Course Projects

Sep 13, 2012

Course Projects

- Covers 50% of your grade
- 10-12 weeks of work
- Required:
  - Serious commitment to project
  - Extra points for working demonstration
  - Project Report
  - Poster presented in poster session
  - Graded by anonymous external reviewers in addition to the course instructors

Project Complexity

- Depends on what you want to do
- Complexity of the project will be considered in grading.
- Projects typically vary from cutting-edge research to reimplementation of existing techniques. Both are fine.

More details

- Projects will be done in teams of 2 or 3
- It is ok to work alone but your project will be no simpler
- If you cannot find teammates, email the TA

- Teams will have to spend a lot of time understanding the problem.

- Team members will also grade each other to make sure that everybody contributes

Incomplete Projects

- Be realistic about your goals.

- Incomplete projects can still get a good grade if
  - You can demonstrate that you made progress
  - You can clearly show why the project is infeasible to complete in one semester

- Remember: You will be graded by peers

Possible projects

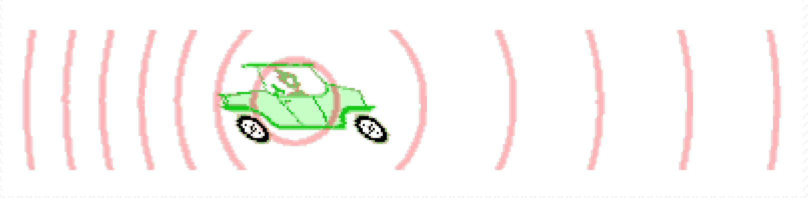
- A list of possible projects will be presented in the rest of this lecture
- You are also free to pick your own project.
- Teams must inform us of their choice of project by (mumble,mumble).
  - The later you start, the less time you have to work on the project

Projects from previous years

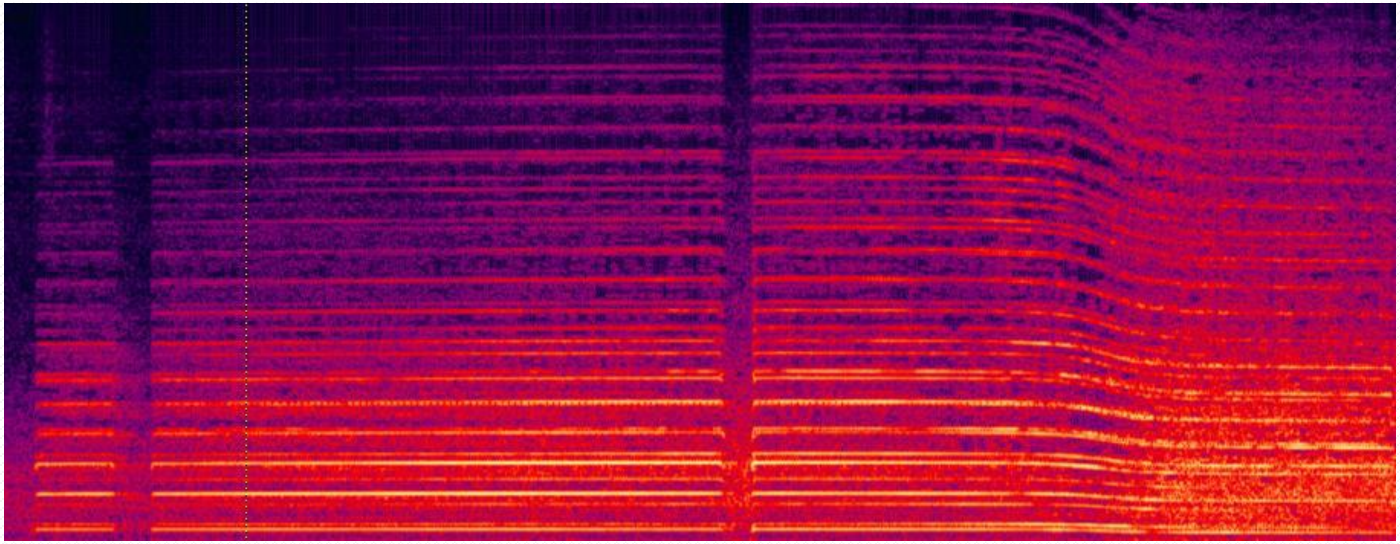
- Non-intrusive load monitoring
- Seam carving
- Statistical Klatt Parametric Synthesis
- Voice Transformation using Canonical Correlation analysis
- Sound source separation and missing feature enhancement
- Counting blood cells in cerebrospinal fluid
- And many more …

