CS 3570 多媒體技術概論
Introduction to Multimedia Technology

- **Class Meeting:** T5F5F6 台達館 108
- **Instructor:** 賴尚宏, 台達館 636
  Phone: ext. 42958
  Email: lai@cs.nthu.edu.tw
- **Office Hours:** T2, F7 or by appointment
- **Teaching Assistants:**

  姚詩軒(dachshund_dog@hotmail.com, ext. 80933)台達館 722室

  陳星宇(andy19933@gmail.com , ext. 80933)台達館 722室

  陳書屏(scarletclaw24@gmail.com , ext. 80933)台達館 722室

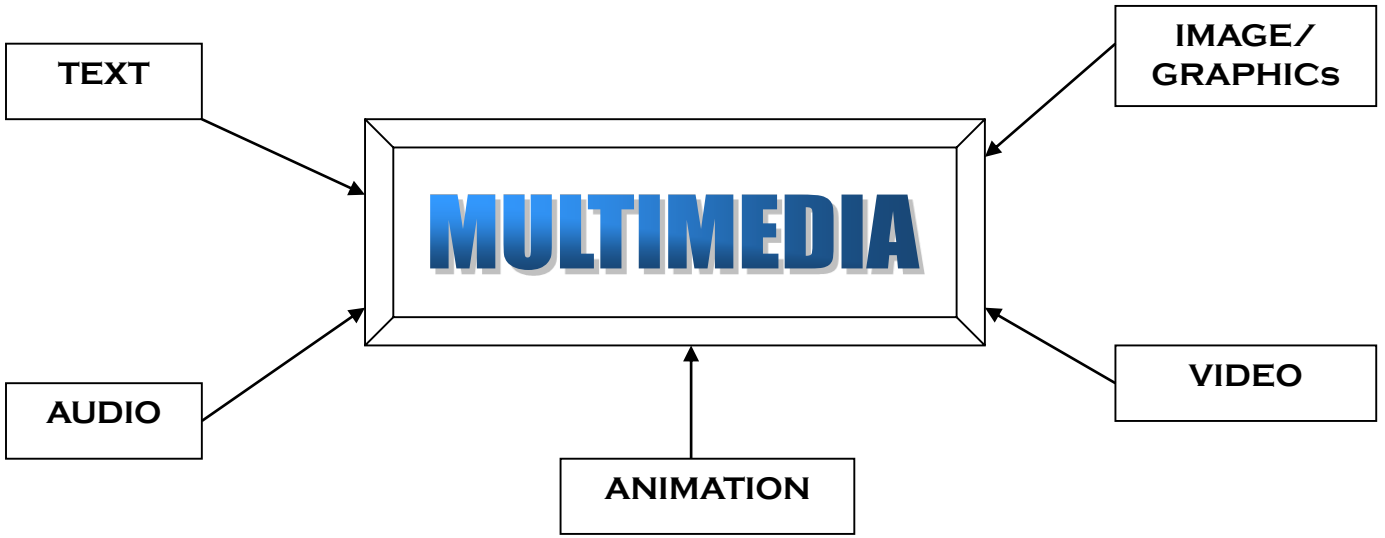
  林宏縉(vtsh.jn@gmail.com , ext. 80933)台達館 722室

Course Objective

- This course will introduce fundamental techniques for digital image/audio/video/graphics representation, compression, processing, and analysis.
- Students will learn the basic knowledge of the multimedia signal processing techniques, and practical implementations of various multimedia applications.

Definition of Multimedia

- Multimedia is a combination of text, image, graphic, sound, animation, and video that is delivered interactively to the user by electronic or digitally manipulated means.



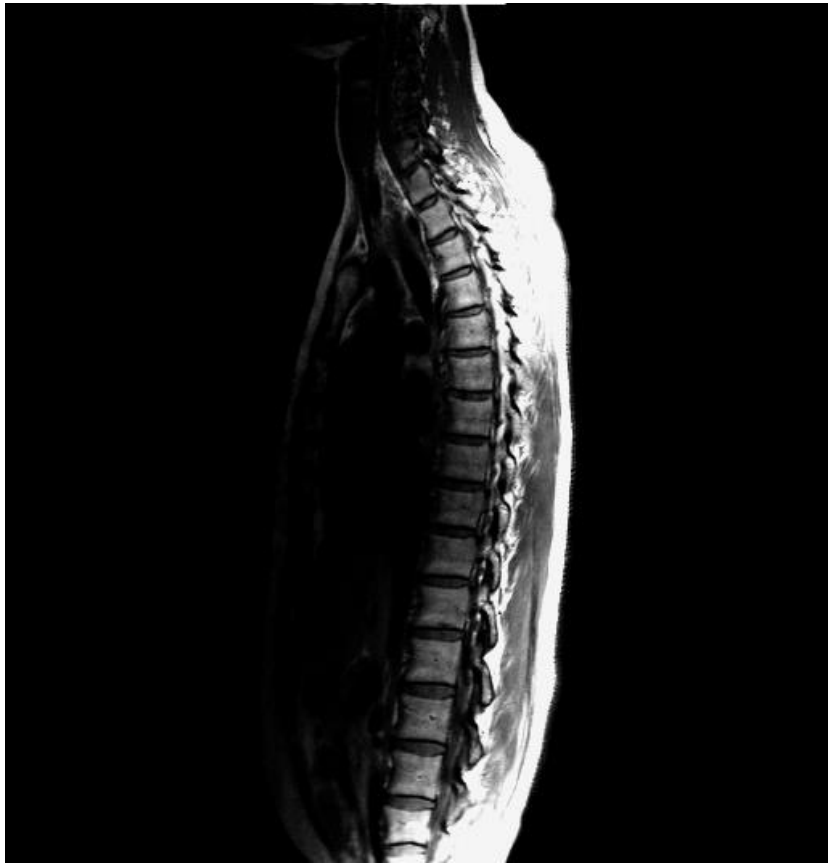
Course Contents

- Digital Data Representation and Communication
- Digital Image Representation & Processing
- Digital Audio Representation & Processing
- Digital Video Representation and Processing
- Computer Graphics
- Machine Learning for Multimedia Analysis
- AR & VR

