
Digital Multimedia

Contact Information

MENG YANG

2581612839@qq.com

Course Materials

■ Prescribed textbook:

- ❑ 数字媒体基础及应用技术, 姬秀娟、周彦鹏、张晓媛、杨艳竹,清华大学出版社, 2014年3月1日
- ❑ Ze-Nian Li, Mark S. Drew. Fundamentals of Multimedia,影印版: 多媒体技术教程, 机械工业出版社, 2004.7

■ Other references:

- ❑ Ze-Nian Li, Mark S. Drew, Jiangchuan Liu, Fundamentals of Multimedia (Second Edition), Springer, 2014

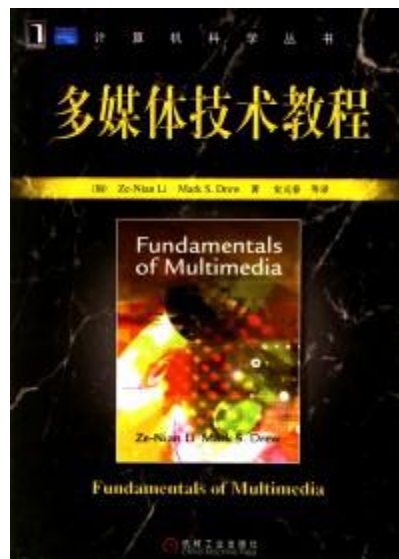
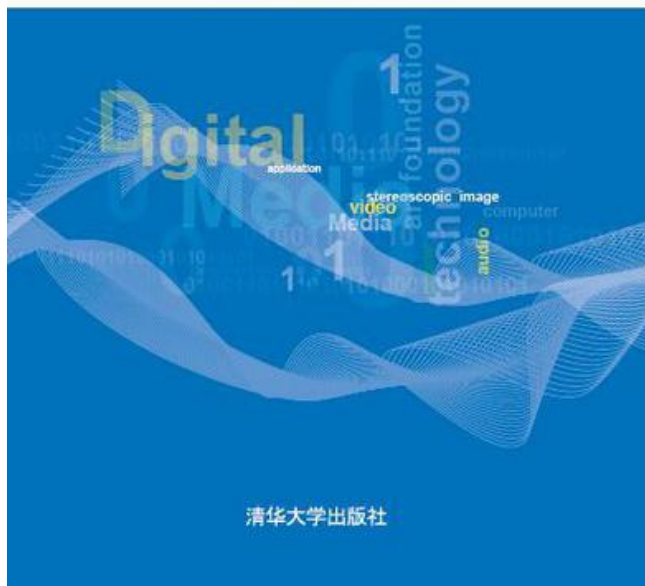
■ Lecture notes will be available online

Digital
Media

21世纪普通高等学校数字媒体技术专业规划教材精选

数字媒体基础 及应用技术

姬秀娟 周彦鹏 张晓媛 杨艳竹 编著



Texts in Computer Science

Ze-Nian Li
Mark S. Drew
Jiangchuan Liu

Fundamentals of Multimedia

Second Edition

 Springer

Course Materials

- We'll cover selected parts of the fundamental concepts using the two books
- Toward the end of the course (last month), students are expected to present topics in more depth as part of the final project.
- Or refer to:

<http://www.cs.ccu.edu.tw/~tsaic/>

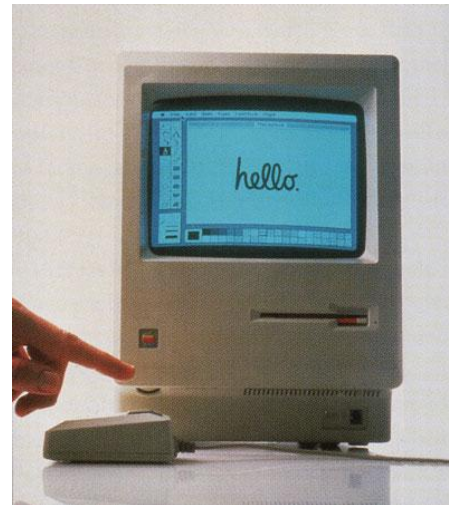
<http://www.cs.bc.edu/~hjiang/>

Course Outline

- 见教学大纲

What is Multimedia?

- “Multimedia” has no strict definition.
- In our context, multimedia indicates the computer technology (multimedia computing) for more efficient communication by using different media types:
 - ❑ Text
 - ❑ Audio and speech
 - ❑ Images
 - ❑ Graphics
 - ❑ Video



