Computer Vision
CS 682

Jana Kosecka
http://cs.gmu.edu/~kosecka/cs682.html
kosecka@cs.gmu.edu
2-9827

Logistics

- **Grading:** Homeworks (about every 2 weeks) 30% Midterm: 30% Final project: 40%
- **Prerequisites:** linear algebra, calculus
- **Required Text:**
- From Images to Geometric Models: Y. Ma, S. Soatto, J.Kosecka and S. Sastry, Springer Verlag 2003
- Introductory Techniques for 3D Computer Vision
  (E. Trucco, A. Verri, Prentice Hall, 1998)
- Computer Vision a Modern Approach
  (D. Forsyth, J. Ponce, Prentice Hall 2002)

- Required Software MATLAB (with Image Processing toolbox)

Biological motivations

Understanding visual sensing modality and its role

- We have no difficulties to navigate, manipulate objects
recognize familiar places and faces

- How can we successfully carry out all these everyday tasks ?
- How does the visual perception mediates these activities ?

- Overall system
Sensation and perception – integrate and interpret sensory readings
Behavior – control muscles and glands to mediate some behavior
Memory – long term memory (declarative – faces, places, events)
short term memory (procedural – associations, skills)
Higher level functions – internal models and representations of the environments

Goal of Computer Vision

- Build machines and develop algorithms which can automatically replicate some funcionalities of biological visual system

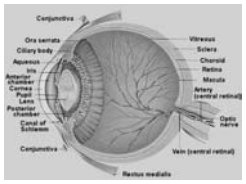
- Systems which navigate in cluterred environments
- Systems which can recognize objects, activities
- Systems which can interact with humans/world

Synergies with other disciplines and various applications
Artificial Intelligence (Robotics, Natural Language Understanding)
Vision as a sensor – medical imaging, Geospatial Imaging, robotics
visual surveilance, inspection

Computer Vision

Visual Sensing

Images I(x,y) – brightness patterns



- image appearance depends on structure of the scene
- material and reflectance properties of the objects
- position and strength of light sources

Challenges/Issues

- About 60% of our brain is devoted to vision
- We see immediately and can form and understand images instantly

- Several strategies for forming a representation of the scenes
- Detailed representations are often not necessary
- Different approaches in the past Animate Vision (biologically inspired), Purposive Vision (depending on the task/purpose)

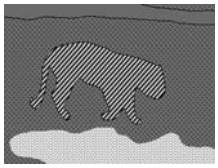
- Recovery of the properties of the environment from single or multiple views

Vision problems (towards image understanding)

- Segmentation
- Recognition
- Reconstruction
- Vision Based Control - Action

Visual Cues

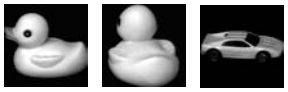
Segmentation – partition image into separate objects



- Clustering and search algorithms in the space of visual cues

Recognition – given an image classify what object it represents

- Face/Object recognition
- Digit recognition
- Activity recognition



- Patter Recognition, Machine Learning techniques

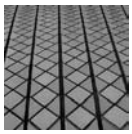
Focus of this course

The focus of the course :

1. Geometry of Single and Multiple Views

Shape and Motion Recovery, Matching, Alignment Problems
Reconstruction (from 2D to 3D)
Visually guided Control
Pictorial cues – shading, texture,
blur, contour
Stereo, motion cues

2. Object Detection and Recognition



1. Geometry of Single and Multiple Views

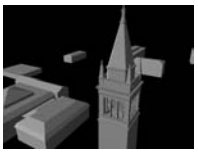
How to reliably recover and represent the geometric model from single image or video and camera motion/pose

Representation issues depends on the task/applications

- Image-based rendering, Computer Graphics
- Virtual and Augmented Reality
- Vision based control, surveillance
- Human computer interaction

- Medical imaging (alignment, monitoring of change)
- Video Analysis

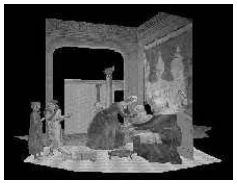
Vision and Computer Graphics
- image based rendering techniques
- 3D reconstruction from multiple views or video
- single view modeling
- view morphing (static and dynamic case)



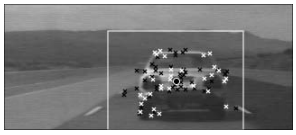
Modeling with Multiple Images

University High School, Urbana, Illinois
Three of Twelve Images

Image Based Rendering – View Interpolation



Vision-Based Control, Surveillance applications
- continuously changing action in response to video input



