Course Projects

Class 9. 22 Sep 2009

Administrivia

- **THURSDAY'S CLASS: WEAN HALL 5403**
  - **Thanks to Ramkumar Krishnan for arranging the room!**
- Almost all submissions of Homework 1 are in
  - Thanks to all students who have submitted
  - Three submissions are still due
- Fernando's lecture
  - Clarifications required? J
- Homework 2 is up on the website
  - Face detection using a single Eigen face
  - Will expand to using multiple Eigen faces in stage 2
    - Complex homework
    - Homework 3 will be very simple: L1 estimation of L2 algebraic operations
      - If (insufficient(time)==true) givenhomework(3) = false

Course Projects

- Covers 50% of your grade
- 9-10 weeks
- Required:
  - A seriously attempted project
  - Demo if possible
  - Project report
  - 20 minute project presentation
- Project complexity
  - Depends on what you choose to do
  - Complexity of project will be considered in grading

Course Projects

- Projects will be done by teams of students
  - Ideal team size: 4
  - Find yourself a team
  - If you wish to work alone, that is OK
    - But we will not require less of you for this
  - If you cannot find a team by yourselves, you will be assigned to a team
  - Teams will be listed on the website
  - All currently registered students will be put in a team eventually
- Will require background reading and literature survey
  - Learn about the the problem
- Grading will be done by team
  - All members of a team will receive the same grade
    - But I retain discretionary powers over this

Projects

- A list of possible projects will be presented to you in the rest of this lecture
- This is just a sampling
- You may work on one of the proposed projects, or one that you come up with yourselves
- Teams must inform us of their choice of project by 29th September 2009
  - The later you start, the less time you will have to work on the project

Projects

- Projects range from simple to very difficult
  - Important to work in teams
- Guest lecturers with project ideas
  - Anatole Gershman (LTI)
  - Alan Black (LTI)
  - Eakta Jain (RI)
  - Fernando De La Torre
    - Not presenting
- Important: Be realistic
  - Partially completed projects will still get grades ***IF:***
    - The work performed is a serious attempt at completing it
  - But only completed projects are likely to result in papers/publications if any

Now.. To our guests..