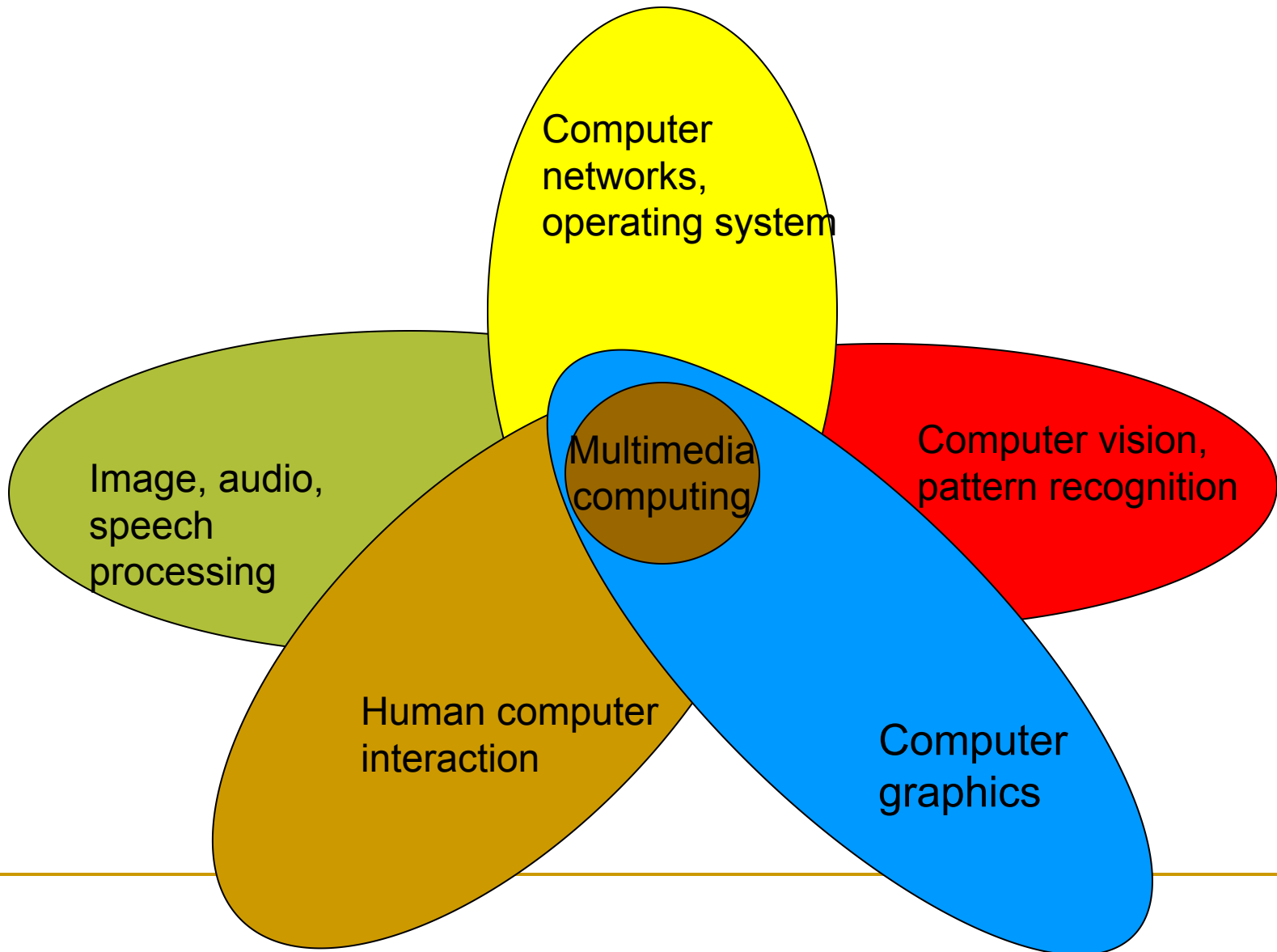
Multimedia System

- Multimedia involves more than simple addition of new data types.
- It integrates a wide range of symbol modes simultaneously into a coherent framework.
- The framework is usually denoted as a multimedia system.

Challenges of Multimedia Computing

- Developing a successful multimedia system is non-trivial.
 - Continuous media types such as video need a lot of space to store and very high bandwidth to transmit.
 - They also have tight timing constraints.
 - Automatically analyzing, indexing and organizing information in audio, image and video is much harder than from text.
 - Multimedia involves many different research areas and needs more complex and more efficient algorithms and hardware platforms.

Multimedia is Multidisciplinary




Example Multimedia Systems

Real Estate

House Flip Book


Page 2 of 6

Next Previous First Last Load...




Cambridge \$1050000

Architectural gem from the height of the American Renaissance. Historic house on academic side street near Harvard. Beautifully proportioned rooms with stunning




Cambridge \$995000

Customized colonial with all-new appointments in exclusive location. The warmth of an older home with the security of newer construction. Prime executive location; perfect



Cambridge \$1025000

Vibrant renovation of impressive Victorian home on national historic register. Exquisite gardens and immaculate grounds. Serious inquiries only.




Cambridge \$990000

You'll never want to leave home again. 3 soaring chimneys and dormers galore in this spacious hideaway.

