CAP 6411 Computer Vision Systems

- Instructor: Dr. Mubarak Shah, shah@cs.ucf.edu, 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 25%, Programs 45% , Homework 10%
- Recommended Book, but not required.
  - Digital Video Processing, A. M. Tekalp, Prentice Hall.

Lecture-1

Multimedia

- Text
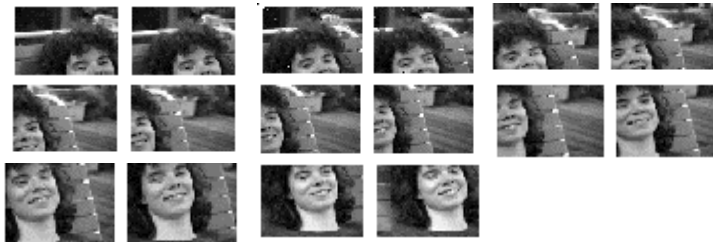- Graphics
- Audio
- Images
- Video

Imaging Configurations

- Stationary camera stationary objects
- Stationary camera moving objects
- Moving camera stationary objects
- Moving camera moving objects

Video

- sequence of images
- clip
- mosaic
- key frames

Sequence of Images

Clip

Mosaic

Key Frames

Clip

Mosaic

Steps in Video Computing

- Acquire   (CCD arrays/synthesize (graphics))
- Process   (image processing)
- Analyze   (computer vision)
- Transmit   (compression/networking)
- Store   (compression/databases)
- Retrieve   (computer vision/databases)
- Browse   (computer vision/databases)
- Visualize (graphics)

Computer Vision

- Measurement of Motion
  - 2-D Motion
    - optical flow
    - point correspondences
  - 3-D Motion
    - structure from motion (sfm)
    - compute 3D translation, 3D rotation
    - shape from motion (depth)

Computer Vision (contd.)

- Scene Change Detection
  - consecutive frame differencing
  - background differencing
    - median filter
    - pfinder
    - W4
    - Mixture of Gaussians

Computer Vision (contd.)

- Tracking
  - people
  - Vehicles
  - animals

Computer Vision (contd.)

- Video Recognition
  - activity recognition
  - gesture recognition
  - facial expression recognition
  - lipreading
- Video Segmentation
  - shots
  - scenes
  - stories
  - key frames

Image Processing

- Filtering
- Compression
  - MPEG-1
  - MPEG-2
  - MPEG-4
  - MPEG-7 (Multimedia Content Description Interface)

Databases

- Storage
- Retrieval
- Video on demand
- Browsing
  - skim
  - abstract
  - key frames
  - mosaics

Networking

- Transmission
- ATM

Computer Graphics

- Visualization
- Image-based Rendering and Modeling
- Augmented Reality

Video Computing

- Computer Vision
- Image Processing
- Computer Graphics
- Databases
- Networks

PART I

Measurement of Motion

Contents

- Image Motion Models
- Optical Flow Methods
  - Horn & Schunck
  - Lucas and Kanade
  - Anandan et al
  - Szeliski
  - Mann & Picard
- Video Mosaics

3-D Rigid Motion

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix}$$

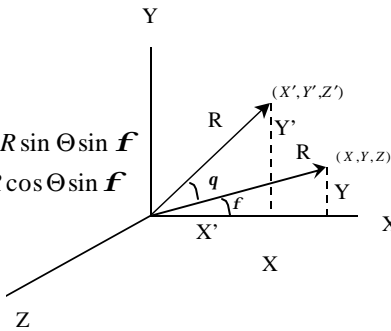Rotation matrix (9 unknowns)   Translation (3 unknowns)

Rotation

$$X = R \cos \boldsymbol{f}$$

$$Y = R \sin \boldsymbol{f}$$

$$X' = R\cos(\Theta + \boldsymbol{f}) = R\cos\Theta\cos\boldsymbol{f} - R\sin\Theta\sin\boldsymbol{f}$$
$$Y' = R\sin(\Theta + \boldsymbol{f}) = R\sin\Theta\cos\boldsymbol{f} + R\cos\Theta\sin\boldsymbol{f}$$

$$X' = X\cos\Theta - Y\sin\Theta$$

$$Y' = X\sin\Theta + Y\cos\Theta$$

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} \cos\Theta & -\sin\Theta & 0 \\ \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