- Alan Black
- Anatole Gershman
- Eakta Jain

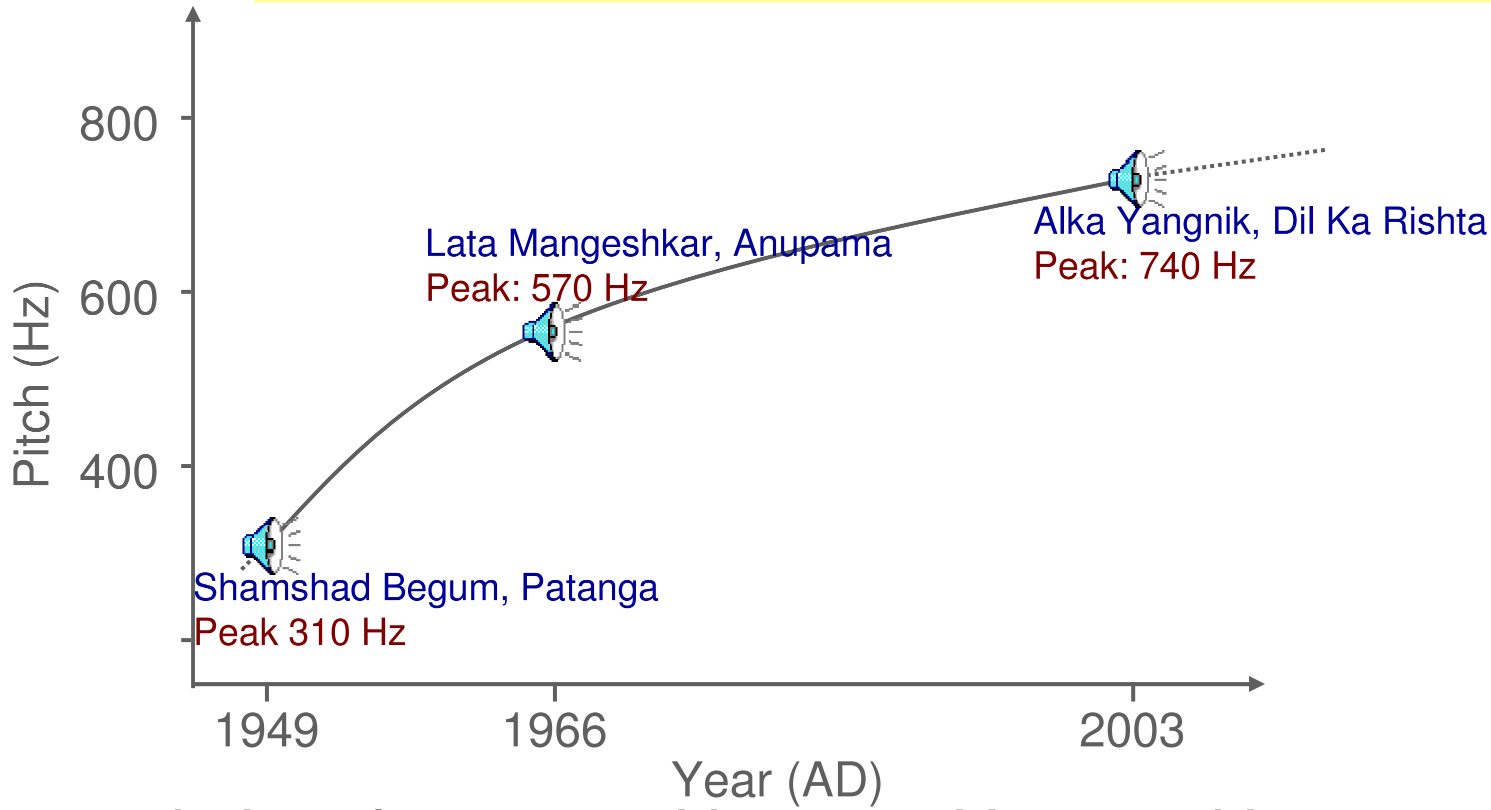
More Project Ideas

- Sound
  - Separation
  - Music
  - Classification
  - Synthesis
- Images
  - Processing
  - Editing
  - Classification
- Video
  - …
  - …

A Strange Observation

- A trend

The pitch of female Indian playback singers is on an ever-increasing trajectory



- Mean pitch values: 278Hz, 410Hz, 580Hz

I'm not the only one to find the high-pitched stuff annoying

- Sarah McDonald (Holy Cow): "... shrieking…"

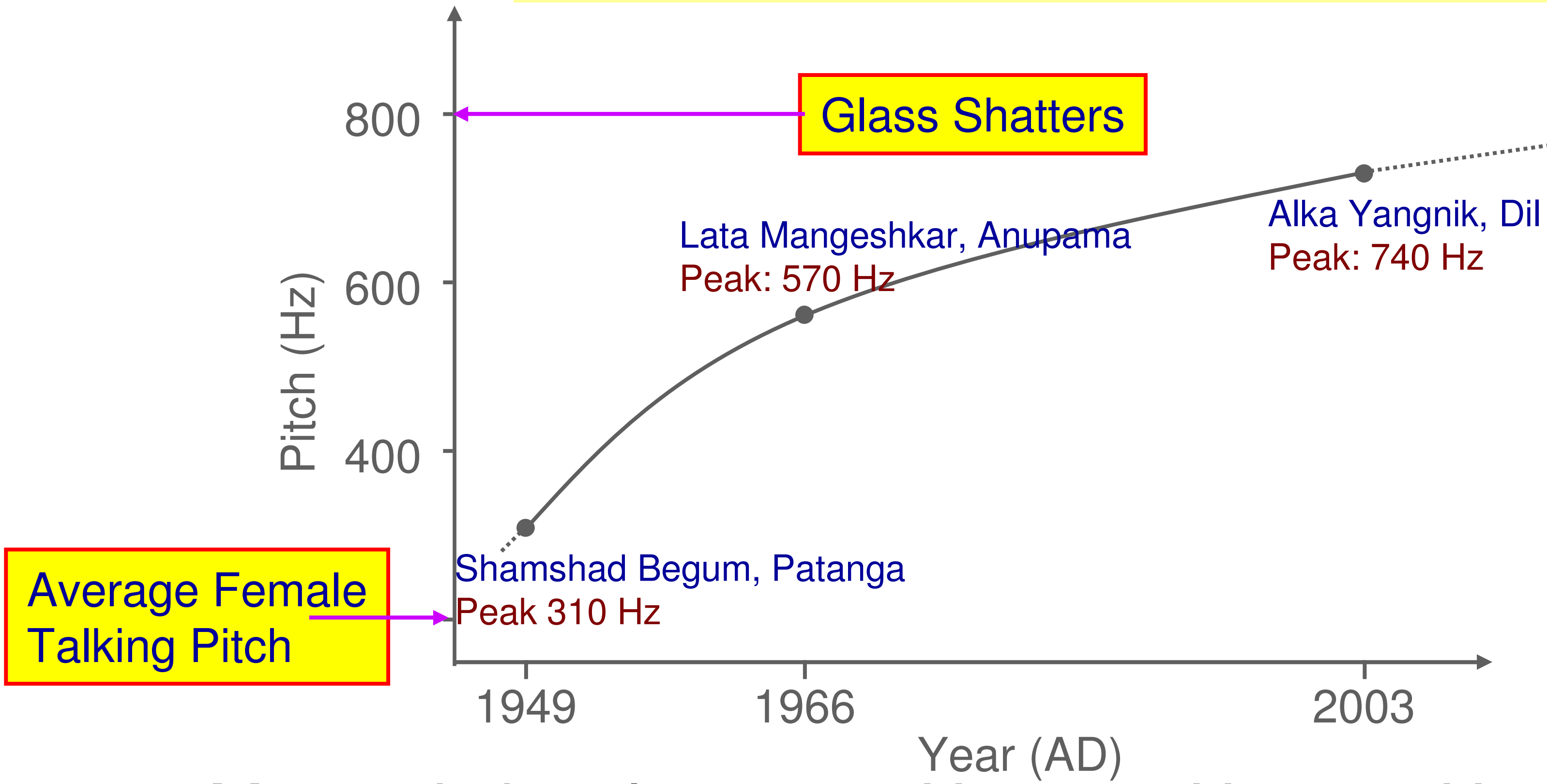
- Khazana.com: "... female Indian movie playback singers who can produce ultra high frequncies which only dogs can hear clearly.."