GTE

Video Conference



Out

Mortgage Calculator

(values are only approximations)

	A	B
0	Home Mortgage Calculations	
1		
2	-----	
3	House Price	120000.00
4	Down Payment	20000.00
5	Interest Rate	10.00
6	Years of loan	30.00
7	-----	
8	Monthly Payments	877.57
9	Total Payments	335925.77
10	Total Interest	215925.77

This prototype real estate application was developed with the AthenaMuse software at GTE Laboratories. House descriptions and color images are retrieved from a multimedia database and placed in a customized listing booklet. The client and Realtor can discuss candidate homes and financing options via desktop videoconferencing and a shared document facility.

Interface:
Russell Sasnett, at
GTE Laboratories.

Example Multimedia Systems

Chronoscope: Musée d'Orsay



The *Chronoscope* application spreads out an artist's works on a timeline. Paintings by different artists of the same period can be studied side-by-side to explore the cross-fertilization of ideas. Based on impressionist paintings from 1848 to 1914, in the collection of the Musée d'Orsay in Paris, France.

Interface:
Matthew Hodges.

Content:
Musée d'Orsay, Paris, France.

Example Multimedia Systems

Navigation Learning Environment



The *Navigation Learning Environment* is a complex simulation designed to teach the basics of coastal navigation with “surrogate travel” techniques. The software can render a view in any direction from the pilot’s perspective, using a database of 360-degree panoramas. Maps and charts help to set a course, while a throttle control determines the rate at which the boat’s position is updated. Interface and content: Matthew Hodges.

Example Multimedia Systems



Photosynth of Microsoft Live Labs.

Multimedia Computing

- Multimedia systems involve some basic enabling techniques:
 - Multimedia data representation and compression.
 - Multimedia data processing and analysis.
 - Transmitting multimedia data through communication networks.
 - Multimedia database, indexing and retrieval.

Data Compression

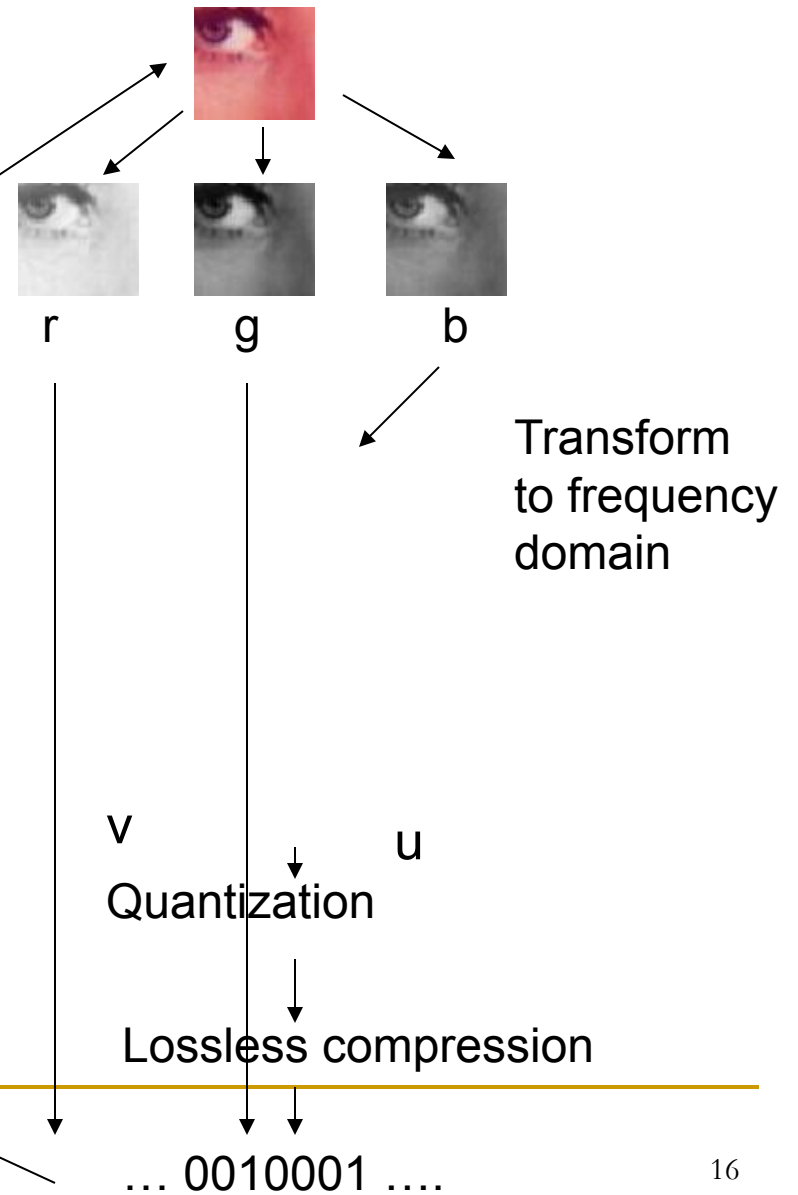
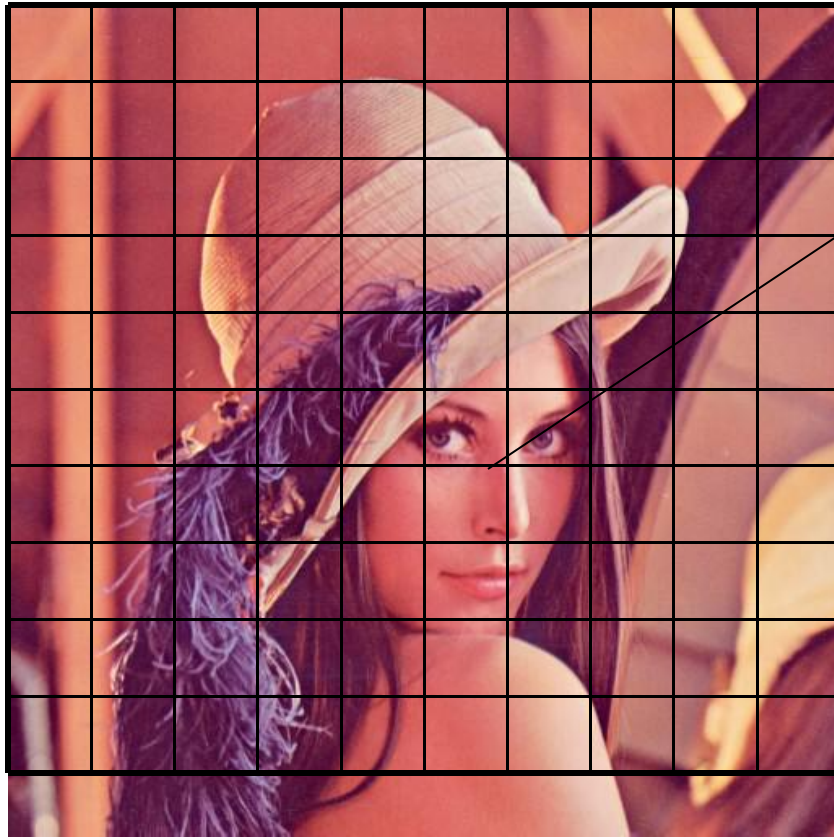