Image Enhancement Example

Adjusting the image histogram to improve image contrast

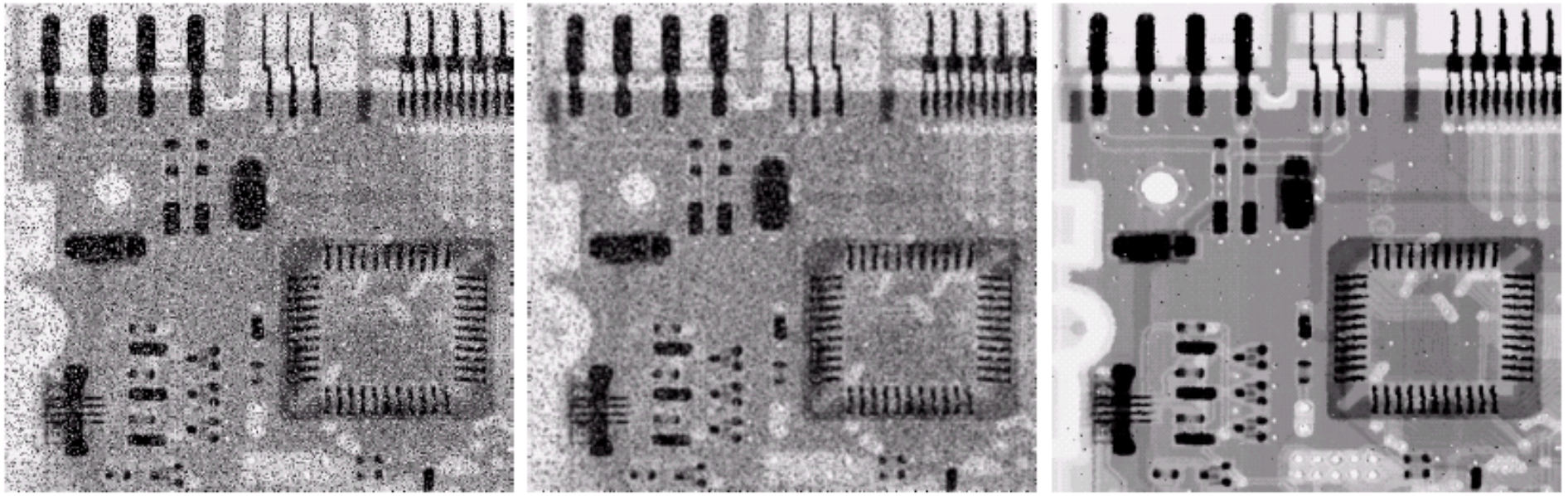
Bias Field Correction for Medical Images



Original MRI

After Correction

Image Denoising Example



a b c

FIGURE 3.37 (a) X-ray image of circuit board corrupted by salt-and-pepper noise. (b) Noise reduction with a 3×3 averaging mask. (c) Noise reduction with a 3×3 median filter. (Original image courtesy of Mr. Joseph E. Pascente, Lixi, Inc.)