- www.roadjunky.com: "... High pitched female singers doing their best to sound like they were seven years old .."

A Disturbing Observation

- A trend

The pitch of female Indian playback singers is on an ever-increasing trajectory



- Mean pitch values: 278Hz, 410Hz, 580Hz

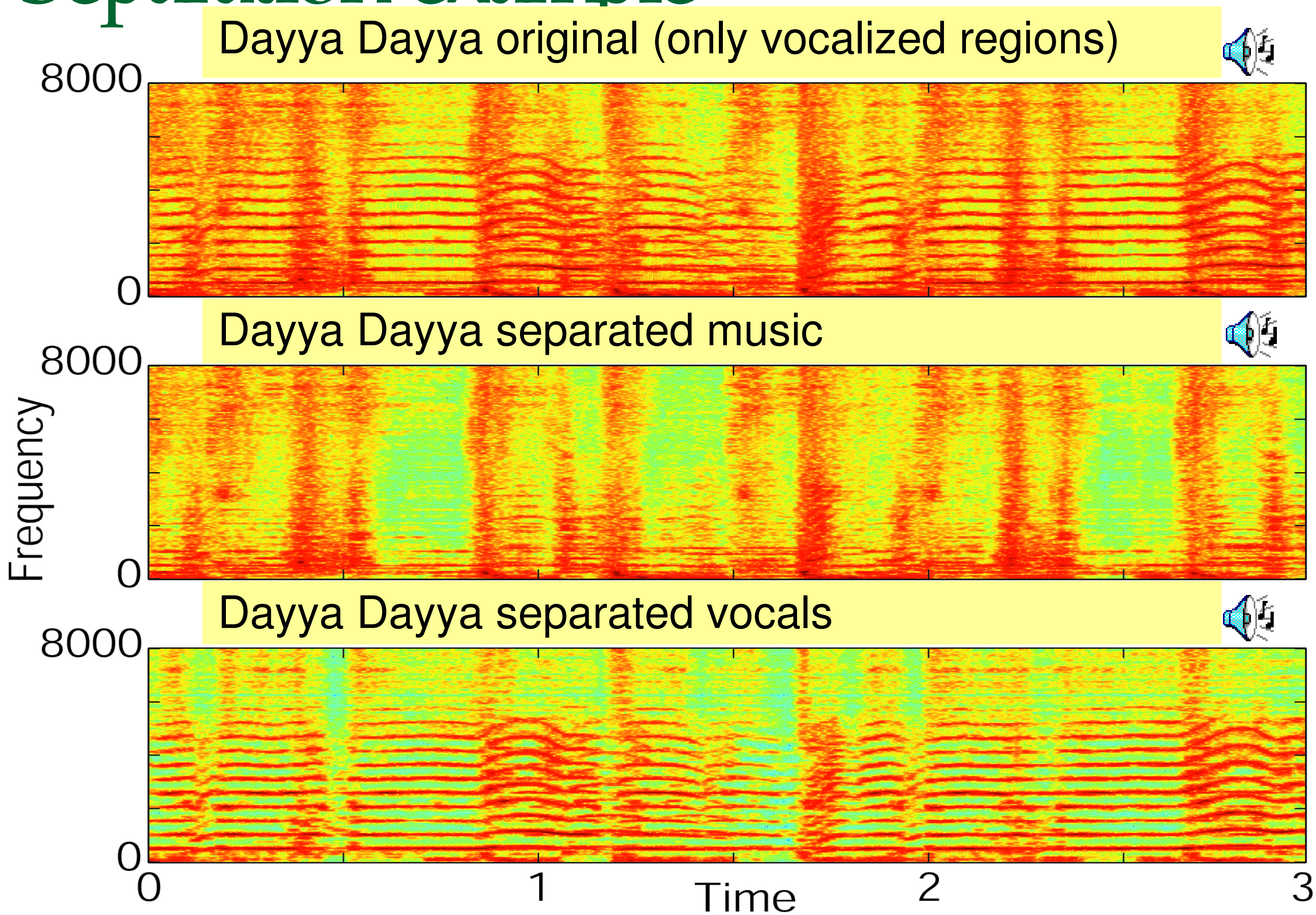
Subjectivity of Taste

- High pitched female voices can often sound unpleasant
- Yet these songs are very popular in India
  - Subjectivity of taste
- The melodies are often very good, in spite of the high singing pitch

