible learners
 - ▣ Participate in classes, constructively
 - Respect for one another (criticize ideas, not people)
 - Humility: none of us knows it all; each of us knows valuable things
 - ▣ Check Canvas site at least once a week!
 - ▣ Notify academics whenever there are difficulties
 - ▣ Notify group partners honestly and promptly about difficulties



Expectations

- To understand
 - ▣ Concepts, contents, and principles
- To be skillful
 - ▣ Practicing algorithms
- To know
 - ▣ Solving problems
 - ▣ Analyzing results
 - ▣ Summarizing ideas
- To be creative

The people who are crazy enough to think they can change the world are the ones who do.

– *Think Different, Apple, 1997*



Knowledge is not enough, we must apply it.
The will is not enough, we must act.

Assessment

- To Pass this course, you must
 - ▣ Score at least 50% overall, and
 - ▣ Score at least 40% in the final exam.
- Late submissions of progressive assessments
 - ▣ Deduct 5% of the assessment per day
 - ▣ Receive zero for being late for more than 10 days
 - ▣ http://sydney.edu.au/policies/showdoc.aspx?recnum=PDO_C2012/267&RendNum=0#page=6
- Academic Integrity
 - ▣ Plagiarism: NO
 - ▣ Outsourcing: NO
 - ▣ See more details on the course website

Using AI tools such as ChatGPT is not allowed!

Similarity detecting software will be used!

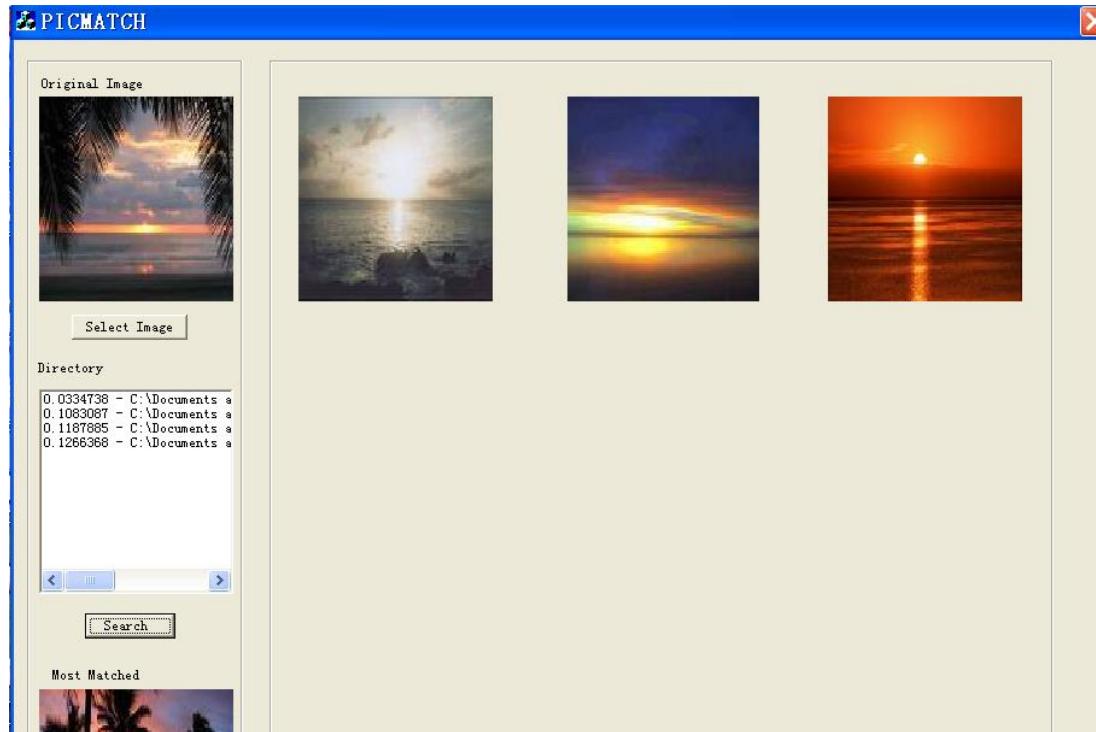
Special Consideration (University Policy)

- If your performance on assessments is affected by illness or misadventure
- Follow proper bureaucratic procedures
 - ▣ Have professional practitioner sign special USyd form
 - ▣ Submit application for special consideration online, upload scans
 - ▣ Note you have only a quite short deadline for applying
 - ▣ http://sydney.edu.au/current_students/special_consideration/
- Also, notify coordinator by email as soon as *anything begins to go wrong*
- There is a similar process if you need special arrangements eg for religious observance, military service, representative sports

Always refer to the latest university policy!



