

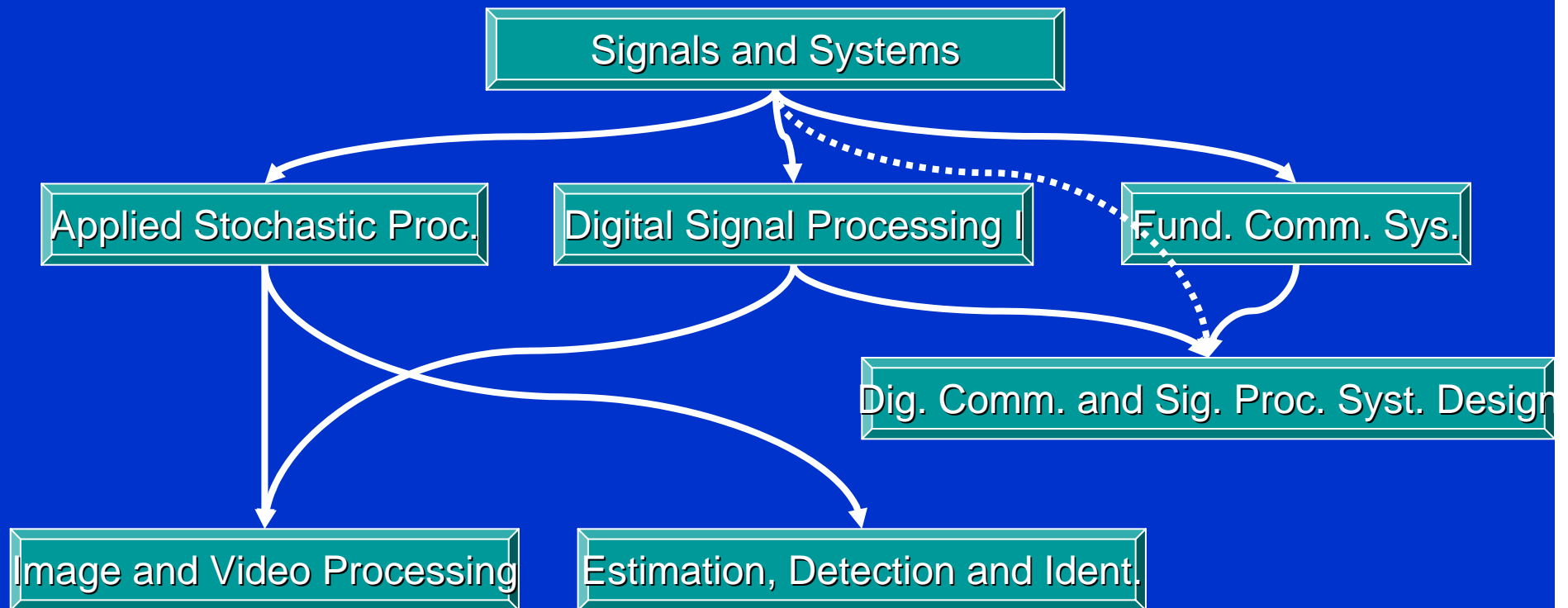
-
-
-
-
-
-
-
-
-
-
-
-

Signals and Systems (18-396),
Image and Video Processing (18-798),
and Life Beyond...

Prof. Tsuhan Chen
tsuhan@cmu.edu



Sample Courses in Signal Processing and Communication



Multimedia Communication

Error Control Coding

Pattern Recognition

Optical Image and Radar Proc.

Digital Signal Processing II

-
-
-
-
-
-
-
-
-
-

Sound



Digital Audio

	Frequency Band (Hz)	Sampling Rate (kHz)	Bits per Sample	Raw Bitrate (kbits/s)
Telephone Speech	300~3400	8	8	64
Wideband Speech	50~7000	16	8	128
Mediumband Audio	10~11000	24	16	384
Wideband Audio	10~22000	48	16	768

– CD: $44.1 \text{ kHz} \times 16 \text{ bits} \times 2 \text{ channels} = 1.411 \text{ Mbits/s}$

MPEG-1 Audio

- ISO/IEC 11172-3 (1988~1991)
 - First high quality audio compression standard
 - Sampling rates: 32, 44.1, 48 kHz
 - CD quality two-channel audio at ~256 kbits/s
 - CD: $44.1 \text{ kHz} \times 16 \text{ bits} \times 2 = 1.411 \text{ Mbits/s}$
 - YES, this is **MP3!!!**
- Quality demonstration
 - Stereo 44.1 kHz at 64 kbits/s
 - Stereo 44.1 kHz at 128 kbits/s
 - Stereo 44.1 kHz at 192 kbits/s
 - Stereo 44.1 kHz at 256 kbits/s

-
-
-
-
-
-
-
-
-
-
-

Image



Image

RGB Color

m n

R = 255 G = 200 B = 200	R = 230 G = 120 B = 234	R = 251 G = 200 B = 190	R = 150 G = 170 B = 253
R = 254 G = 133 B = 200	R = 253 G = 100 B = 120	R = 244 G = 222 B = 214	R = 248 G = 220 B = 242
R = 151 G = 140 B = 139	R = 204 G = 203 B = 202	R = 0 G = 0 B = 0	R = 149 G = 244 B = 130
R = 190 G = 170 B = 90	R = 151 G = 148 B = 149	R = 159 G = 149 B = 150	R = 124 G = 110 B = 123

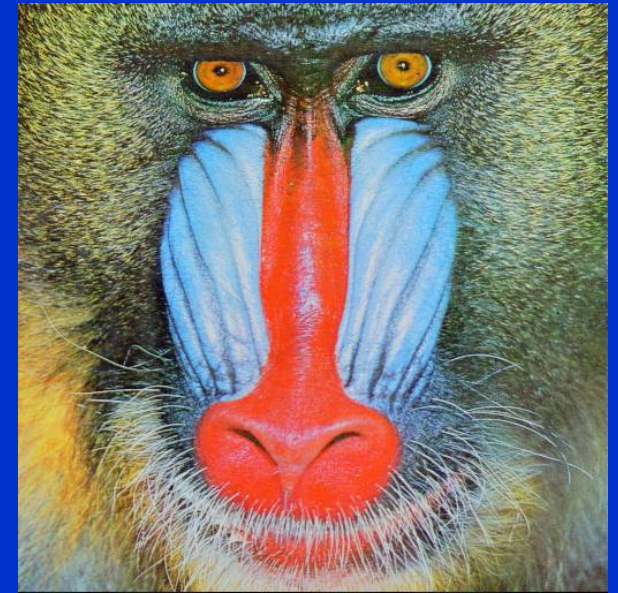
Sample Images



Lena



Pepper



Baboon

$512 \times 512 \times 3 \text{ bytes} = 768\text{KB}$
With JPEG, $\sim 32\text{KB}$

Sampling

Spatial Subsampling

Aliasing!!!