The Doppler Effect

- The observed frequency of a moving sound source differs from the emitted frequency when the source and observer are moving relative to each other



The Doppler Effect

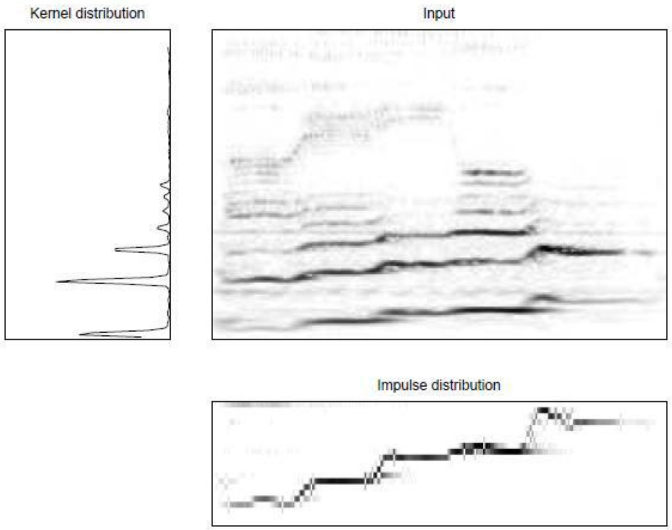


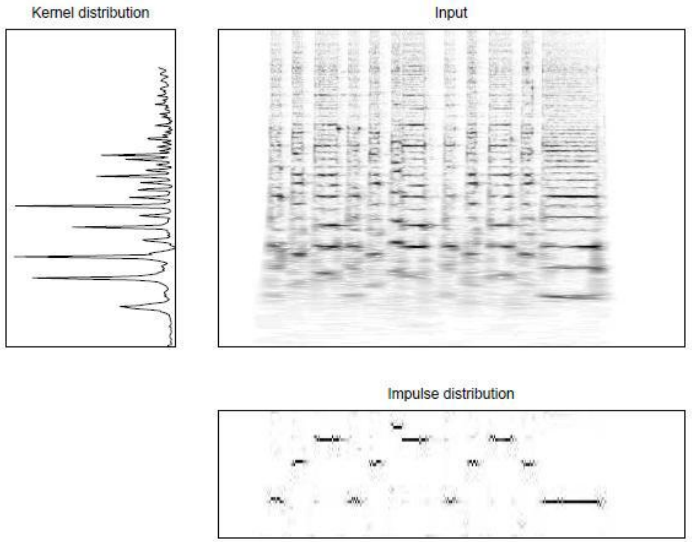
- Spectrogram of horn from speeding car
  - Tells you the velocity
  - Tells you the distance of the car from the mic

Problem

- Analyze audio from speeding automobiles to detect velocity using the Doppler shift
- Find the frequency shift and track velocity/position
- Supervisor: Dr. Rita Singh