Raw image takes about 6M bytes
(without header information)



24k bytes with jpeg, Q=50

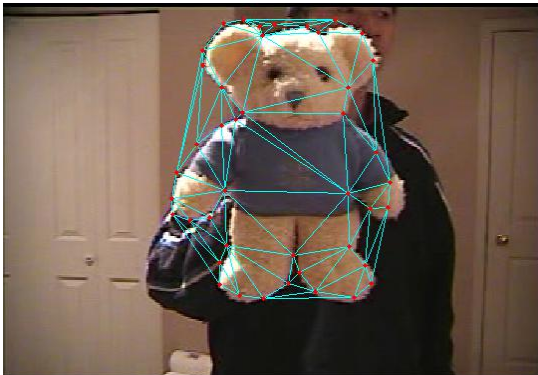
“JPEG” Illustration



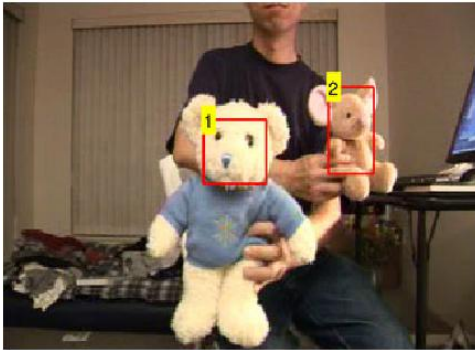
Media Processing and Analysis

- In applications such as digital library, automatic data analysis has to be done to extract semantic meanings from audios, images and videos.
 - Based on media processing methods such as
 - ❑ object tracking (face, eyes),
 - ❑ object recognition,
 - ❑ gesture recognition, etc.,
- we can build more effective human computer interfaces.