Original (256×256)

(64×64)

(32×32)



MSE = 2058

MAE = 24

CR = 16:1

MSE = 3924

MAE = 36

CR = 64:1

Sampling

Spatial Subsampling w/Averaging

Original (256×256)



(64×64)



(32×32)



MSE = 1010

MAE = 18

CR=16:1

MSE = 1643

MAE = 26

CR=64:1

Quantization

Original (24bit)



(12-bit)



MSE = 9670
MAE = 78
CR = 2:1

(6-bit)



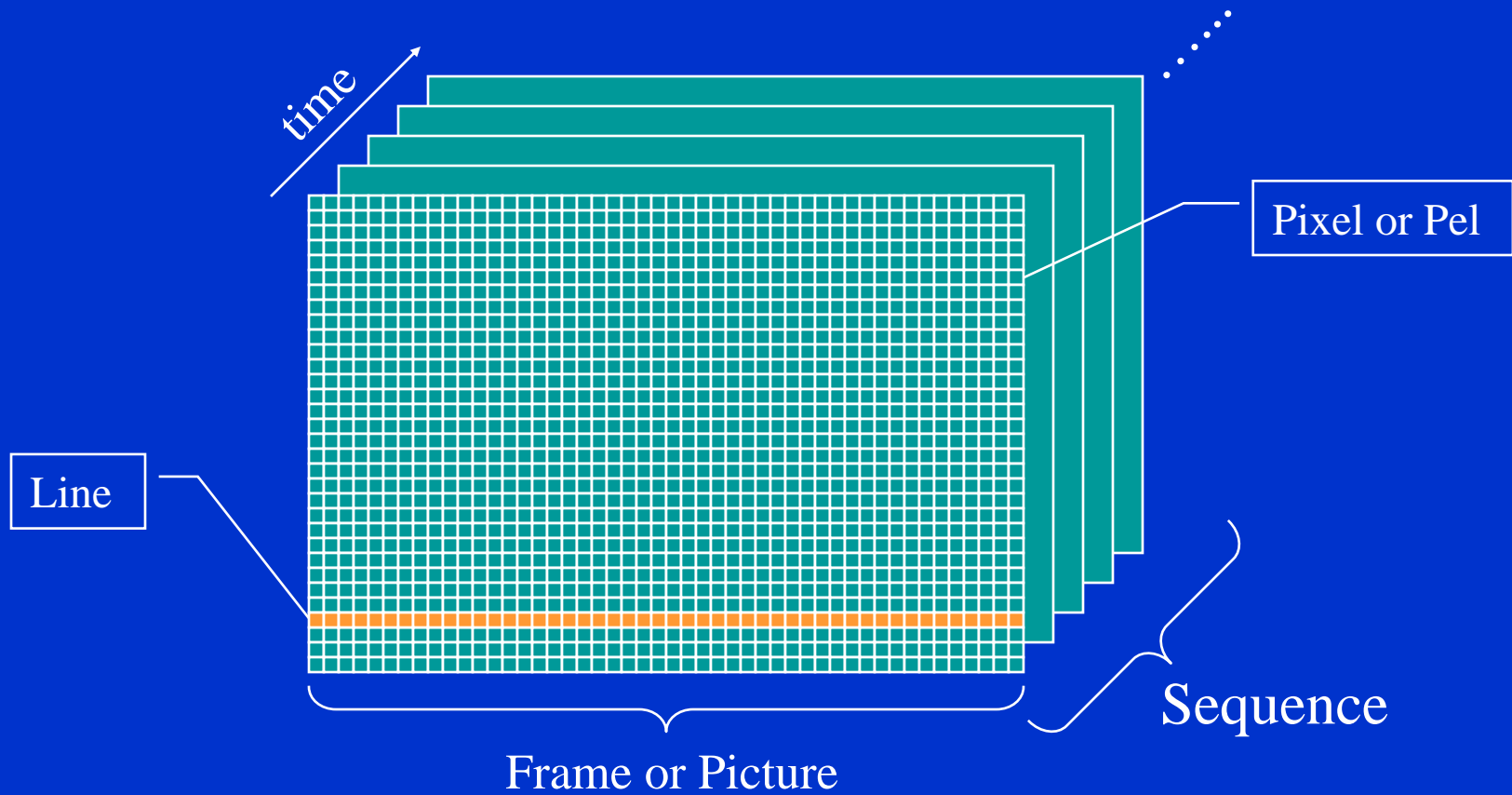
MSE = 10381
MAE = 82
CR = 4:1

-
-
-
-
-
-
-
-
-
-
-
-

Video



Video



Video Data

- Video

	Pels/line	Lines	Frames/s	Bytes/pel	Bit rate
Video Telephony (CIF)	352	288	10	1.5	12.2 Mbits/s
Broadcast TV (ITU-R 601 4:2:2)	720	480	30	2	166 Mbits/s
HDTV	~1280	~720	60	2	885 Mbits/s

- So, we need MPEG-1 (VCD etc.), MPEG-2 (DVD etc.), MPEG-4 (some camcorders, etc.)