Sample Projects of Previous Students



Sample Projects of Previous Students

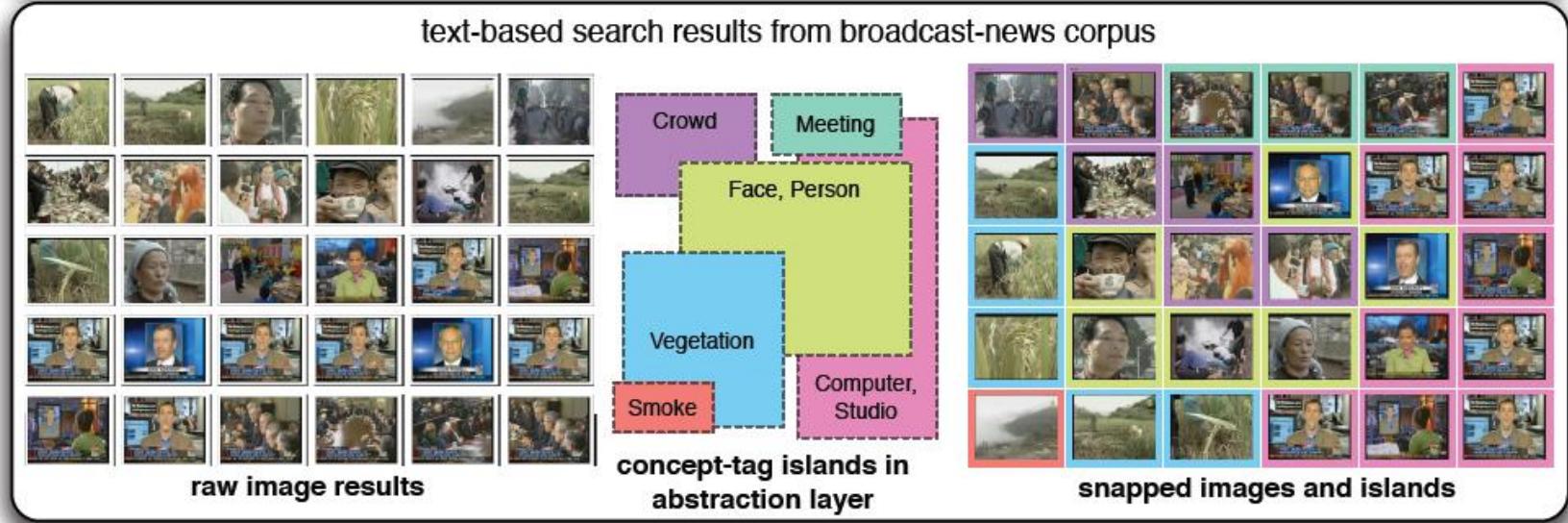
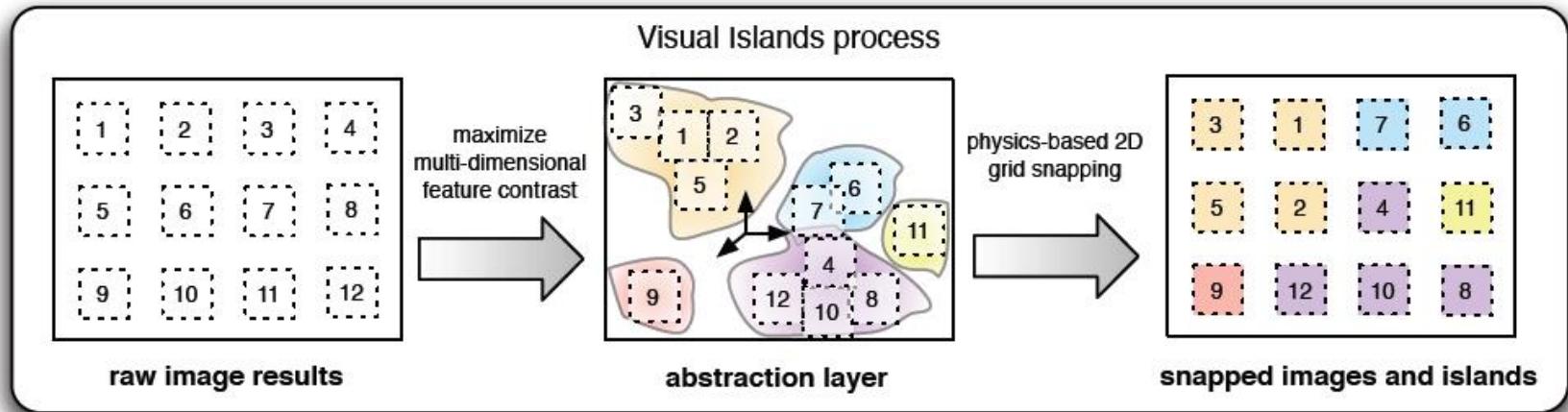
[Simple Search](#) [Advanced Search](#) [Similar Search](#)