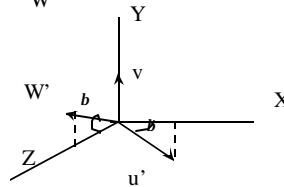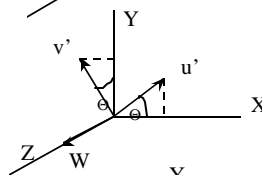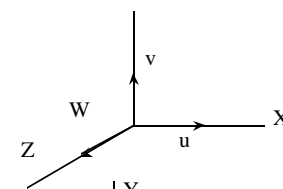# Rotation (continued)

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R = \begin{bmatrix} \cos\Theta & -\sin\Theta & 0 \\ \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R = \begin{bmatrix} \cos\boldsymbol{b} & 0 & -\sin\boldsymbol{b} \\ 0 & 1 & 0 \\ \sin\boldsymbol{b} & 0 & \cos\boldsymbol{b} \end{bmatrix}$$

13

Euler Angles

$$R = R_Z^a R_Y^b R_X^g = \begin{bmatrix} \cos a\cos b & \cos a\sin b\sin g - \sin a\cos g & \cos a\sin b\cos g + \sin a\sin g \\ \sin a\cos b & \sin a\sin b\sin g + \cos a\cos g & \sin a\sin b\cos g - \cos a\sin g \\ -\sin b & \cos b\sin g & \cos b\cos g \end{bmatrix}$$

if angles are small($\cos\Theta \approx 1 \quad \sin\Theta \approx \Theta$

$$R = \begin{bmatrix} 1 & -a & b \\ a & 1 & -g \\ -b & g & 1 \end{bmatrix}$$

CAP 6411 Computer Vision Systems

- Instructor: Dr. Mubarak Shah, shah@cs.ucf.edu, 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 25%, Programs 45% , Homework 10%
- Recommended Book, but not required.
  - Digital Video Processing, A. M. Tekalp, Prentice Hall.

Lecture-1

Multimedia
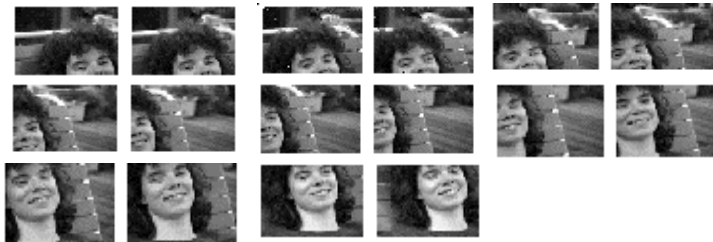
- Text
- Graphics
- Audio
- Images
- Video

Imaging Configurations

- Stationary camera stationary objects
- Stationary camera moving objects
- Moving camera stationary objects
- Moving camera moving objects

Video

- sequence of images
- clip
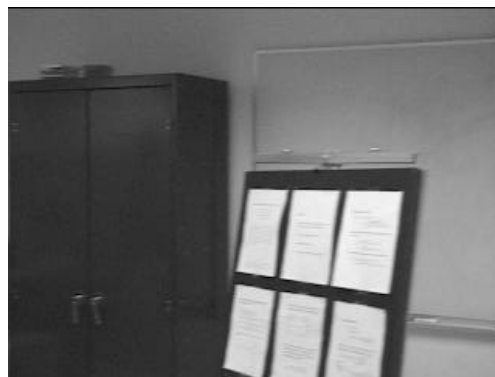- mosaic
- key frames

Sequence of Images

Clip

Mosaic

Key Frames

Clip

Mosaic

Steps in Video Computing

- Acquire   (CCD arrays/synthesize (graphics))
- Process   (image processing)
- Analyze   (computer vision)
- Transmit   (compression/networking)
- Store   (compression/databases)
- Retrieve   (computer vision/databases)
- Browse   (computer vision/databases)
- Visualize   (graphics)

Computer Vision

- Measurement of Motion
  - 2-D Motion
    - optical flow
    - point correspondences
  - 3-D Motion
    - structure from motion (sfm)
    - compute 3D translation, 3D rotation
    - shape from motion (depth)

Computer Vision (contd.)