-
-
-
-
-
-
-
-
-
-
-
-

Computer Graphics



Face Animation

- Wire-frame mesh model with texture mapping



Demo



Computer Vision



-
-
-

Face Tracking

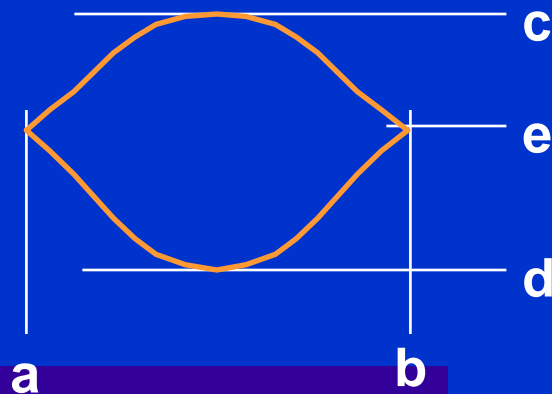
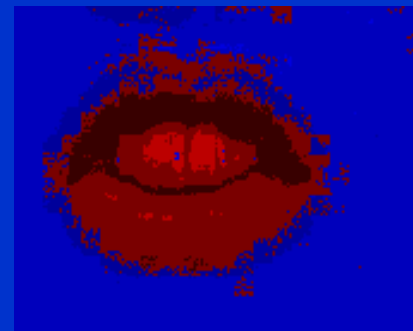
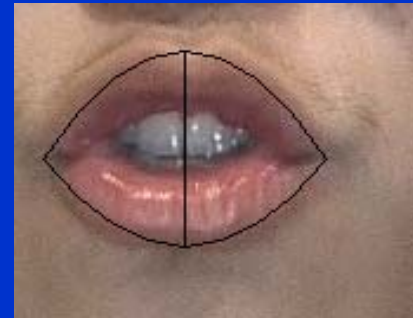
Use color information to segment target vs. non-target pixels

Use deformable template to track the target



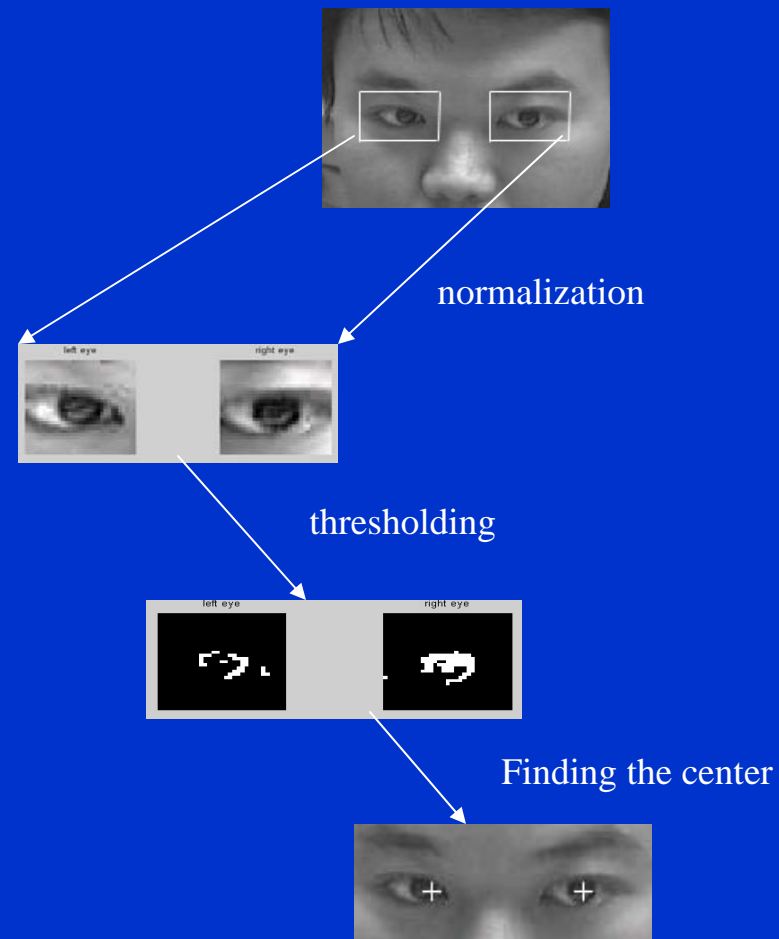
Lip Tracking

- Use a Gaussian mixture with three Gaussians to model the color distribution of the mouth
- Template: two parabolas defined by $\lambda = (a, b, c, d, e)$



Eye Tracking

- Find the center of the darkest region in the search window



Tracking in a Car...

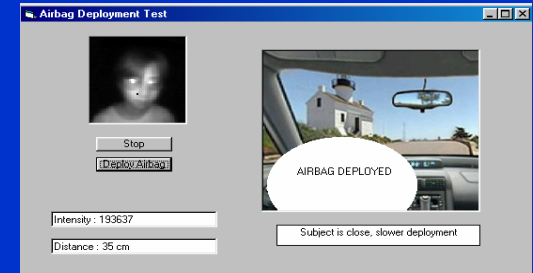
Face/Eye/Hand Tracking



Driver Verification:
Security and User Preference



Gesture-Controlled
Map Browsing



Airbag Deployment Control

-
-
-
-
-
-
-
-
-
-
-

Higher Dimensions?