[中文](#)

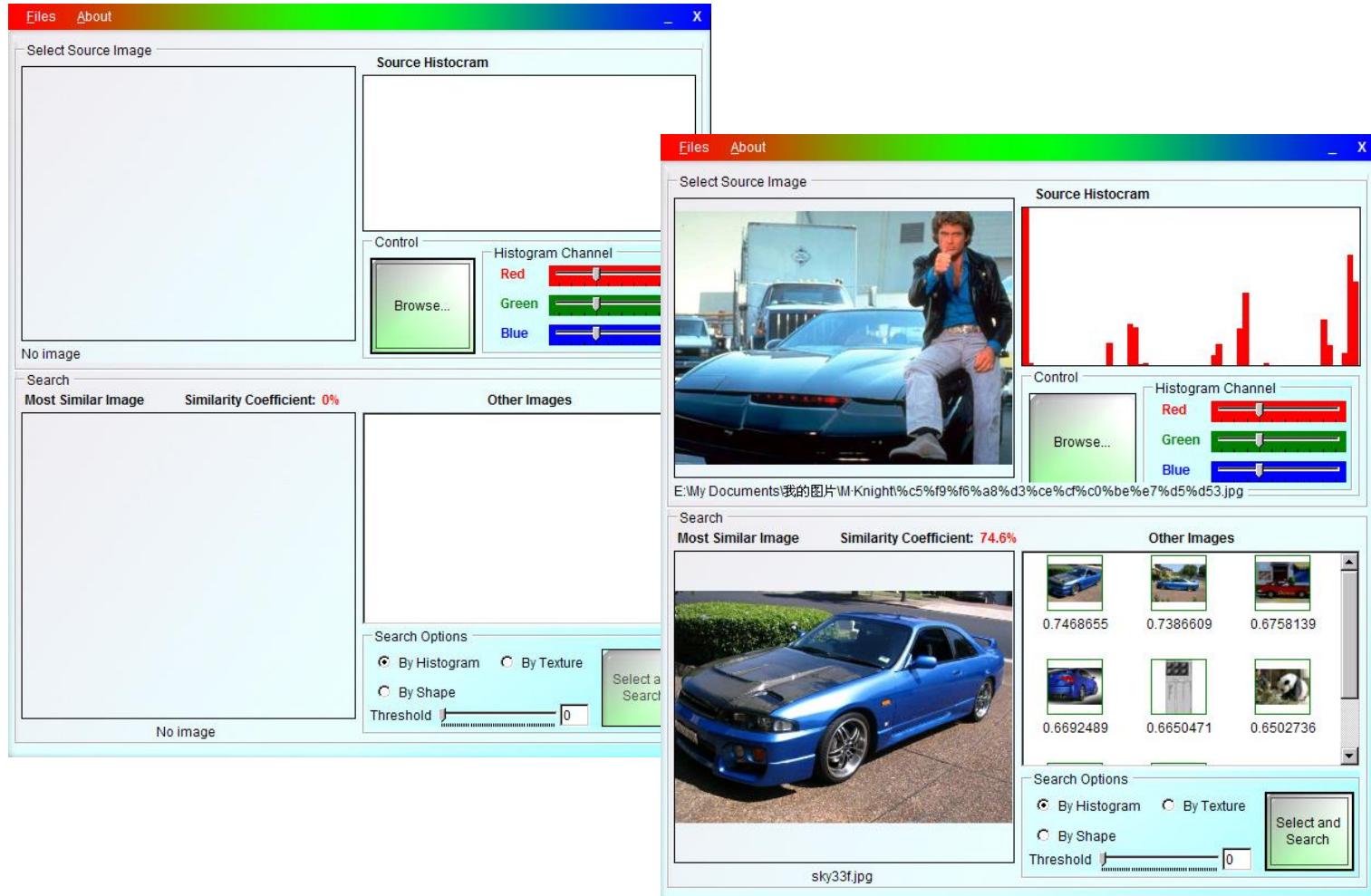
Picture Picker



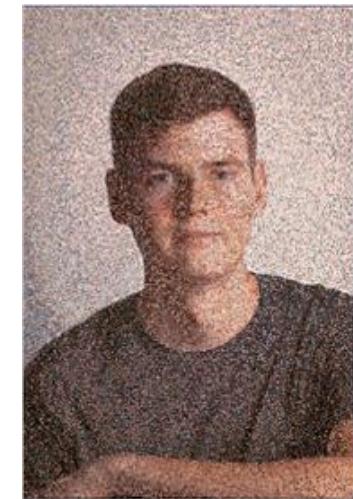
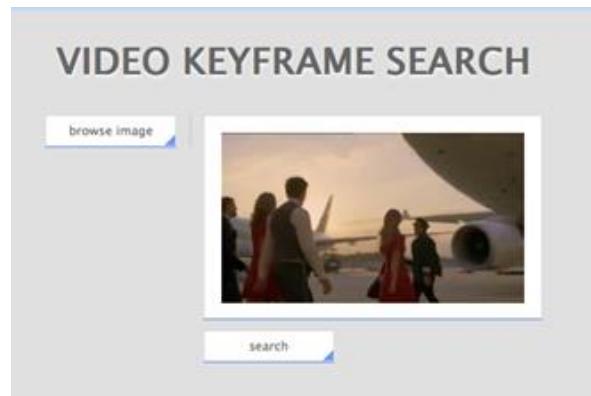
Sample Projects of Previous Students



Sample Projects of Previous Students

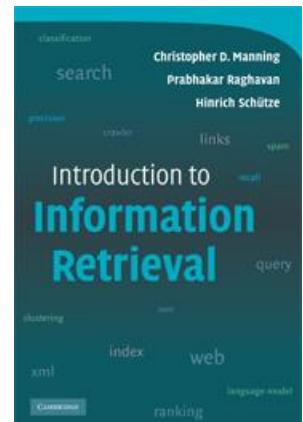
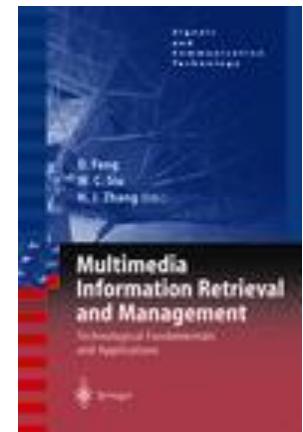


Sample Projects of Previous Students



Resources

- Lecture and Lab notes
- Key reference books
 - ▣ Multimedia Information Retrieval and Management - Technological Fundamentals and Applications, D. Feng, W. C. Siu, and H. J. Zhang, Springer-Verlag, 2003.
 - ▣ Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze, Introduction to information retrieval, Cambridge University Press, 2008. [Library]
 - <http://nlp.stanford.edu/IR-book/information-retrieval-book.html>
- Online resources
 - ▣ ACM digital libraries: <http://dl.acm.org/>
 - ▣ IEEE digital libraries: <http://ieeexplore.ieee.org/>



Projects on Multimedia (12cpt/18cpt)

- Multimedia Information Retrieval
 - ▣ Image/video/3D/audio/text content analysis
 - ▣ Summarization
 - ▣ Recommendation
- Human Motion Analysis, Modeling, Animation, and Synthesis
- Surgical Video Analysis
- Medical Image Analysis
- Hyperspectral Image Analysis
-



Need to Know

- Lab starts from next week (Week 02).
 - Preparation can start from this week.
- Think of a **Multimedia Retrieval** application