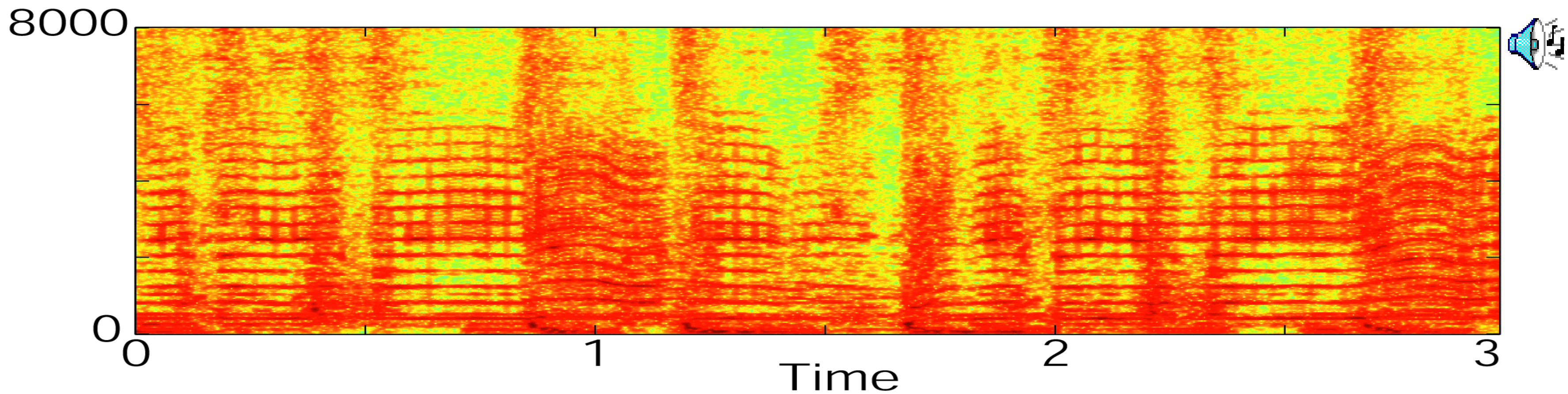
“Personalizing” the Song

- Retain the melody, but modify the pitch
  - To something that one finds pleasant
  - The choice of “pleasant” pitch is personal, hence “personalization”
- Must be able to separate the vocals from the background music
  - Music and vocals are mixed in most recordings
  - Must modify the pitch without messing the music
- Separation need not be perfect
  - Must only be sufficient to enable pitch modification of vocals
  - Pitch modification is tolerant of low-level artifacts
    - For octave level pitch modification artifacts can be undetectable.