-
-
-

7D Plenoptic Function

$$f(V_x, V_y, V_z, \theta, \psi, \lambda, t)$$

[Adelson'91]

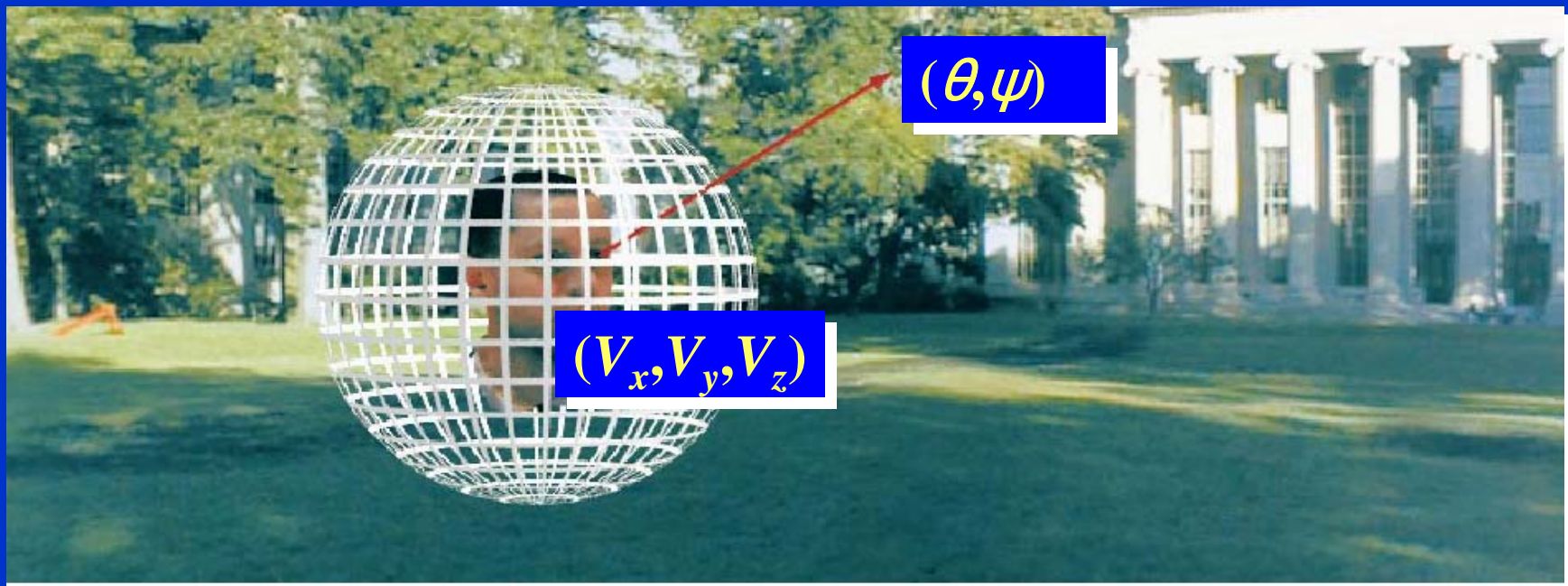
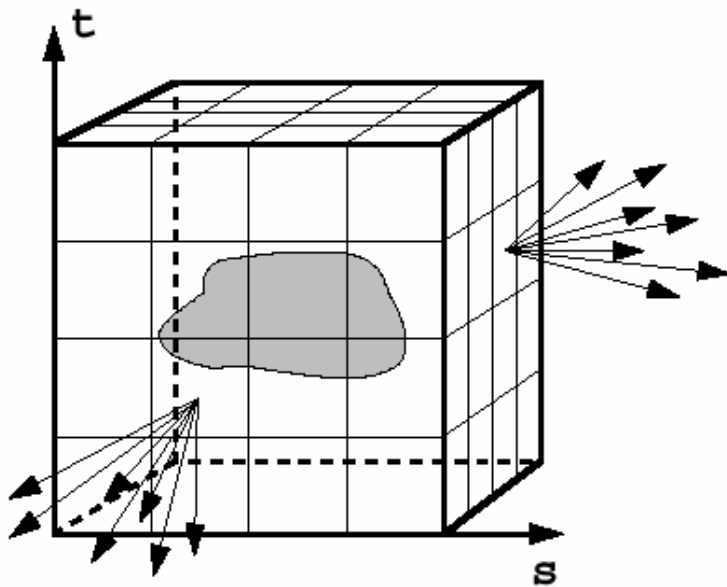


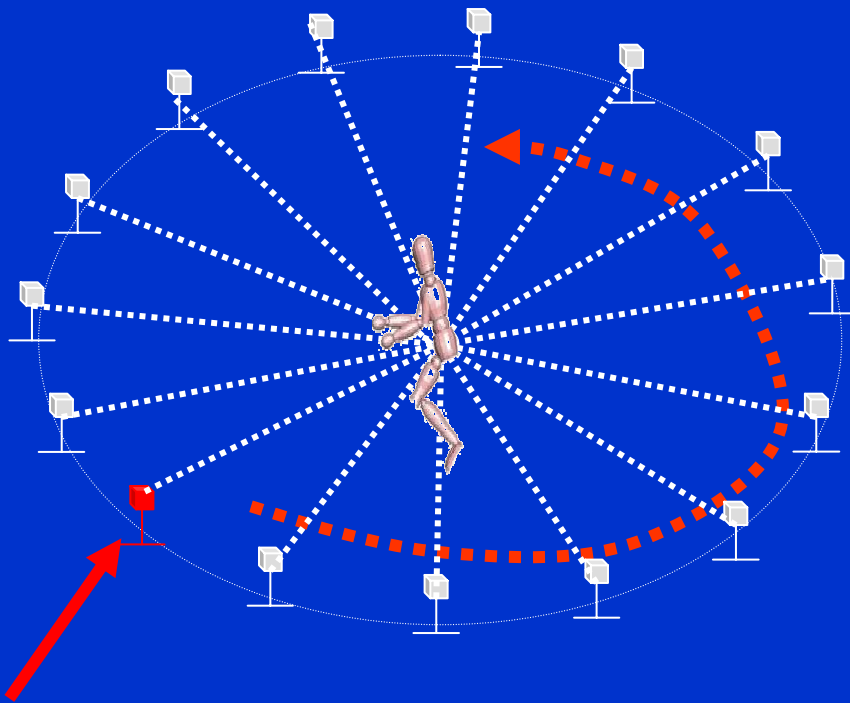
Image-Based Rendering



- Plenoptic Function [Adelson'91] [McMillan'95]
- Lumigraph/Lightfield [Gortler/Grzeszczuk'96] [Levoy'96]
- Concentric Mosaics [Shum99]



“The Matrix”



EyeVision



→ 4D IBR
(incl. time)