Application of image filtering

Bilateral Filtering

Original image

Filtered image

From B. Weiss, Fast Median and Bilateral Filtering, SIGGRAPH'2006

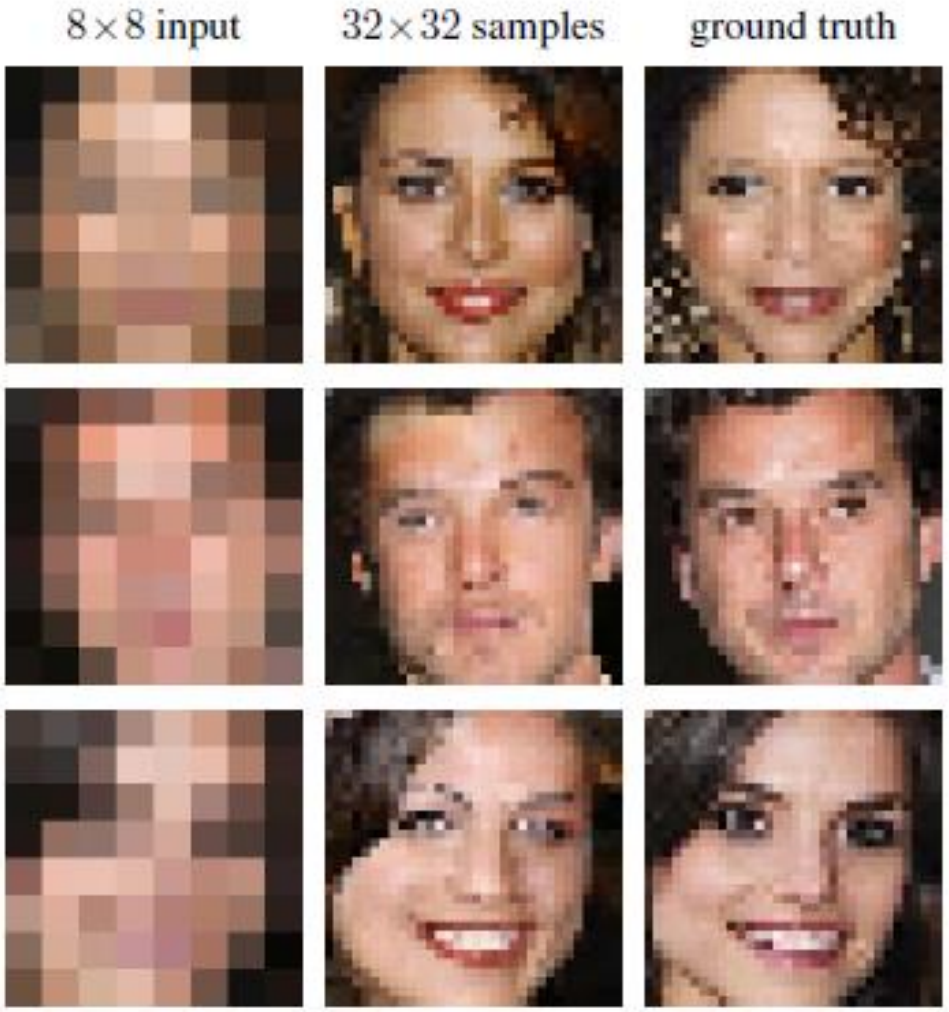
Image Super-Resolution

original

4x scaling by bi-cubic interpolation

4x scaling by an advanced method

Image Super-Resolution



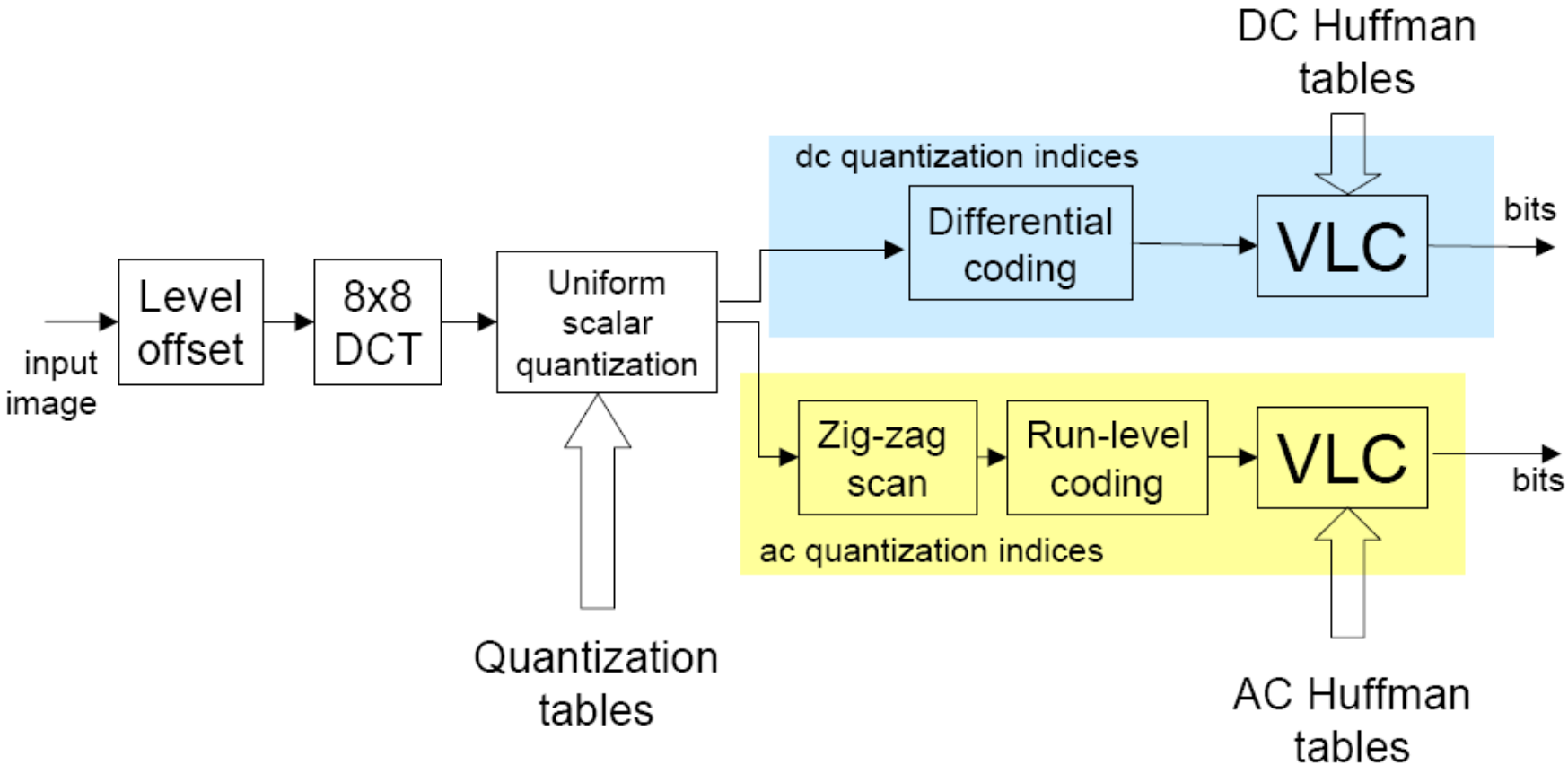
From “Pixel Recursive Super Resolution”, Dahl et al., https://arxiv.org/abs/1702.00783