Separation example

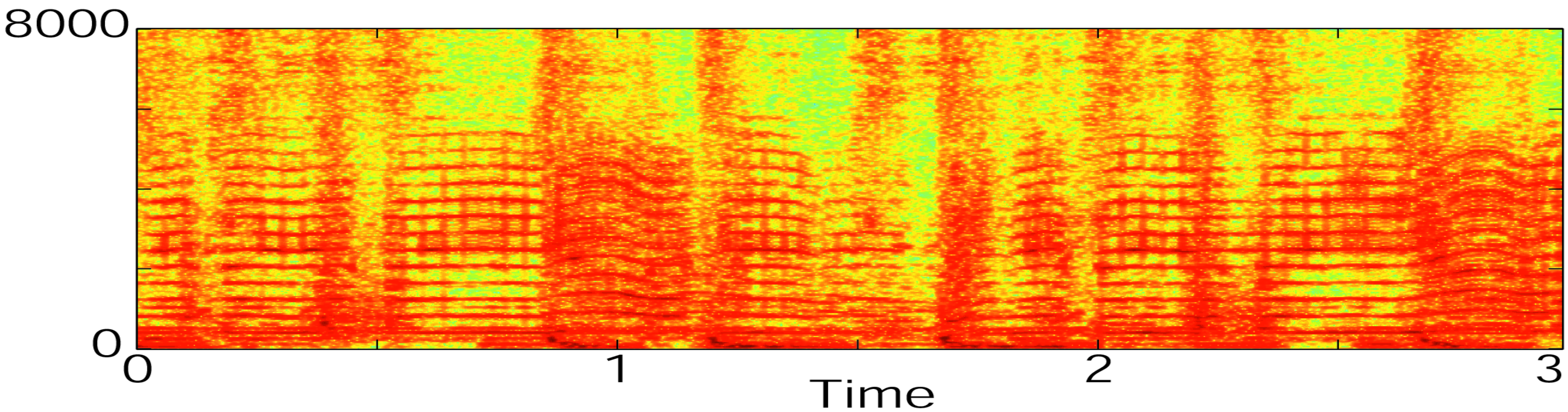


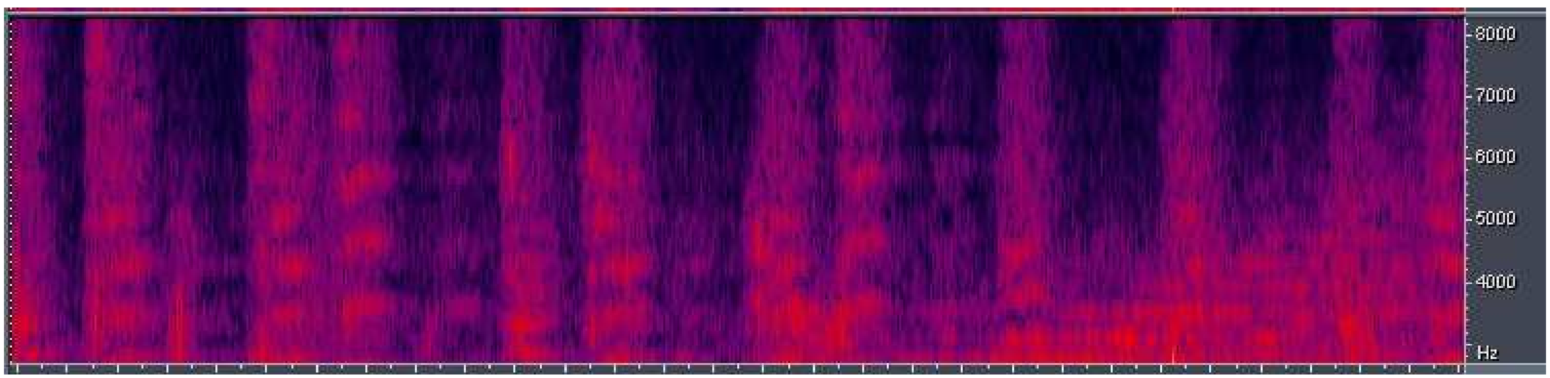
Some examples



- Example 1: Vocals shifted down by 4 semitones

Some examples





- Example 1: Vocals shifted down by 4 semitones
- Example 2: Gender of singer partially modified

Projects..

- Several component techniques
- Illustrate various ML *and* signal processing concepts

- Signal separation
  - Latent variable models
  - Non-negative factorization
- Signal modification
  - Pitch and spectral modification
  - Phase and phase estimation

Song "Personalizer"

- Modify vocals as desired
  - Mono or Stereo
  - "Knob" control to modify pitch of vocals
- Given a song
  - Separate music and song
  - Modify pitch as required
  - Adjust parameters for minimal artifacts
  - Add..
- Issues:
  - Separation
  - Modification
  - Use of appropriate statisical model and signal processing

Talk-Along Karaoke

- Pick a song that features a prominent vocal lead
  - Preferably with only *one* lead vocal
- Build a system such that:
  - User talks the song out with reasonable rhythm
  - The system produces a version of the song with the user *singing* the song instead of the lead vocalist
    - i.e. The user's singing voice now replaces the vocalist in the song
- No. of issues:
  - Separation
  - Pitch estimation
  - Alignment
  - Pitch shifting

Dereverberation

Sound recorded in an Auditorium

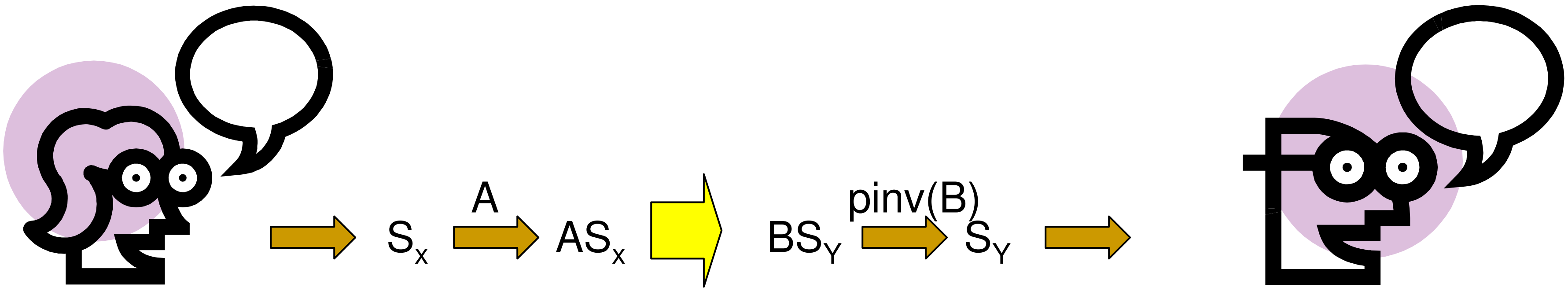
Dereverberated (with artifacts)