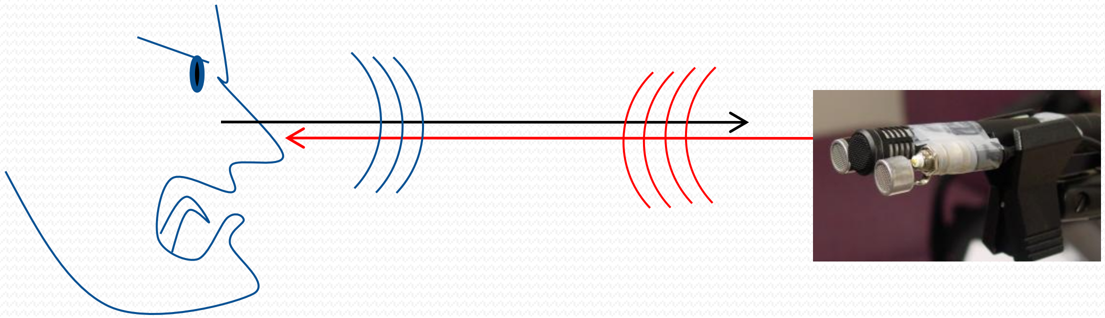
Pitch Tracking





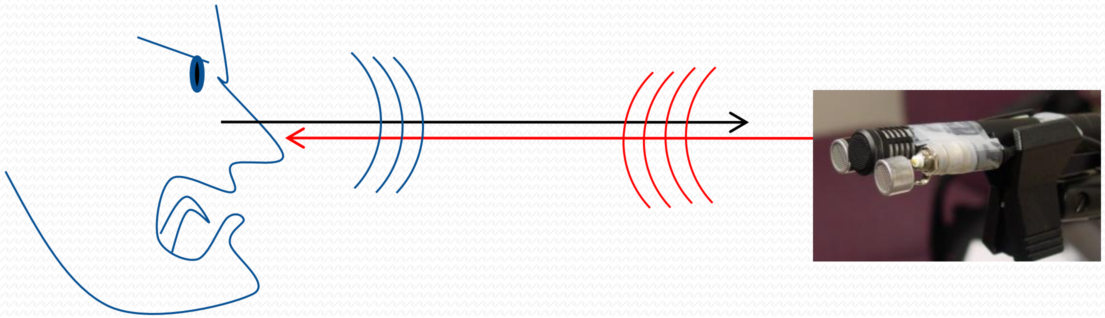
- Frequency shift invariant latent variable analysis
- Combined with Kalman filtering
- Estimate the velocity of *multiple* cars at the same time

More on Doppler



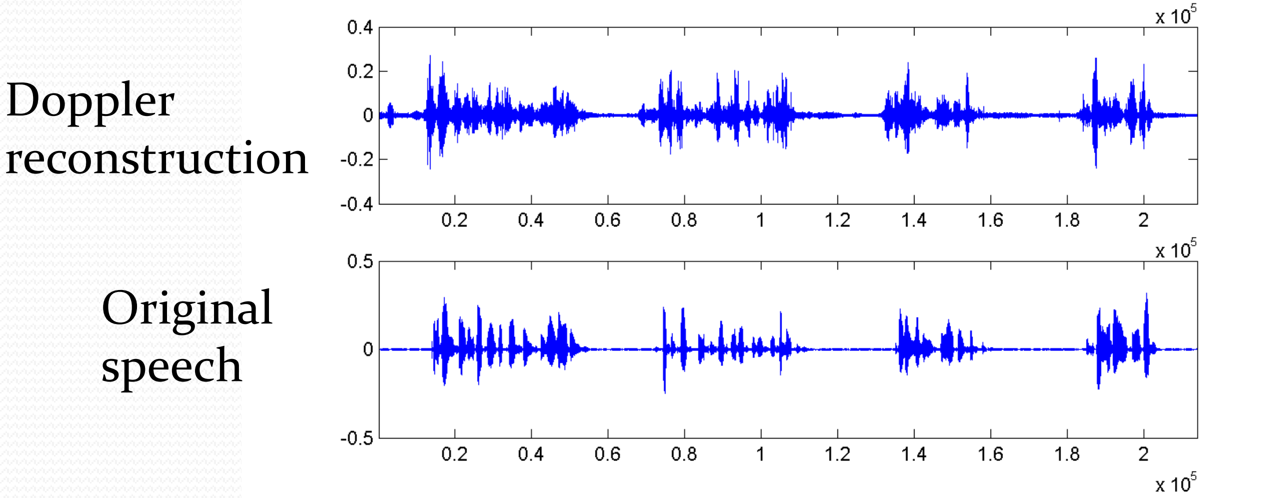
- Reflections of a 40khz tone from a speaker's face have Doppler shifts
- These capture facial movements related to speech
- They represent articulator movements of the speaker
- Prior work:
  - Recognizing the speaker from the Doppler measurements
  - Resynthesizing the *speech* from the Doppler measurements of the speaker's face

Identifying talking faces



- Beam ultrasound on talker's face
- Capture and analyze reflections
- Identify subject

Synthesizing Sound from ultrasound observations



- Subject mimes sound but does not produce any sound
- Can we produce sound with just the ultrasound observations?

New Doppler Problem

- Can we learn to derive articulator information from speech by considering its relationship to Doppler signal
- Can this be used to improve automatic speech recognition performance