[Kanade'01]

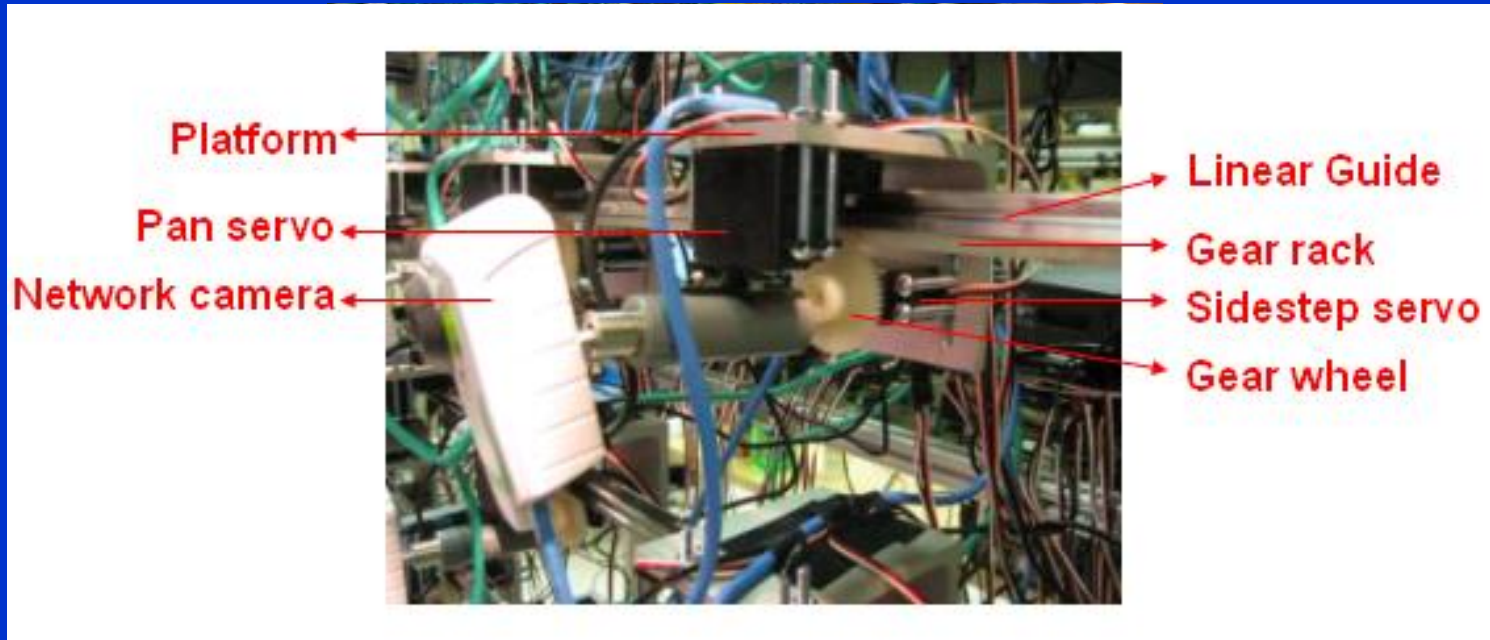


Before Correction



After Correction

Self-Reconfigurable Camera Array

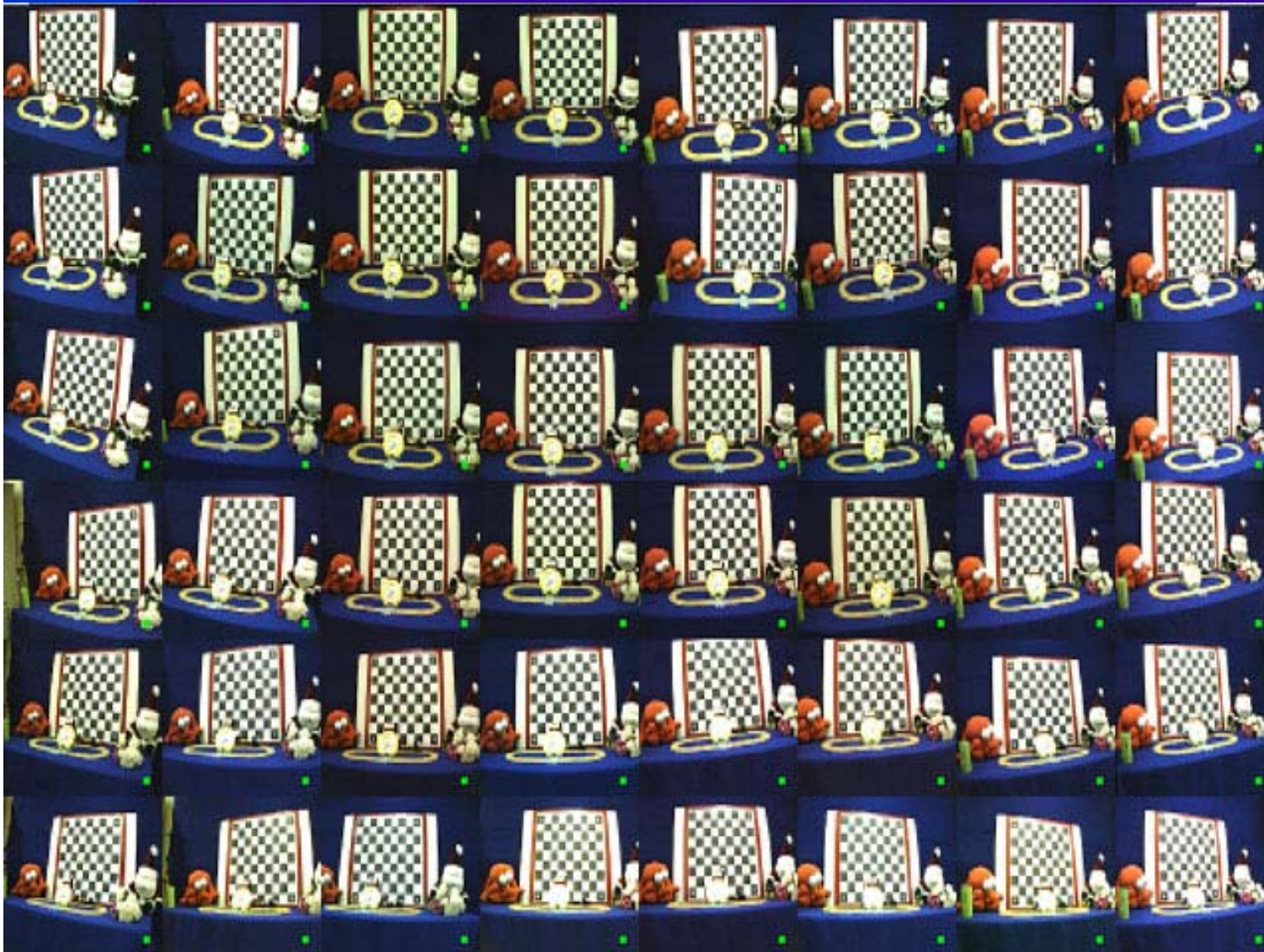


[Levoy, Stanford]

[Zhang and Chen, CMU]

[McMillan, MIT]





Results

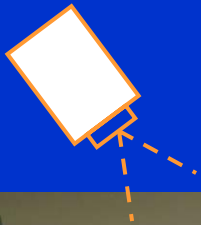
- **Real-time** capturing/calibration/rendering
 - 48 webcams → sensor network
 - 2 step-motors each (translation and pan)