- Develop a *supervised* technique that can dereverberate a noisy signal
  - Will work with artificially reveberated data
- Issues:
  - Modeling the data
  - Learning parameters
  - Overcomplete representations

Real-time music transcription

- Proposed by Siddharth Hazra
- Discover sheet music for a guitar on-line, as it is played

Voice transformation with Canonical Correlation Analysis



- Canonical correlation Analysis:
  - Given spectra S_x from speaker X
  - And spectra S_y from speaker Y
  - Find transform matrices A and B such that AS_x *predicts* BS_y
- Will *transform* the voice of speaker X to that of speaker Y
- Issues:
  - CCA
  - Voice transformation

The Doppler Ultrasound Sensor

- Using the Doppler Effect

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
  - Discovery attributed to Christian Doppler (1803-1853)

Person being approached by a police car hears a higher frequency than a person from whom the car is moving away

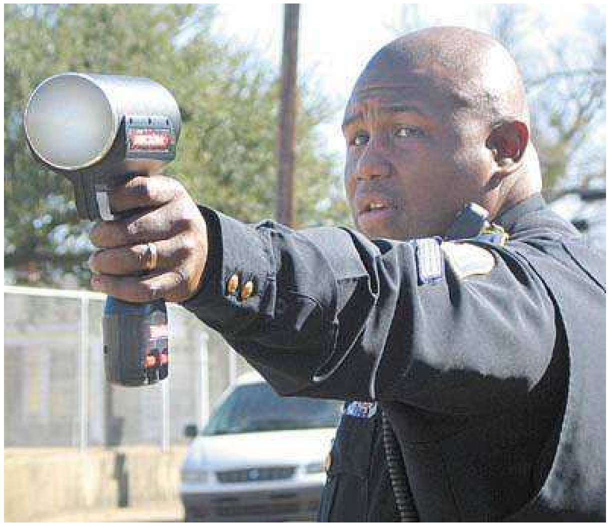
Observed frequency

- The relationship of actual to percieved frequencies is known
- Case 1: The source is moving with velocity *v*, but the listener is static
  - Observed frequency is:

$$f' = \frac{c_{sound} f}{c_{sound} - v}$$

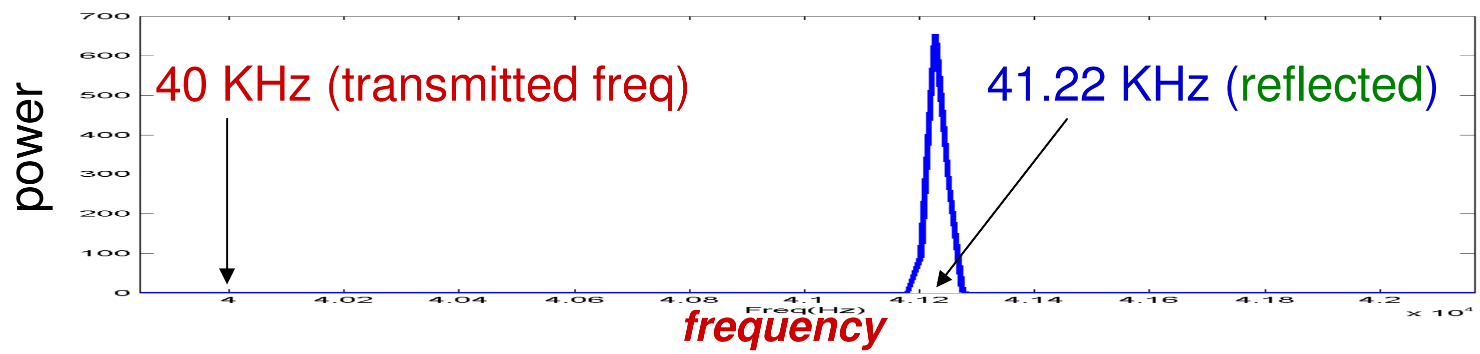
- Case 2: The observer is emitting the signal which is reflected off the moving object
  - Observed frequency is:

$$f' = \frac{(c_{sound} + v) f}{c_{sound} - v}$$

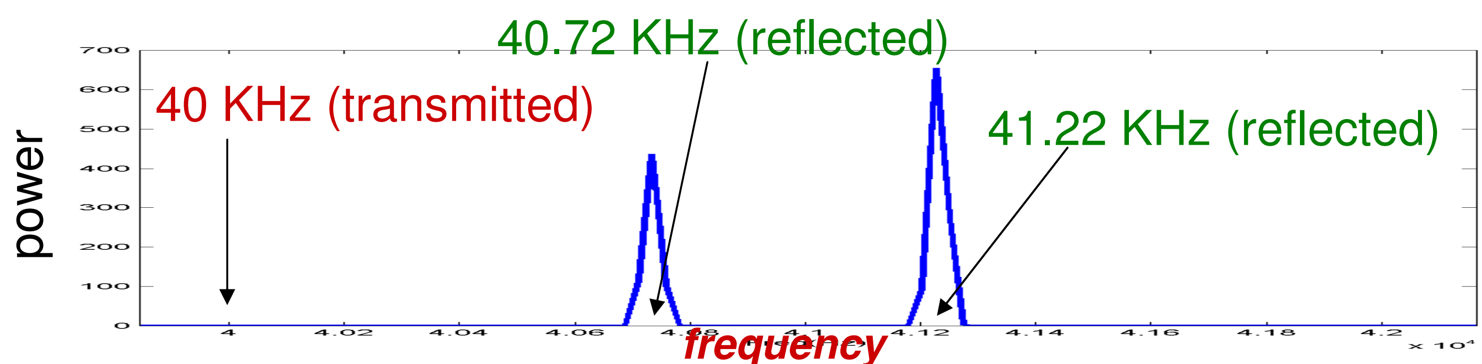


Doppler Spectra

- 40 Khz tone reflected by an object approaching at approximately



- 40 Khz tone reflected by two objects, one approaching at approximately 5m/s and another at 3m/s

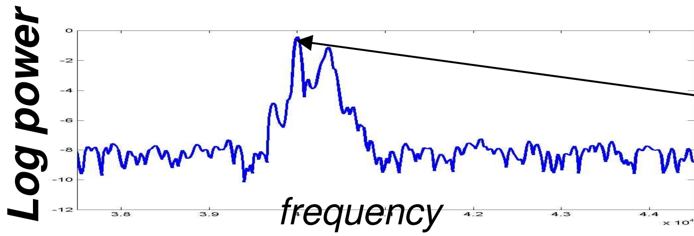


Multiple velocities result in multiple reflected frequencies

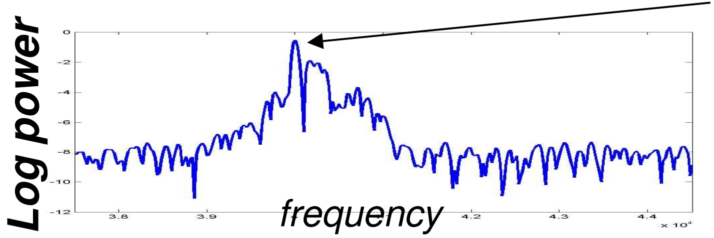
Doppler from Walking Person

- Human beings are articulated objects
- When a person walks, different parts of his body move with different velocities. The combination of velocities is characteristic of the person
  - These can be measured as the spectrum of a reflected Doppler signal

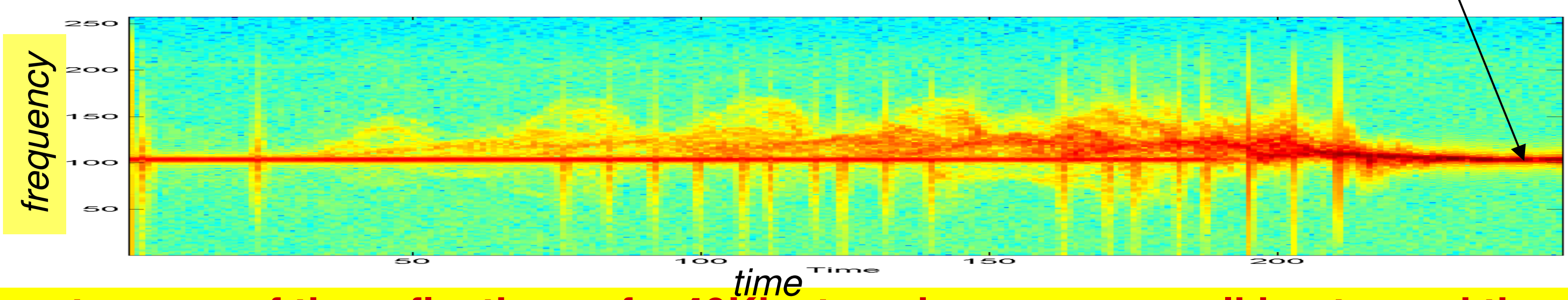
Peak stride: *Frequencies are less spread out*