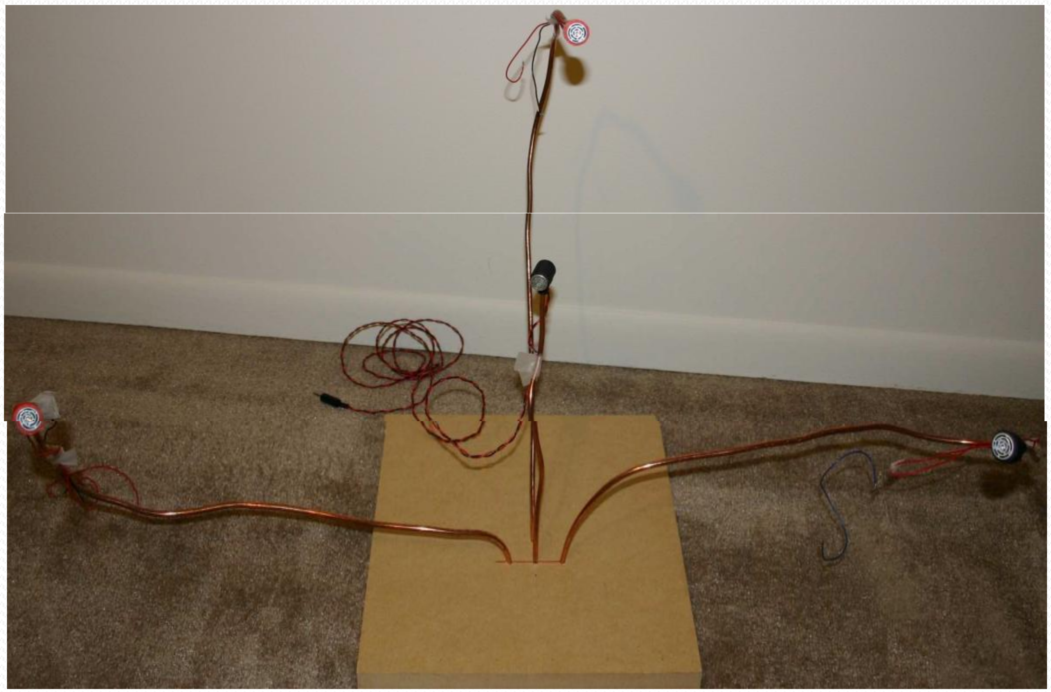
- Procedure
  - Learn a deep neural network to learn the mapping
  - Use the network as a feature computation module for speech recognition
    - Augments conventional features

- Supervisor: Bhiksha Raj

Doppler from walking person

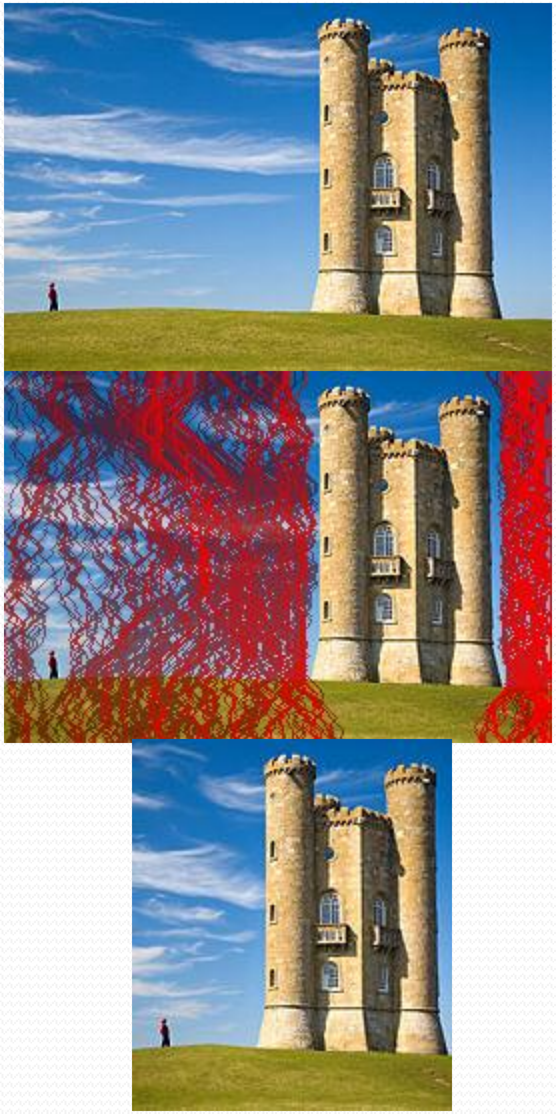
- Gait recognition
- Beam ultrasound at walking subjects
- Capture reflections
- Determine identity of the person