Tracking Object by Image Matching



Multiple Object Tracking



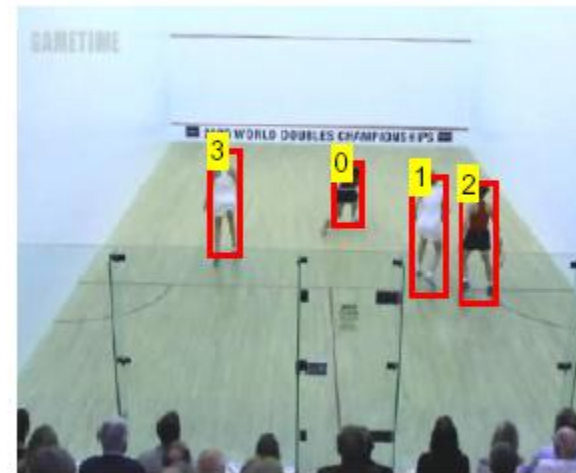
Toys



Squash

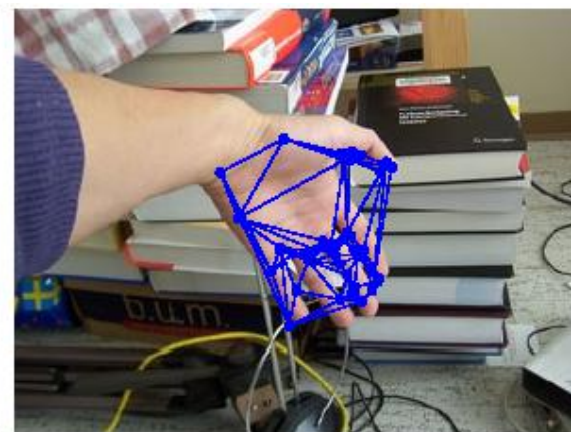
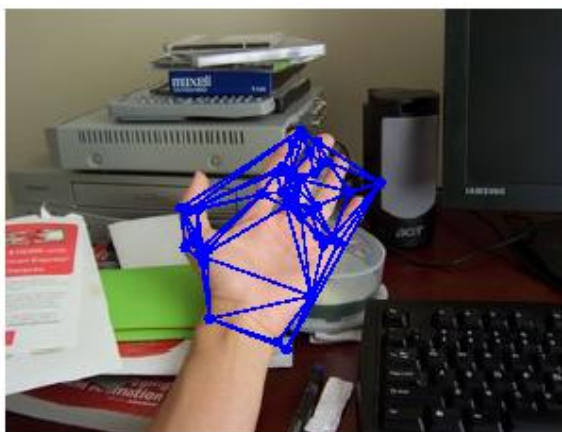
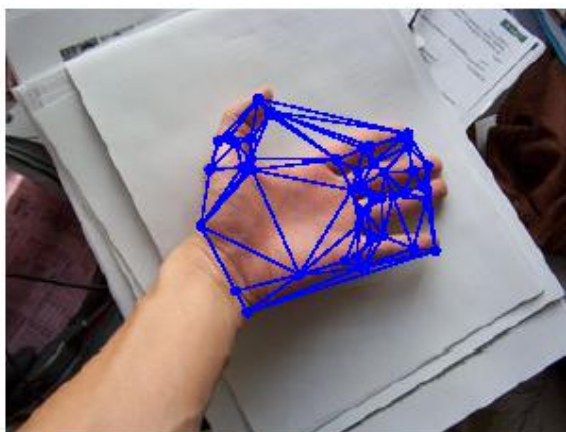
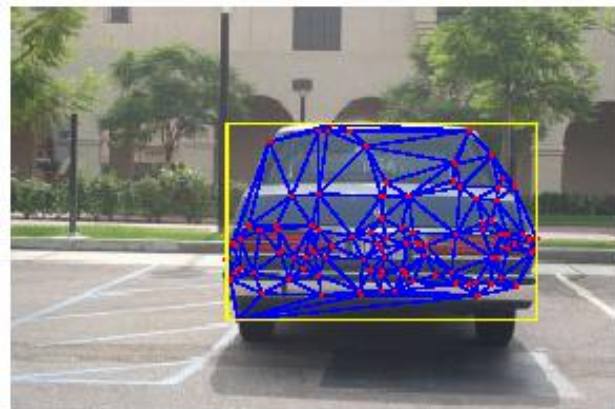


Lab

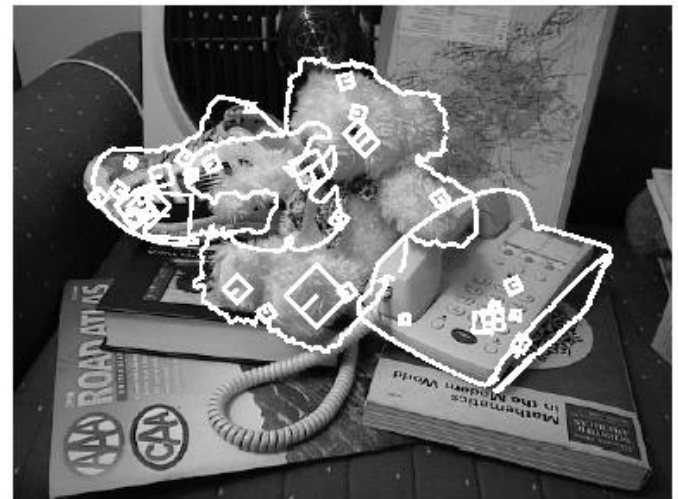


Double Squash

Finding Objects in Clutter



Object Recognition



David G. Lowe, "Object recognition from local scale-invariant features," ICCV99.

Posture Detection in Images



(a) Top 23 matches for figure skating posture 1. The first image is the exemplar.