Image Compression

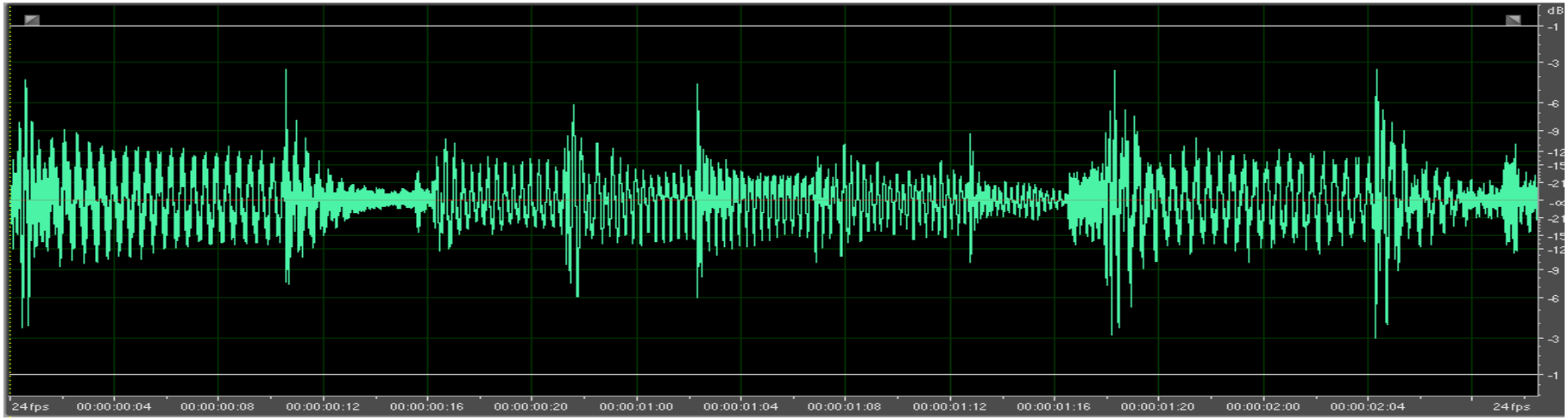
24k bytes with JPEG (Q=50)

6M bytes with raw image format
(without compression)

JPEG Image Compression



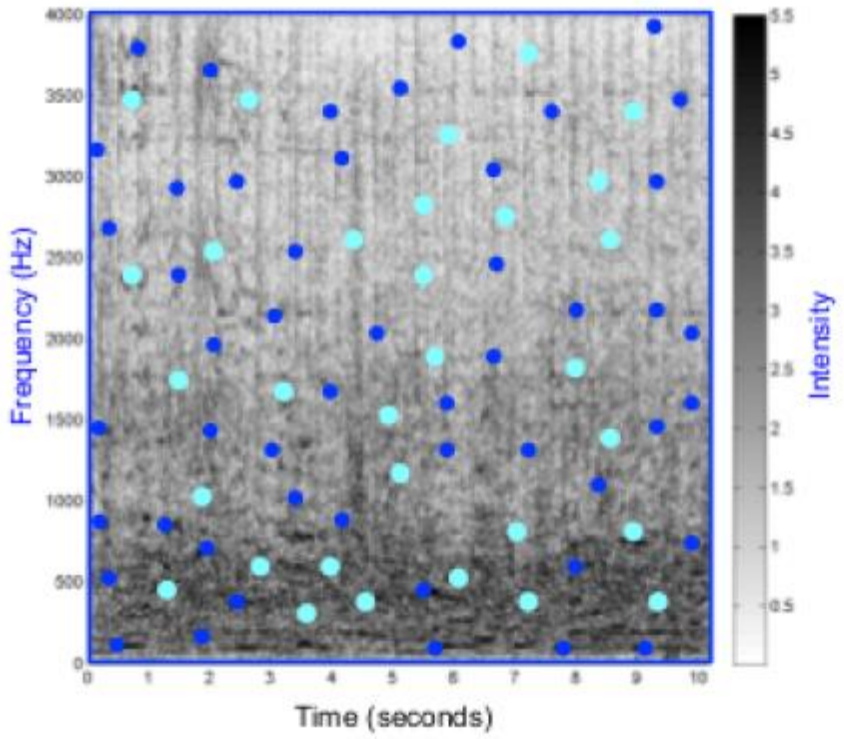
Audio Signal Processing



- Audio compression
- Noise reduction
- Frequency-domain processing

Audio Recognition/Matching

Shazam audio fingerprints:



Steps:

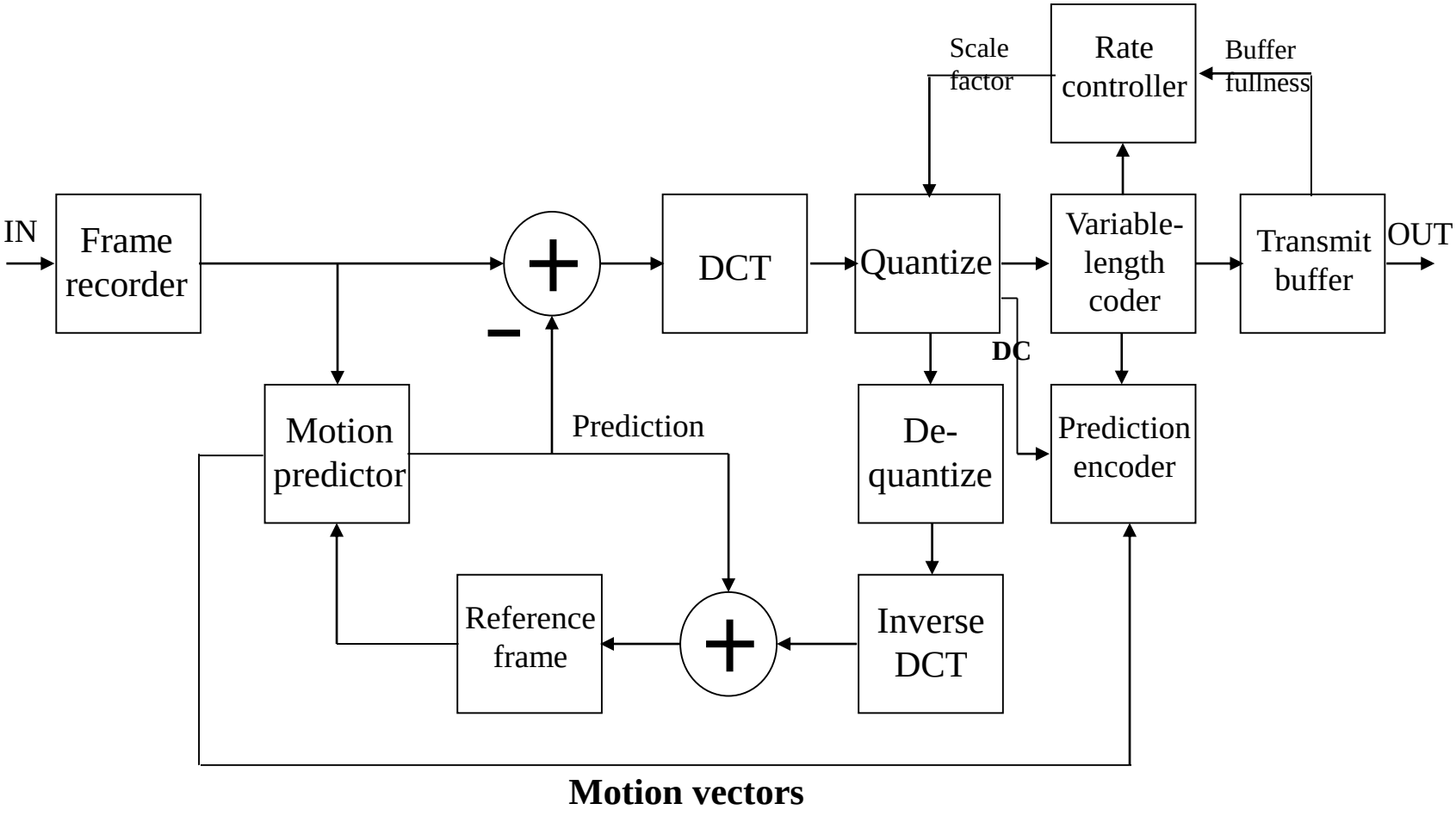
1. Spectrogram
2. Peaks / differing peaks

Robustness:

- Noise, reverb, room acoustics, equalization
- Audio codec
- Superposition of other audio sources