- Scene Change Detection
  - consecutive frame differencing
  - background differencing
    - median filter
    - pfinder
    - W4
    - Mixture of Gaussians

Computer Vision (contd.)

- Tracking
  - people
  - Vehicles
  - animals

Computer Vision (contd.)

- Video Recognition
  - activity recognition
  - gesture recognition
  - facial expression recognition
  - lipreading
- Video Segmentation
  - shots
  - scenes
  - stories
  - key frames

Image Processing

- Filtering
- Compression
  - MPEG-1
  - MPEG-2
  - MPEG-4
  - MPEG-7 (Multimedia Content Description Interface)

Databases

- Storage
- Retrieval
- Video on demand
- Browsing
  - skim
  - abstract
  - key frames
  - mosaics

Networking

- Transmission
- ATM

Computer Graphics

- Visualization
- Image-based Rendering and Modeling
- Augmented Reality

Video Computing

- Computer Vision
- Image Processing
- Computer Graphics
- Databases
- Networks

PART I

Measurement of Motion

Contents

- Image Motion Models
- Optical Flow Methods
  - Horn & Schunck
  - Lucas and Kanade
  - Anandan et al
  - Szeliski
  - Mann & Picard
- Video Mosaics

3-D Rigid Motion

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix}$$

Rotation matrix (9 unknowns)   Translation (3 unknowns)

12

Rotation
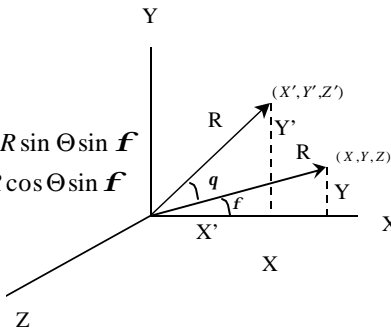
$$X = R\cos f$$

$$Y = R\sin f$$

$$X' = R\cos(\Theta + f) = R\cos\Theta\cos f - R\sin\Theta\sin f$$
$$Y' = R\sin(\Theta + f) = R\sin\Theta\cos f + R\cos\Theta\sin f$$

$$X' = X\cos\Theta - Y\sin\Theta$$

$$Y' = X\sin\Theta + Y\cos\Theta$$

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} \cos\Theta & -\sin\Theta & 0 \\ \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$
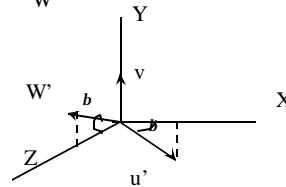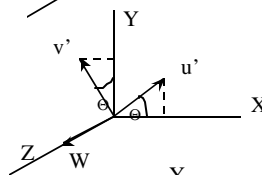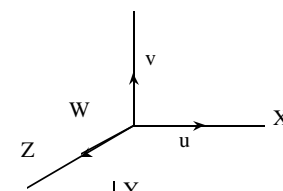

Rotation (continued)

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R = \begin{bmatrix} \cos\Theta & -\sin\Theta & 0 \\ \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R = \begin{bmatrix} \cos b & 0 & -\sin b \\ 0 & 1 & 0 \\ \sin b & 0 & \cos b \end{bmatrix}$$

Euler Angles

$$R = R_Z^a R_Y^b R_X^g = \begin{bmatrix} \cos a \cos b & \cos a \sin b \sin g - \sin a \cos g & \cos a \sin b \cos g + \sin a \sin g \\ \sin a \cos b & \sin a \sin b \sin g + \cos a \cos g & \sin a \sin b \cos g - \cos a \sin g \\ -\sin b & \cos b \sin g & \cos b \cos g \end{bmatrix}$$

if angles are small($\cos \Theta \approx 1$ $\sin \Theta \approx \Theta$

$$R = \begin{bmatrix} 1 & -a & b \\ a & 1 & -g \\ -b & g & 1 \end{bmatrix}$$