- Building the next version...
 - More mobile and wireless

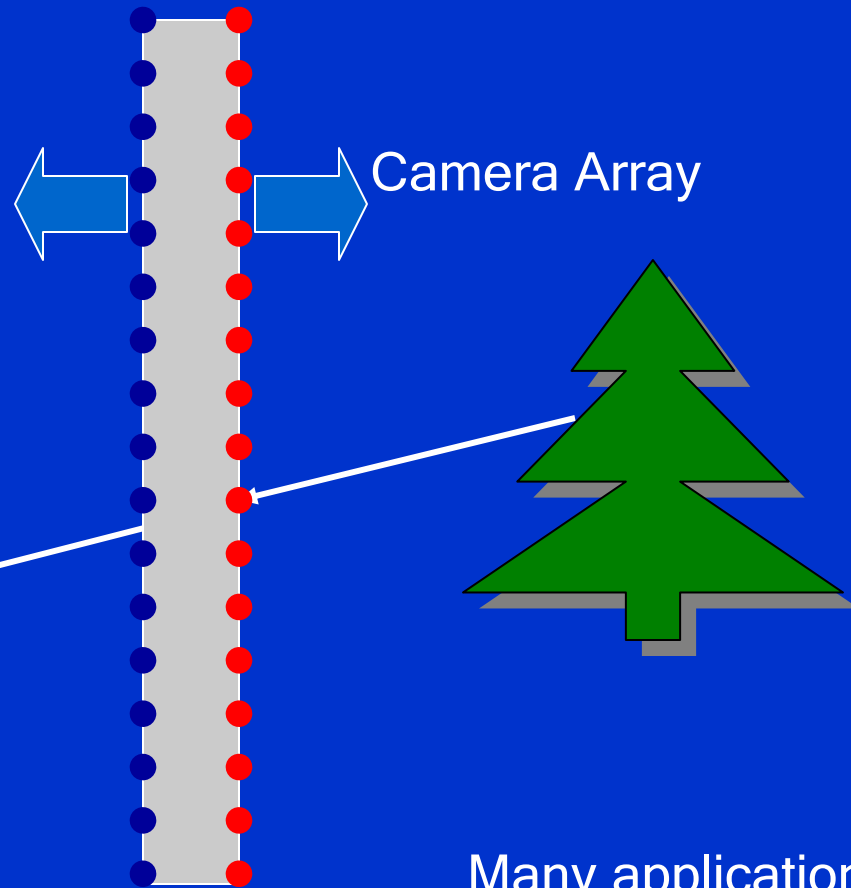
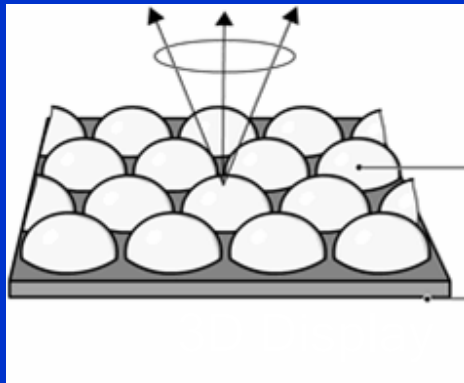
-
-
-

Ongoing: Mirror/Lens Array



This is lightfield/lumigraph!

Future: “Transparent Material”



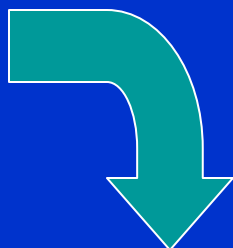
Many applications...

-
-
-
-
-
-
-
-
-
-
-

Information Retrieval (Pattern Recognition)