http://www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf

Video Compression



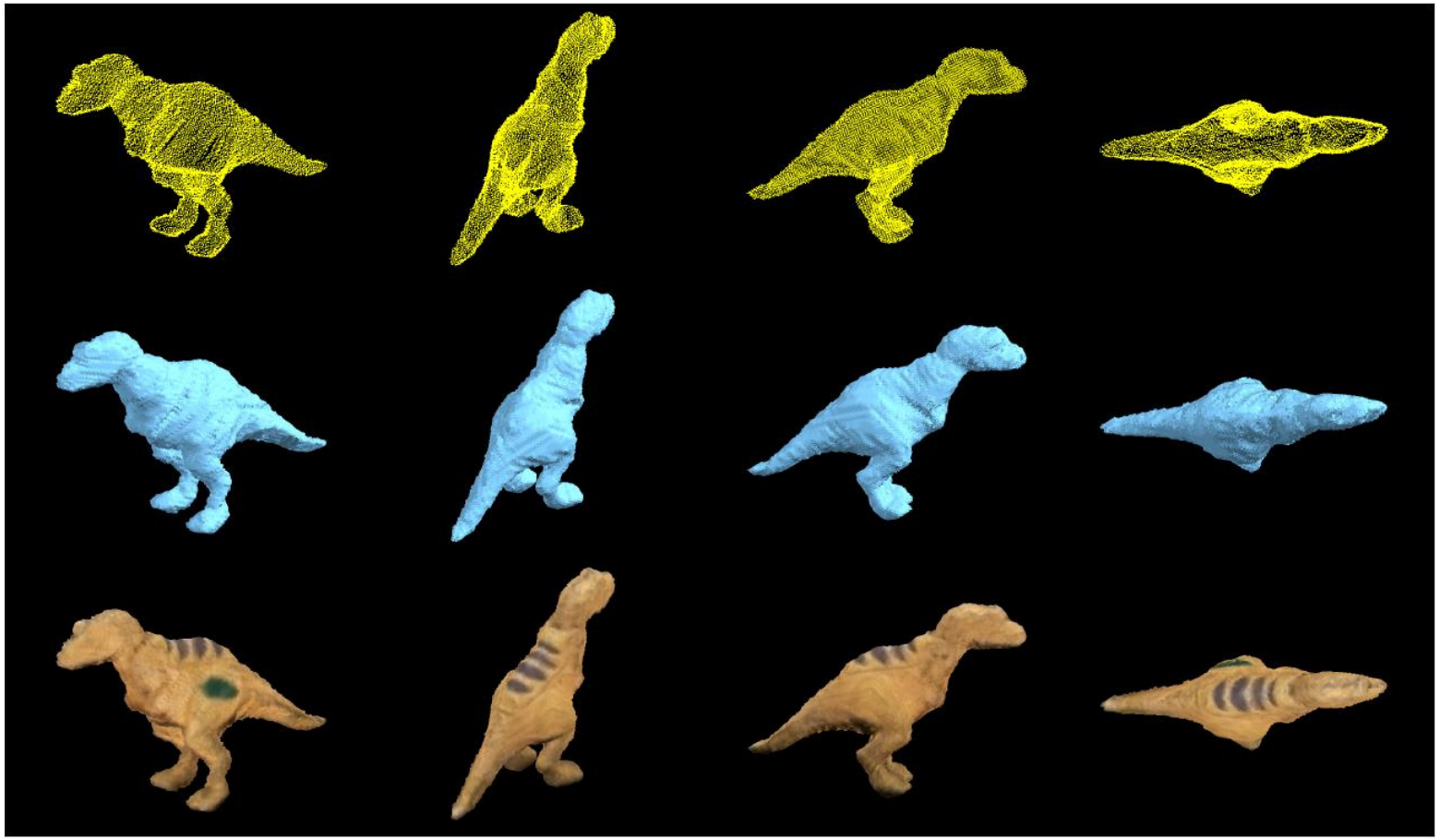
MPEG Video Encoder

Video Stabilization

Adopted from:
http://public.hronopik.de/vid.stab/features.php?lang=en

3D Computer Graphics

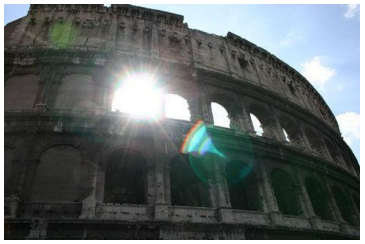
- 3D Modeling
- Image Rendering



Machine Learning for Multimedia Analysis