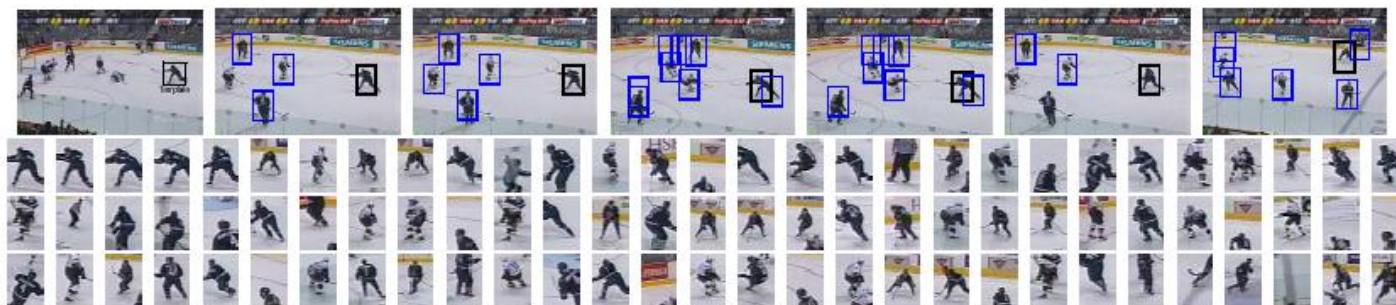
(b) Top 23 matches for figure skating posture 2. The first image is the exemplar.



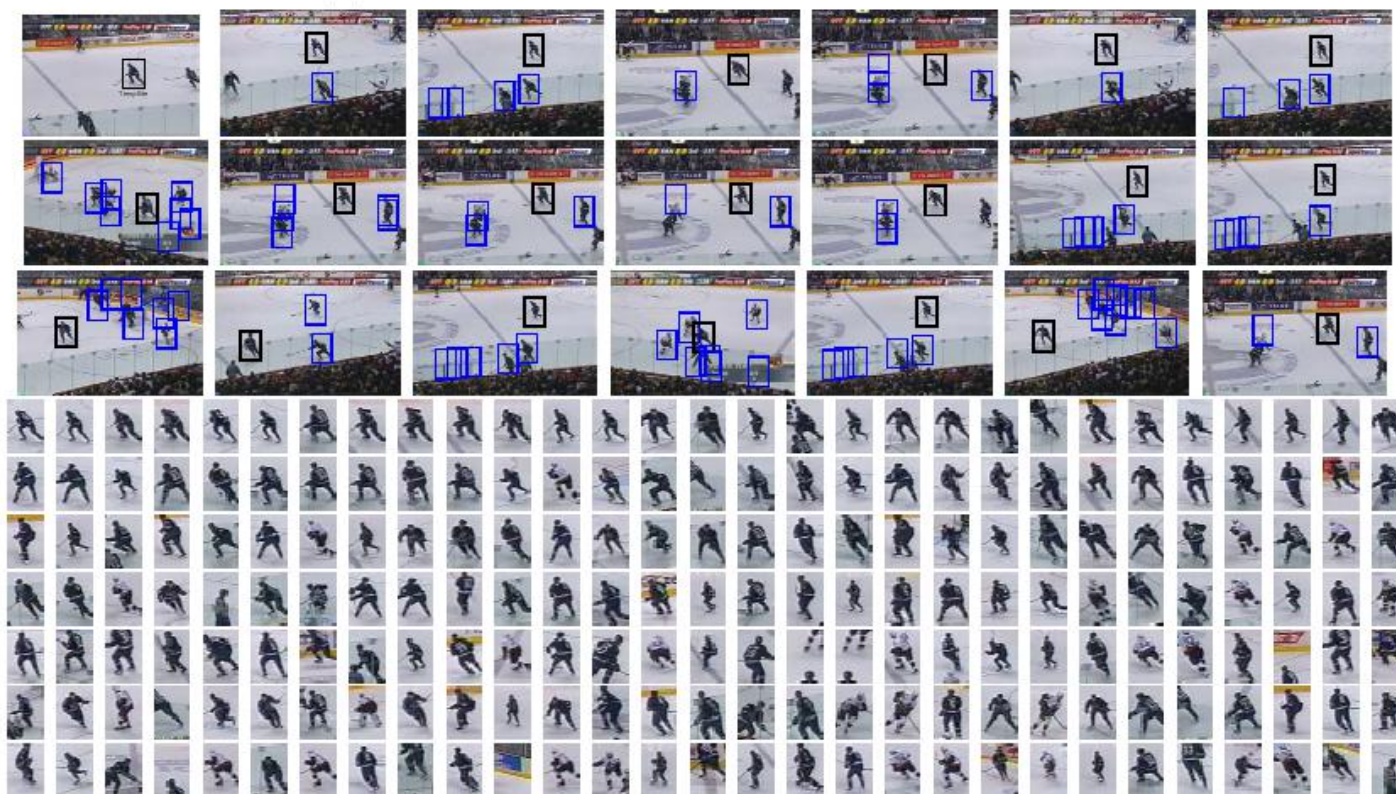
(c) Top 23 matches for figure skating posture 3. The first image is the exemplar.