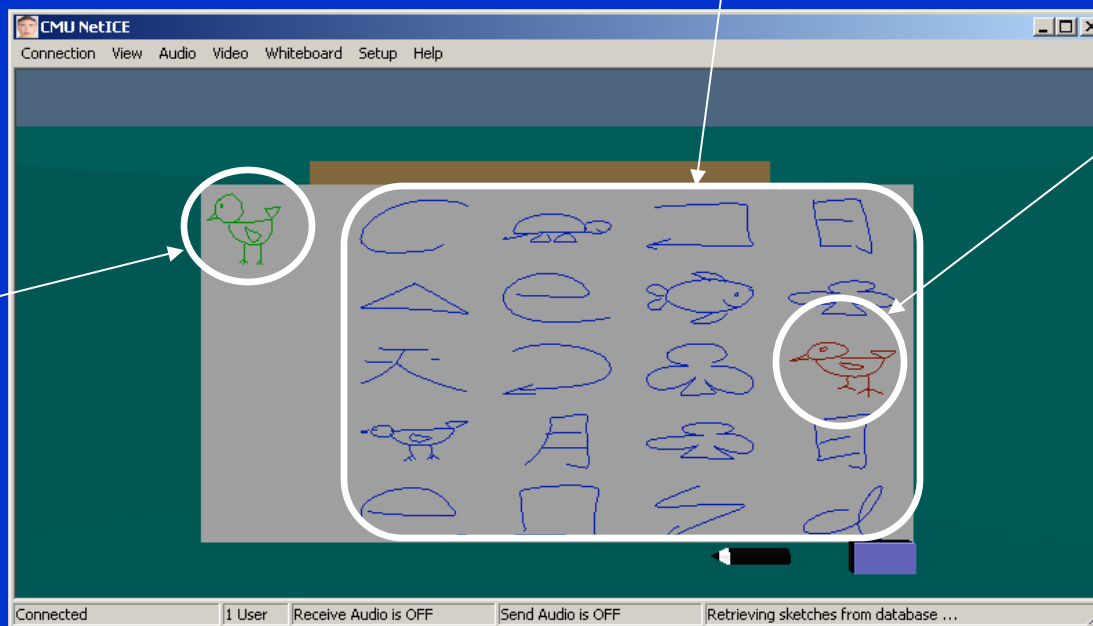


Hand-Drawn Sketch Retrieval



Page stored in
Database

Query
Sketch



Similar
Sketch

Trademark Retrieval




TrademarkRetrieval - D:\demo\TrademarkRetrieval\TrademarkDatabase2.mdf

Database View Adjust Experiment Options Help

Retrieved Trademarks

Page 1

Prev. Next Rand Feedback

049866 [Score=1.0000]  NR ————— R	049484 [Score=0.8209]  NR ————— R	049879 [Score=0.8098]  NR ————— R	044021 [Score=0.3741]  NR ————— R
037318 [Score=0.3669]  NR ————— R	049901 [Score=0.3646]  NR ————— R	044135 [Score=0.3636]  NR ————— R	044148 [Score=0.3616]  NR ————— R
049891 [Score=0.3546]  NR ————— R	040700 [Score=0.3534]  NR ————— R	045808 [Score=0.3515]  NR ————— R	048895 [Score=0.3499]  NR ————— R

Query

Carnegie Mellon University
Advanced Multimedia Processing Lab

Trademark Inside Database
Load Sketch
Sketching My Own Query

Congratulations! Your query results are shown on the right hand side! You are using 'Query by internal trademark'.

To change the query trademark, left click another thumbnail on the right hand side.

Quick reference:

- Click button "Rand" to randomly browse the database, click "Prev." and "Next" to turn pages
- Right click the thumbnail pictures to view the extracted sketch
- Left click reserved for picking up trademarks as queries when query by trademark from inside the database
- Move the sliders toward R (relevant) or NR (nonrelevant) to give the feedback, press the "Feedback" button to do relevant feedback

Trademark Retrieval

TrademarkRetrieval - D:\demo\TrademarkRetrieval\TrademarkDatabase2.mdf

Database View Adjust Experiment Options Help

Carnegie Mellon University
Advanced Multimedia Processing Lab

Hand-Drawn Query

Save Sketch

Query the database by:

Database













The results are shown on the right hand side! You are under 'Query by user sketch' mode.

To sketch another query, click the 'Sketching My Own Query' again.

BTW, your sketch is nice! :)

Retrieved Trademarks

Page 1

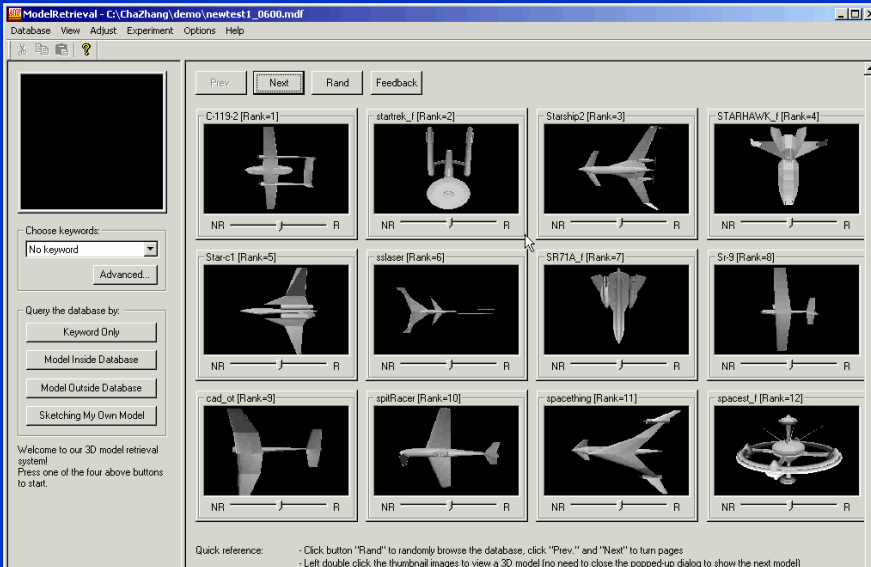
040239 [Score=1.0000]  NR ————— R	040506 [Score=0.9460]  NR ————— R	040352 [Score=0.9018]  NR ————— R	040195 [Score=0.8938]  NR ————— R
040441 [Score=0.8596]  NR ————— R	040194 [Score=0.8379]  NR ————— R	039824 [Score=0.8296]  NR ————— R	039911 [Score=0.6751]  NR ————— R
042457 [Score=0.4333]  NR ————— R	049583 [Score=0.4258]  NR ————— R	040299 [Score=0.4234]  NR ————— R	040824 [Score=0.4147]  NR ————— R

Quick reference:

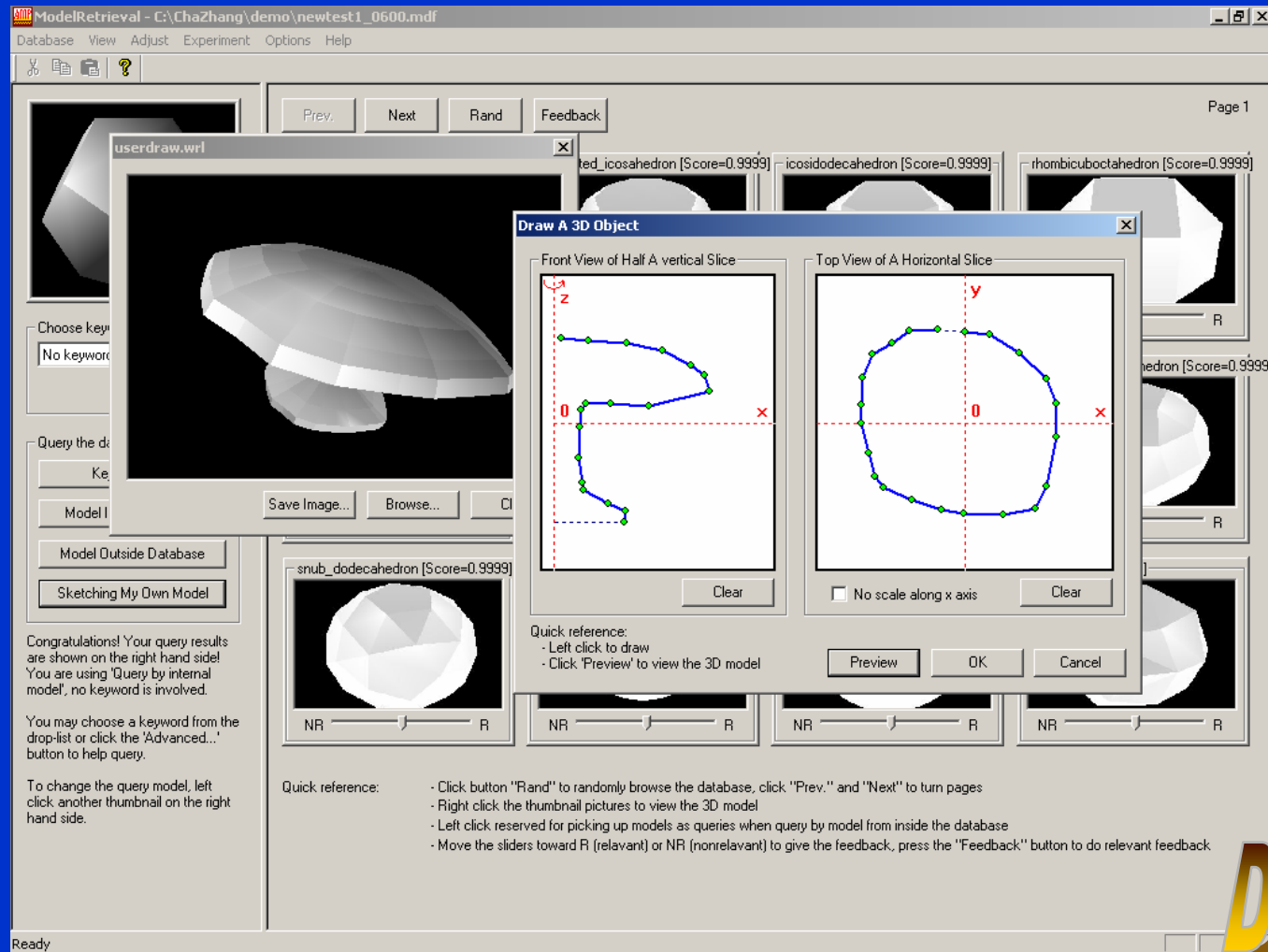
- Click button "Rand" to randomly browse the database, click "Prev." and "Next" to turn pages
- Right click the thumbnail pictures to view the extracted sketch
- Left click reserved for picking up trademarks as queries when query by trademark from inside the database
- Move the sliders toward R (relevant) or NR (nonrelevant) to give the feedback, press the "Feedback" button to do relevant feedback

Ready

3D Object Retrieval



Sketched 3D Query too...



Demo

3D Protein Structures too...

The screenshot displays a web application for protein structure search. The main window, titled "ModelRetrieval - C:\msk000\myprotein200.nsf", features a menu bar with "Database", "View", "Generate DB", "Experiment", "Options", and "Help". Below the menu are navigation buttons: "Prev.", "Next", "Reset", and "Feedback".

The interface is divided into several sections:


- Left Panel:** Contains a large 3D protein structure visualization. Below it are buttons for "Browse...", "Model Info...", and "Search database".
- Classification:** A text block providing details for the selected model: "Classification: --OXYGEN TRANSPORT--", "Database Code: 105M", "Database Date: 19-DEC-97", and "Number of Atoms: 1386".
- Grid of Results:** A 3x3 grid of smaller 3D protein structure visualizations. Each visualization is labeled with a PDB ID and a score: "pdb105m [Score=1.0000]", "pdb111m [Score=0.9958]", "pdb112m [Score=0.9846]", "pdb106m [Score=0.9789]", "pdb101m [Score=0.9744]", and "pdb108m [Score=0.9578]". Each visualization includes a "NR" label and a horizontal slider between "NR" and "R".
- Detail Window:** A pop-up window titled "c:\msk200\pdb111m.ent" is open over the "pdb111m" result. It shows a larger 3D protein structure and its classification: "Classification: --OXYGEN TRANSPORT--", "Database Code: 111M", "Database Date: 24-DEC-97", and "Number of Atoms: 1433". Below the structure are buttons for "Save Image..." and "Browse...".

Demo

-
-
-

Summary

- Signals and Systems
- Image and Video Processing
- Computer Vision
- Computer Graphics
- Pattern Recognition
- Information Retrieval



Signals and Systems (18-396),
Image and Video Processing (18-798),
and Life Beyond...



Prof. Tsuhan Chen
tsuhan@cmu.edu