Automated Driving

Visual navigation

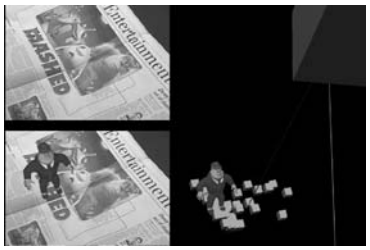
Automated Landing

Visual surveillance

wide area surveillance, traffic monitoring
Interpretation of different activities

Virtual and Augmented Reality, Human computer Interaction

Virtual object insertion
various gesture based interfaces
Interpretation of human activities
Enabling technologies of intelligent homes, smart spaces

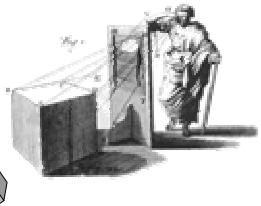


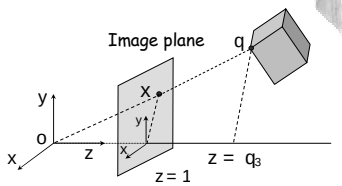
Topics

- Image Formation, Representation of Camera Motion, Camera Calibration
- Image Features – filtering, edge detection, point feature detection
- Image alignment, 3D structure and motion recovery, stereo
- Analysis of dynamic scenes, detection, tracking
- Object detection and object recognition

Basic Ingredients – Course Overview

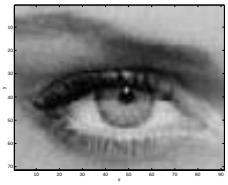
Rigid body motion
Image Formation
Perspective Projection



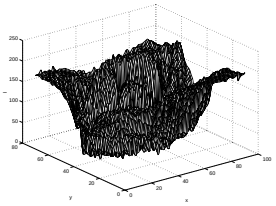


Feature Detection and Correspodence

Image

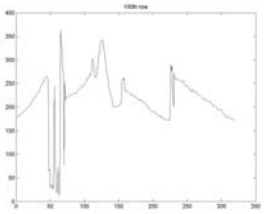


Brightness values



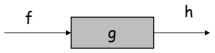
I(x,y)

Original Image



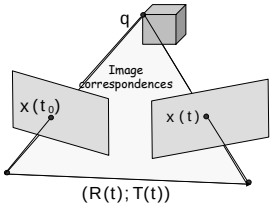
100th row of the image

Low level global operations on the image (signal) – linear time invariant system

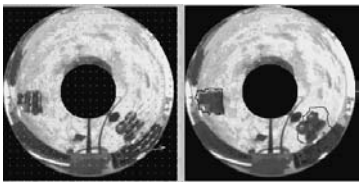


Two view motion estimation
Recovery of 3D Structure

Epipolar constraint

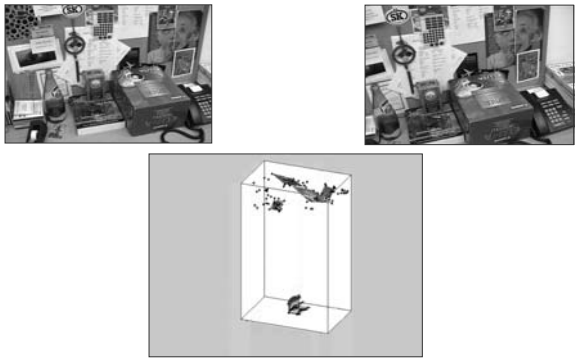


Motion segmentation

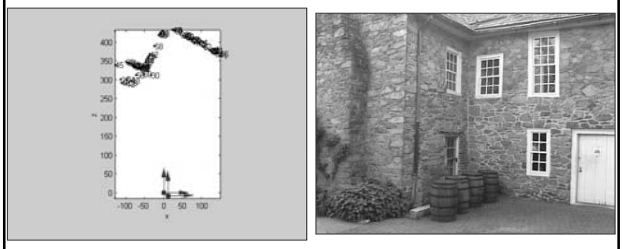


Examples

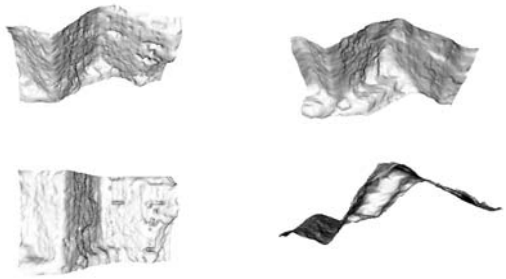
Example - Euclidean multi-view reconstruction



Euclidean Reconstruction



Texture mapping, hole filling



Texture mapping

Modeling choices
And
Model recovery

How to use prior scene knowledge