Gesture recognizer

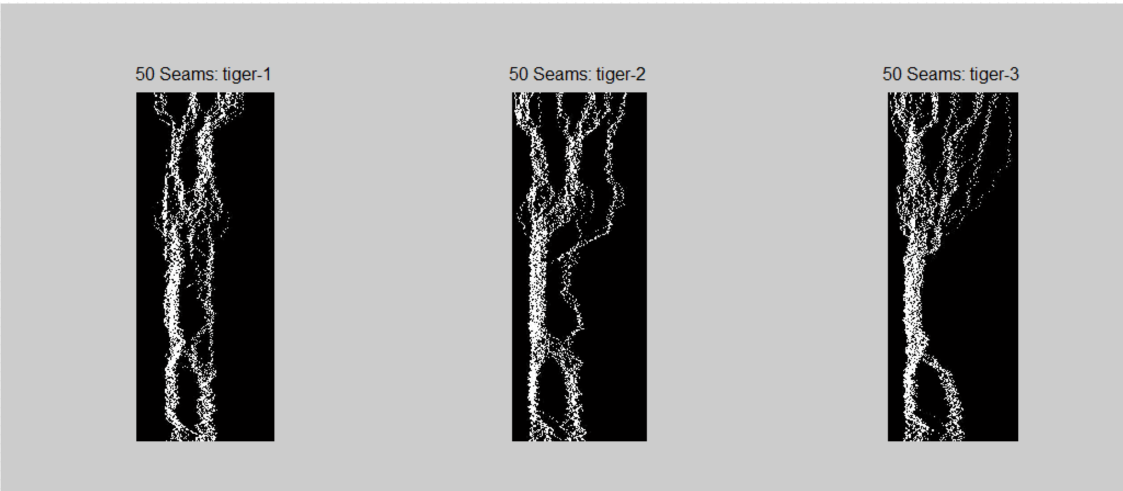


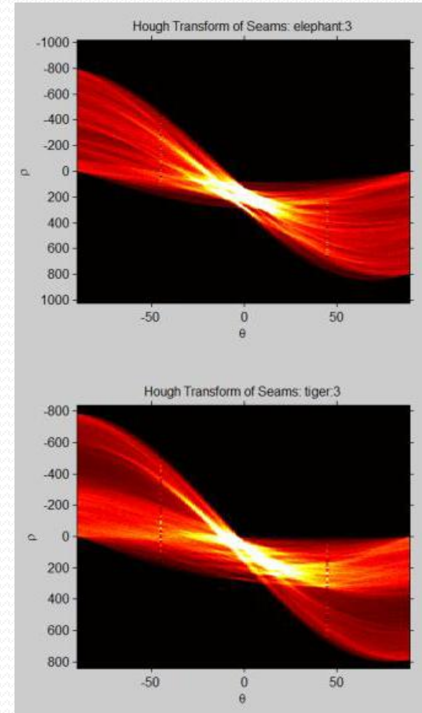
- Recognizing gestures and the actions that constitute a gesture

Seam Carving



Seam carving for word spotting (Rita Singh)





- Seams in spectrograms: Word specific
- Characterize seams to recognize/detect words
    - Combine with traditional methods for improved performance

Song lyric recognition (Rita Singh)

- Recognize lyrics in songs

- Conventional Automatic Speech recognition won't work
  - Stylized voices
  - Overlaid music
  - Mispronunciations

- Can assume any framework
  - Select lyrics from a collection of lyrics
  - Know words but not lyrics

De-reverberation

- Develop a *supervised* technique that can dereverberate a noisy signal
  - Know what is spoken and has prior information about speaker
  - Will work with artificially reverberated data

- Issues:
  - Modeling the data
  - Learning parameters
  - Overcomplete representations

Sound Classification

- Identifying cars from their sound
- Simple problem: Can we build a system that can identify the make (and possibly model) of a car by listening to it?
  - Can you make out the difference between a V6 and a V8 engine?
- Issues:
  - Gathering training data
  - Modeling

Face Recognition

- Similar to the face detector, but now we want to *recognize* the faces too
  - Who was it that walked by my office?
- Variety of existing techniques available
- Can be combined with face detection

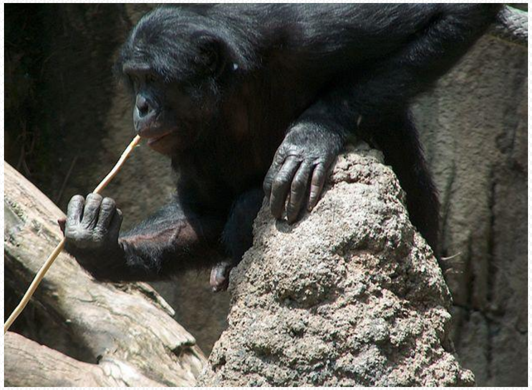
Recognizing the gender of a face

- A hard problem
- Even humans are bad at this

Image Manipulation: Filling in

- Some images are often occluded
- Search a database to find objects that best fit into the occluded region

