CAP5415 Computer Vision

- Instructor: Dr. Mubarak Shah, shah@cs.ucf.edu, office:238 CSB, http://www.cs.ucf.edu/class/cap6411
- Office Hours:
  - 2PM to 3PM Mon, 4PM-5PM Tu, 3PM-4PM Thurs
- Grading
  - Mid term 20%, Final 30%, homework 10%, programs 30%, term paper 10%
- Class notes
  - Fundamental of Computer Vision, Mubarak Shah, available on the webpage
- Text Book
  - Introductory Techniques for 3D computer vision, E. Trucco and A. Verri, Prentice Hall
- Other suggested Books

Computer Vision

- Image Analysis
- Image Understanding
- Video Analysis
- Video Understanding
Image Formation

- Light Source
- Camera
- Surface reflectance
- Surface shape

Perspective Projection

\[
\begin{align*}
-y &= \frac{f}{Z} \\
y &= -\frac{fY}{Z} \\
x &= -\frac{fX}{Z}
\end{align*}
\]
**Orthographic Projection**

Image Plane

(X,Y,Z)

World point

y = Y
x = X

**Image**

- 2-D array of numbers (intensity values, gray levels)
- Gray levels 0 (black) to 255 (white)
- Color image is 3 2-D arrays of numbers
  - Red
  - Green
  - Blue
- Resolution (number of rows and columns)
  - 128X128
  - 256X256
  - 512X512
  - 640X480
Video

• Sequence of frames
• 30 frames per second

Digitization

• TV camera is analog, need
  – A to D converter
  – Frame grabber
• Digital Cameras do not need digitization
  – JVC (MPEG through fire wire)
  – Sony (MPEG through fire wire)
Image Formats

- TIF
- PGM
- PBM
- GIF
- JPEG
- MPEG
- Quick Time

Digital TV

- Networks started broadcasting limited DTV programs in Nov 98.
- All commercial stations are supposed to switch to DTV by 2002
- All stations are supposed to switch to DTV by 2003
- Govt wants broadcasters’ NTSC channels returned by 2006 for auctioning!
Digital TV

• CBS carried few NFL games last year
• CBS and ABC plans
  – evening news
  – movies
  – rest of the day up-convert standard TV
• NBC
  – no broadcast yet
  – plans for “Tonight Show” this fall!

Digital TV

• CBS and NBC use 1080i (1920x1080), which is 995Mb/s at 30 fps
• ABC and Fox use 720p (1280x720), which is 424Mb/s at 30 fps
• 6 MHz channel assigned to each network can carry 19.4Mb/s
• Need 50:1 compression ratio!
Computer Vision

• Shape from X (Recover 3-D shape from 2-D image(s))
  – Stereo
  – Motion
  – Shading
  – Texture
  – Contours
Stereo

http://www.vision3d.com/stereo.html
Renault Stereo Pair

Depth Map
Stereo Pair

Candy
Dinosaur

Shark
Shape from Shading

Lambertian Model

\[ f(x, y) = n.L = (n_x, n_y, n_z).(l_x, l_y, l_z) \]

\[ f(x, y) = n.L = \frac{1}{\sqrt{p^2 + q^2 + 1}}(-p, -q, 1).(l_x, l_y, l_z) \]
Sphere

\[ z = \sqrt{\left( R^2 - x^2 - y^2 \right)} \]

\[ p = \frac{\partial z}{\partial x} = -\frac{x}{z} \]

\[ q = \frac{\partial z}{\partial y} = -\frac{y}{z} \]

\[ (n_x, n_y, n_z) = \frac{1}{R} (x, y, z) \]
Vase

Visual Motion
Image from Hamburg Taxi seq

optical flow
Video Mosaic

Video Mosaic
Video Mosaic

Sprite
JPEG

Original 64K 13K 5K Difference

Model-Based Image Coding
Synthesizing Realistic Facial Expressions
Compression

Figure 14: Left to Right: Mesh with uncompressed textures, compressed to 400 kbps/sec, and compressed to 200 kbps/sec.

FACIAL EXPRESSIONS

RAISE EYE BROWS   SMILE
FACIAL EXPRESSIONS

DISGUST

ANGER

Lipreading
Human Behavior Recognition

Key Frames Sequence 1 (350 frames), Part 1
Detecting Driver Alertness
Eye Tracking

Tamara Miller

Results

..\..\d drive\STUDENTS\TAMARAM\2latest.html

..\..\d drive\STUDENTS\TAMARAM\3latest.html

..\..\d drive\STUDENTS\TAMARAM\8latest.html

..\..\d drive\STUDENTS\TAMARAM\9latest.html
Determining 3D Face Orientation

Alper Yilmaz
Discriminating human and animal motion

Nan Li
Discriminating human and animal motion

A method is presented to:

- Remove commercials from interview videos
- Segment interviews into host and guest shots

A clip of Larry King interview
A Short Connectivity Graph

Start

One story

Another story

Commercials

End

Shots detected as ‘Host’

Shots detected as ‘Guest’