Query image | **Ranking list**

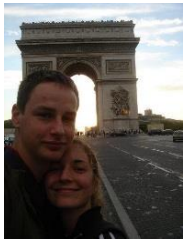


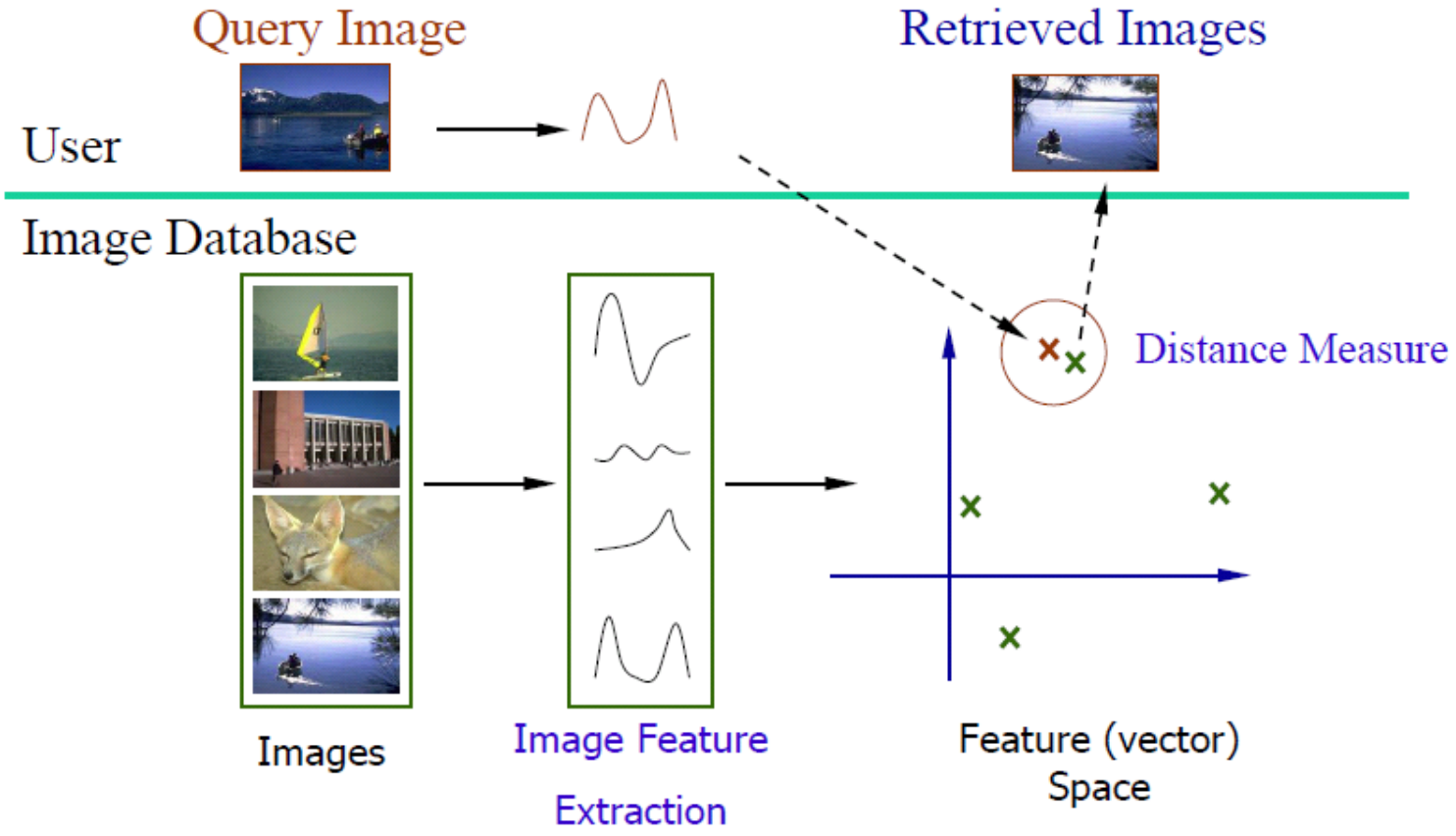






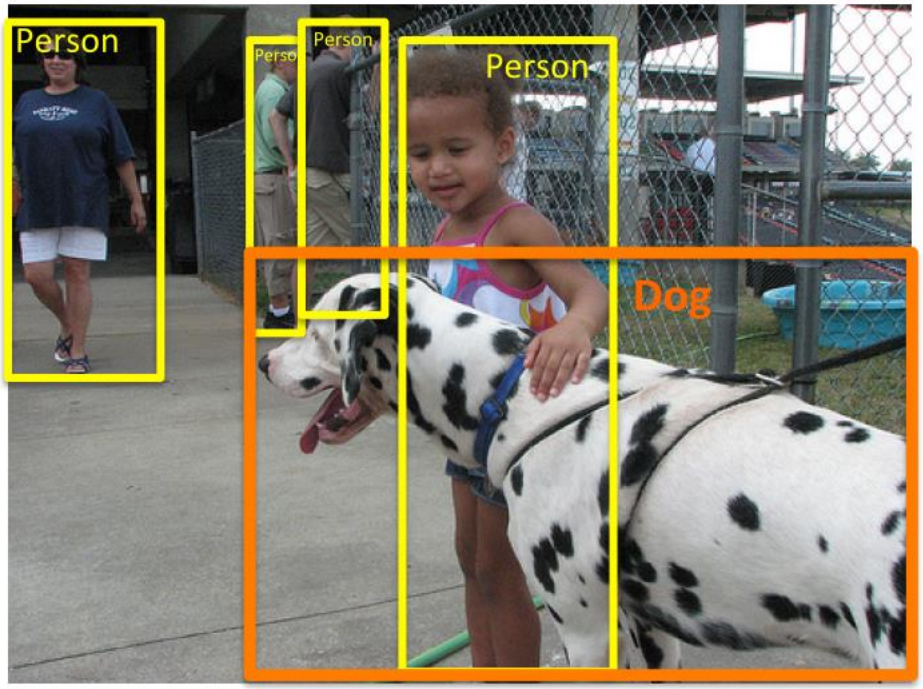



Content-Based Image Retrieval (CBIR) from Image Database



IMAGENET Large Scale Visual Recognition Challenge (ILSVRC)

200 object classes	456,567 images	DET
1000 object classes	1,431,167 images	CLS-LOC