Bonobo ‘speech’ analysis



- Bonobos and chimpanzees are humans’ closest living relatives
- Bonobos vocalize in a way similar to humans
- Need to make sense of several Terabytes of data where bonobos interact with humans
- Supervisor: Prof. Alan Black

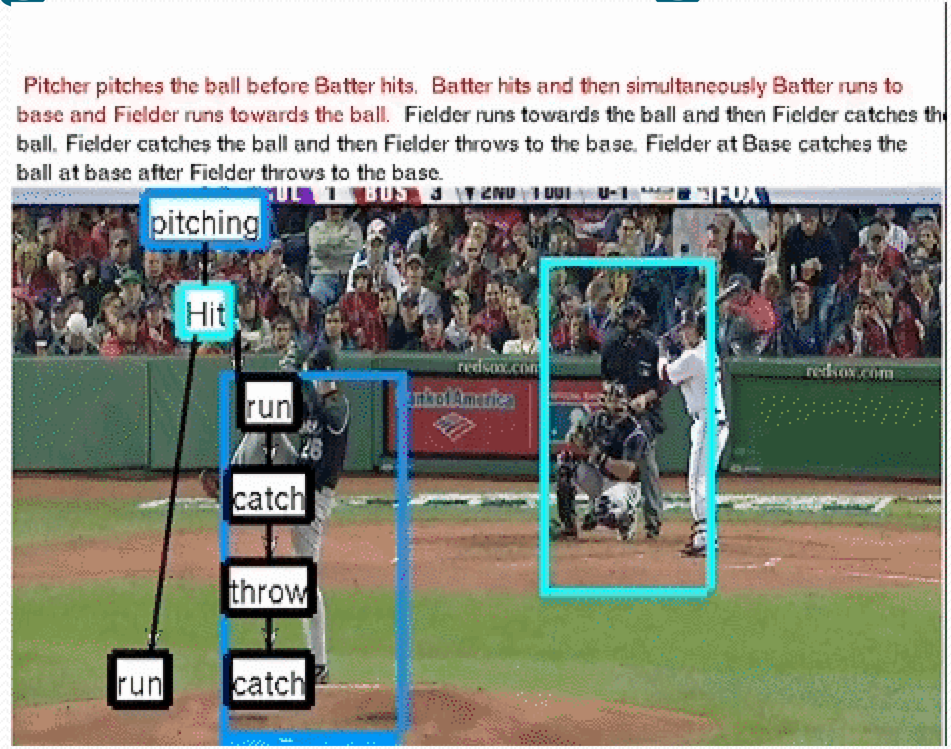
Detecting buses

- Detecting buses that stand at Forbes and Craig so that you can stay in your office in Gates and work until the bus comes.
- Need to use the audio or visual data to detect the presence of buses in video.
- Supervisor: Prof. Alan Black + possibly others

Emotion detection from audio/images

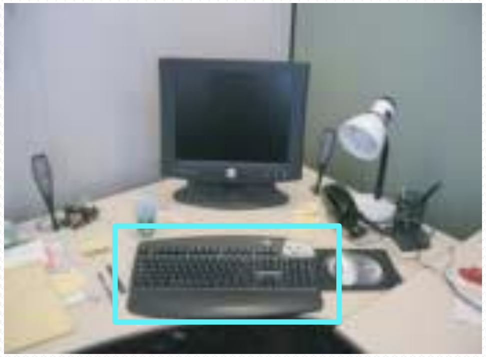
- Detecting and recognizing the emotion in faces
- Doing the same from voices

Assigning Semantic tags to video



- http://www.cs.cmu.edu/~abhinavg/Home.html

Object detection and Clustering



- Detect various types of objects in images
  - Supervised: You know what objects to detect
  - Unsupervised: Detect objects based on motion

Scene segmentation with audio

- Identify change of scene with audio alone
  - A set of speakers is scene specific
  - The background conditions change
  - Detect when the change is significant

Scene segmentation with video

- Automatically detect discontinuity in the narrative with video alone
  - Automatic shot change detection

- Scene change detection. A scene may have multiple shots

Some more ideas will be put on the website

Questions?