Template
Image

Dealing with Multiple Objects

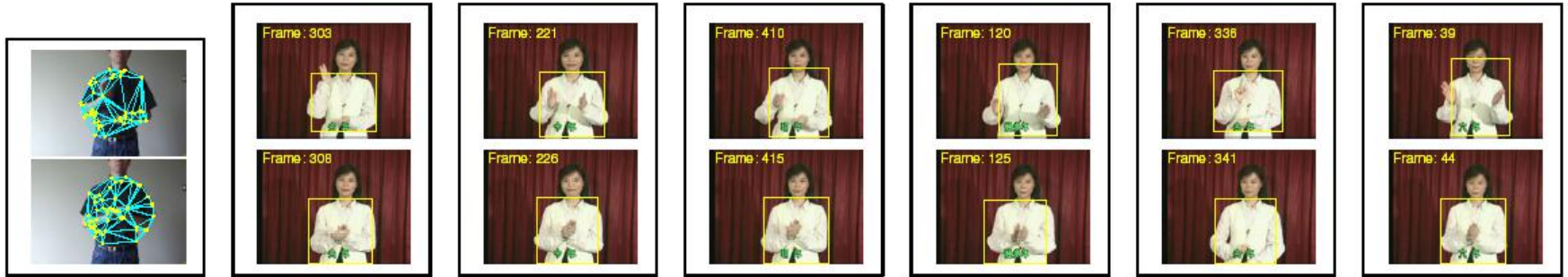


(a) Locating shooting posture in video with exemplar 1



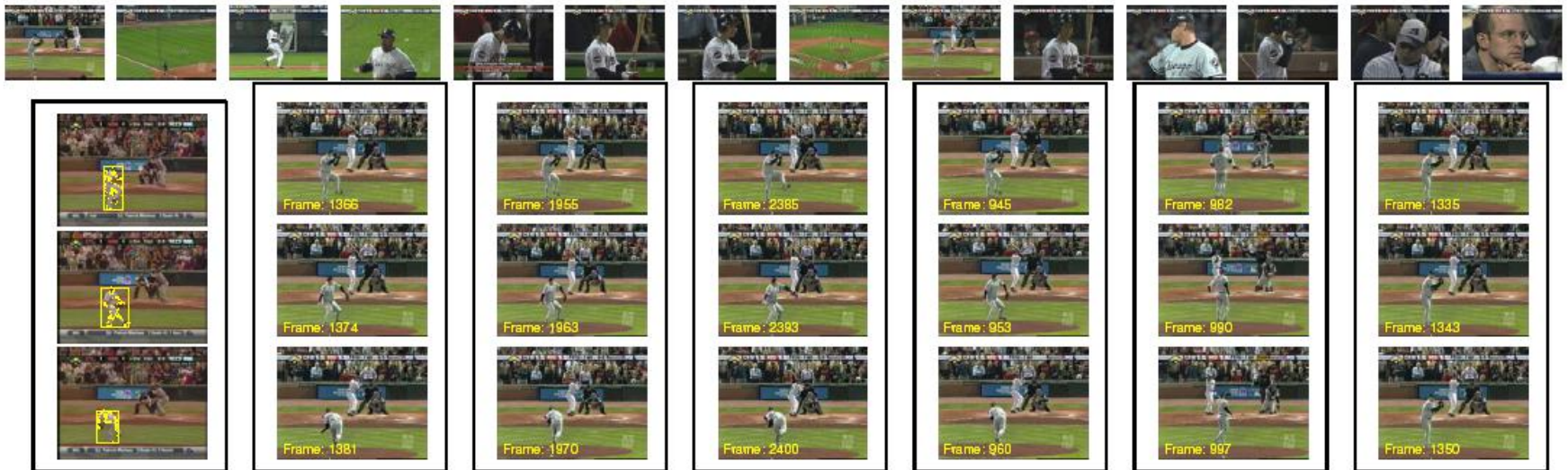
(b) Locating postures in video with exemplar 2

Action Detection



(a) Templates (b) Cost:43.9 (c) Cost:45.5 (d) Cost:46.0 (e) Cost:46.2 (f) Cost:46.4 (g) Cost:46.5

Sign language gesture detection



(a) Templates (b) Cost:29.29 (c) Cost:29.68 (d) Cost:29.91 (e) Cost:30.11 (f) Cost:30.17 (g) Cost:30.19

Detecting actions in baseball sequence

Finding Action Clusters in Image

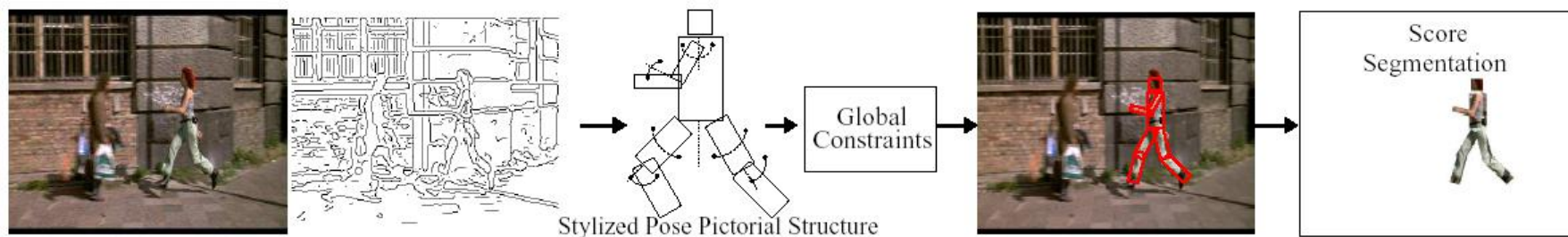
D:



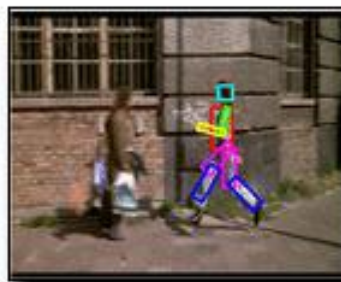
Finding Action Clusters (con't)



Action Recognition



video1



video2

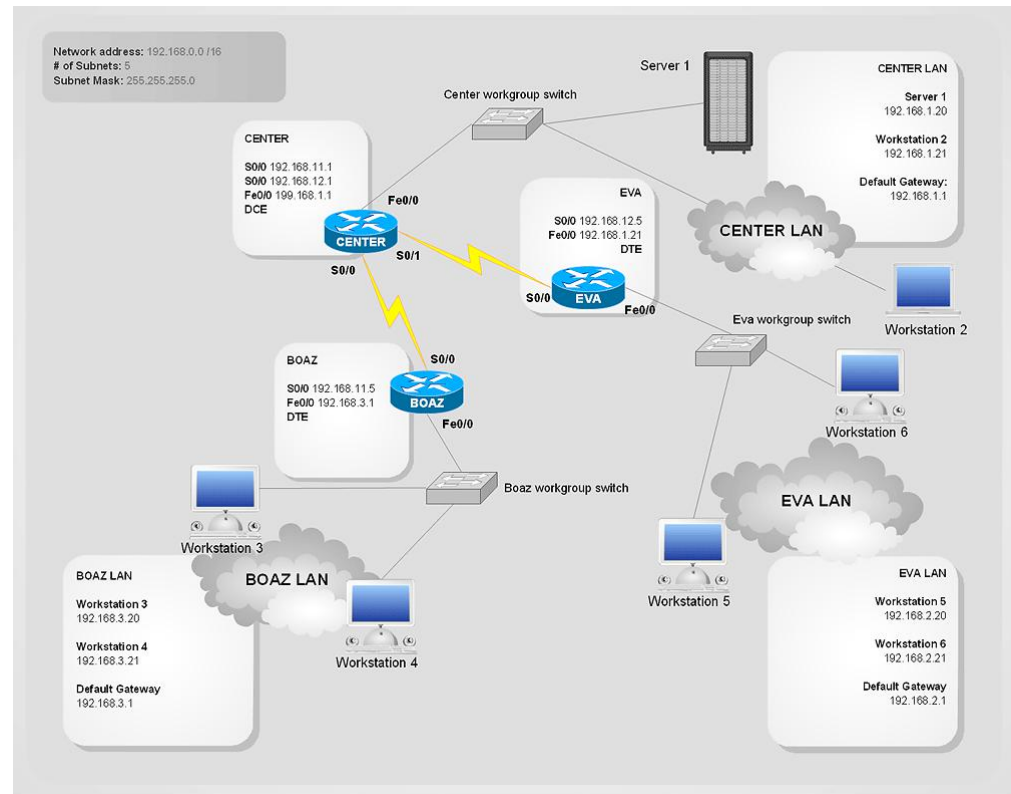


video3

Ramanan, D., Forsyth, D. A., Zisserman, A. "**Strike a Pose: Tracking People by Finding Stylized Poses**"
Computer Vision and Pattern Recognition (CVPR), San Diego, CA, June 2005.

Media Delivery

- Transmitting multimedia data across the network is another topic in multimedia computing.
- There are different issues when we transmit video and audio through packet network.
 - ❑ Quality of service.
 - ❑ Synchronization.
 - ❑ Error and congestion control.
 - ❑ Session setup and book keeping.



Multimedia Database and Indexing

- Multimedia database has to deal with large media files.
- Multimedia data needs new data structures, indexing and searching methods.
- Content based multimedia retrieval is still an ongoing research topic.

Useful Resources

- Journals
 - IEEE Multimedia
 - IEEE Transaction on Multimedia
 - IEEE Transaction on Image Processing
 - IEEE Transaction on Pattern Recognition and Machine Intelligence
- Conferences
 - ACM Multimedia
 - International Conference on Multimedia and Expo (ICME)
 - IEEE Computer Vision and Pattern Recognition (CVPR)
 - International Conference on Computer Vision (ICCV)

Announcement for all

- 期末考试（60%）+平时成绩（40%）
- 平时成绩 包括平时作业 +考勤
- 杜绝抄袭，按时提交
- 认定为抄袭者，此次作业0分
- 不按时提交者，此次作业0分
- 考勤不到者，按照学校规定处理