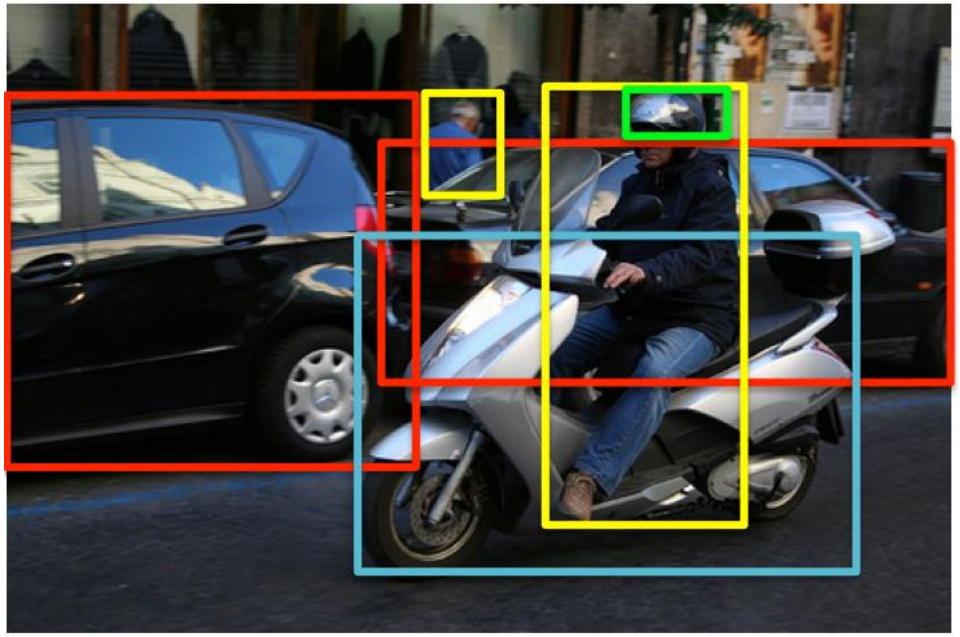


Dramatic improvement thanks to Deep Learning

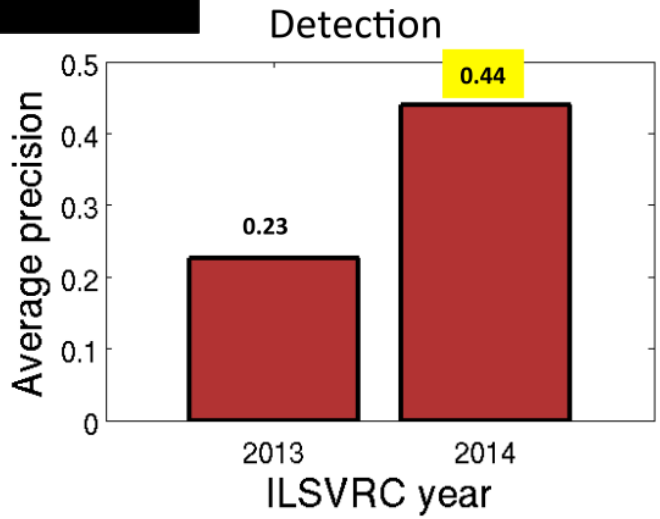
http://image-net.org/challenges/LSVRC/

ILSVRC Object Detection Task

Fully annotated 200 object classes across 121,931 images



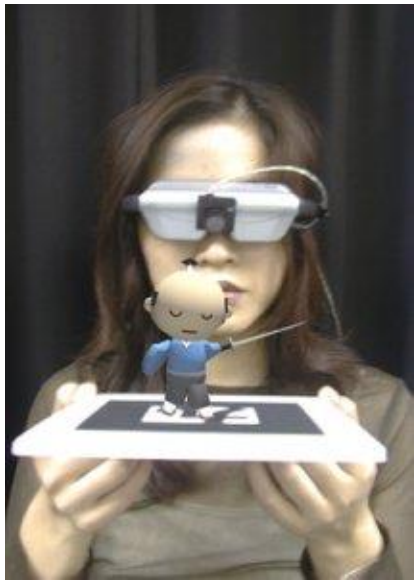
Person
Car
Motorcycle
Helmet



Allows evaluation of generic object detection in cluttered scenes

Augmented Reality (AR)

- A combination of
  - a real scene viewed by a user and
  - a virtual scene/object generated by a computer that augments the scene with additional information.
- Usually require 3D models for the virtual object as well as precise 3D pose estimation of the real scene.



Virtual Reality (VR)

- Inducing targeted behavior in an organism by using artificial sensory stimulation, while the organism has little or no awareness of the interference.

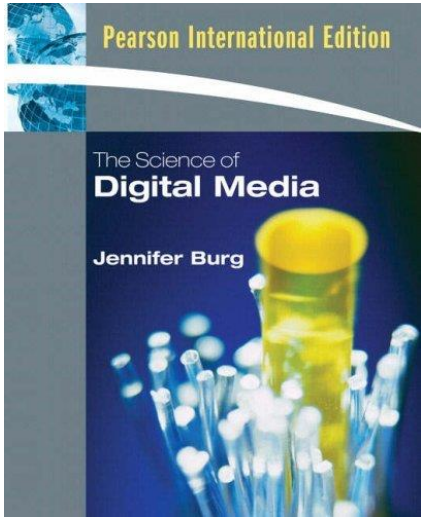
The user, wearing a VR headset, flaps his wings while flying over virtual San Francisco, while a motion platform and fan provide additional sensory stimulation. The figure on the right shows the stimulus presented to each eye.

References

The Science of Digital Media

Jennifer Burg

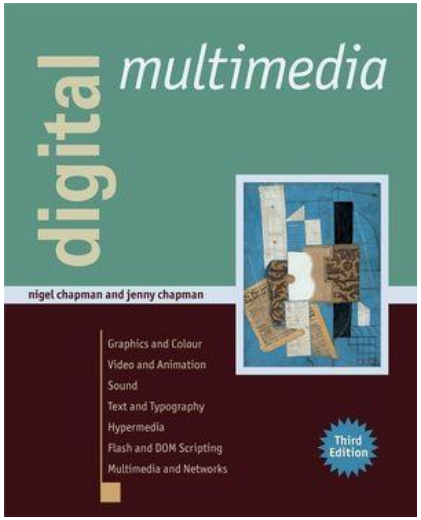
Pearson Prentice Hall, 2010



Digital Multimedia, 3rd Edition

Nigel Chapman and Jenny Chapman

Wiley, 2009.



Prerequisites

- Linear Algebra
- Probability
- Basic programming skills

Grading

Midterm Exam. (April 28)	30%
Homeworks (4)	40%
Final Project	20%
Quizzes	5%
Class Participation	5%

Homework Policy

- Homeworks will involve programming assignments (in Matlab, C, or C++).
- Discussion of homework is encouraged, but you have to write your own. Copying is **strictly** prohibited.
- Homework should be submitted before the announced due time. Scores of late homeworks will be reduced by 20% per day.

Final Project

- Each student is required to do a final project of a topic from a list of suggested topics.
- You can form a team of 2-3 students to do the final project.

Course Webpage

- http://cv.cs.nthu.edu.tw/courses.php
- Important information and course slides will be posted on the NTHU iLMS system.
- Questions and discussions for this course are encouraged to post on the iLMS system.

Class Participation

- Class attendance is required and treated as the basic requirement for class participation.
- Asking questions is strongly encouraged.
- There will be a couple of quizzes in class during the semester.

CS 3570 Classroom Rule

- No eating is permitted.
- No sleeping during the class.
- Disturbance to others in class should be minimized.
- Cell phone should be turned off during the class.