Mid stride: *Frequencies are more spread out*

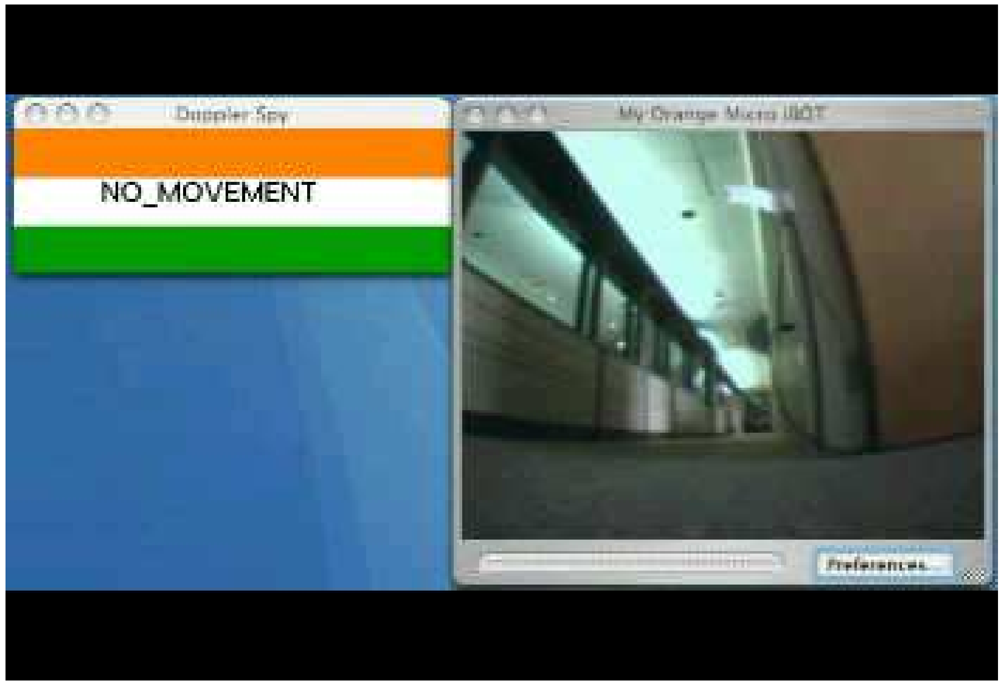


Peaks at the incident frequency (40KHz) from reflections off static objects in environment



spectrogram of the reflections of a 40Khz tone by a person walking toward the sensor
The spikes in the spectrogram are measurement artefacts

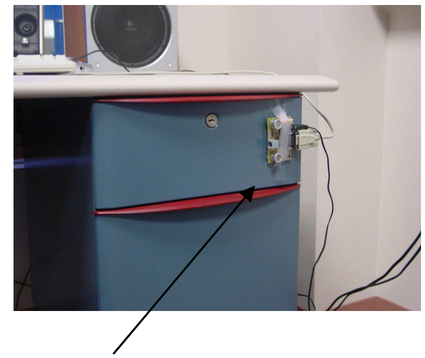
Identifying moving objects



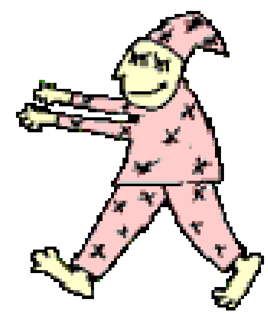
- Doppler spectra are signatures of the motion
  - The pattern of velocities associated with the movement of an object are unique

Gait Recognition

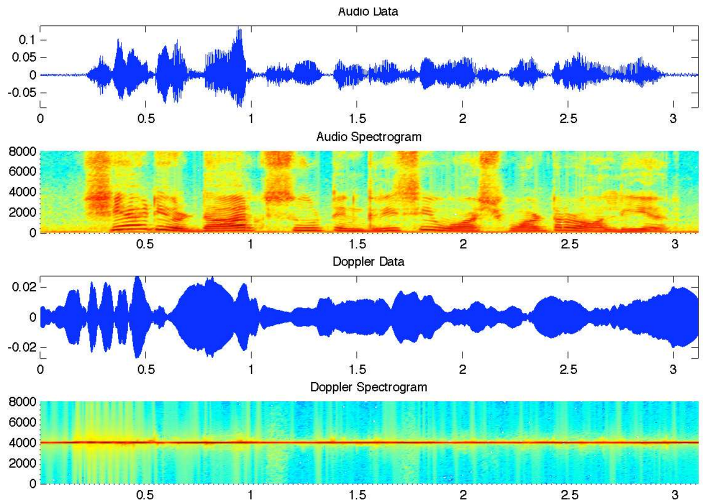
- Beam Ultrasound at a walking subject
- Capture reflections
- Determine identity of subject from analysis of reflections
- Issues:
  - Type of Signal Processing
  - Type of classifier
  - Hardware..



Doppler sensor

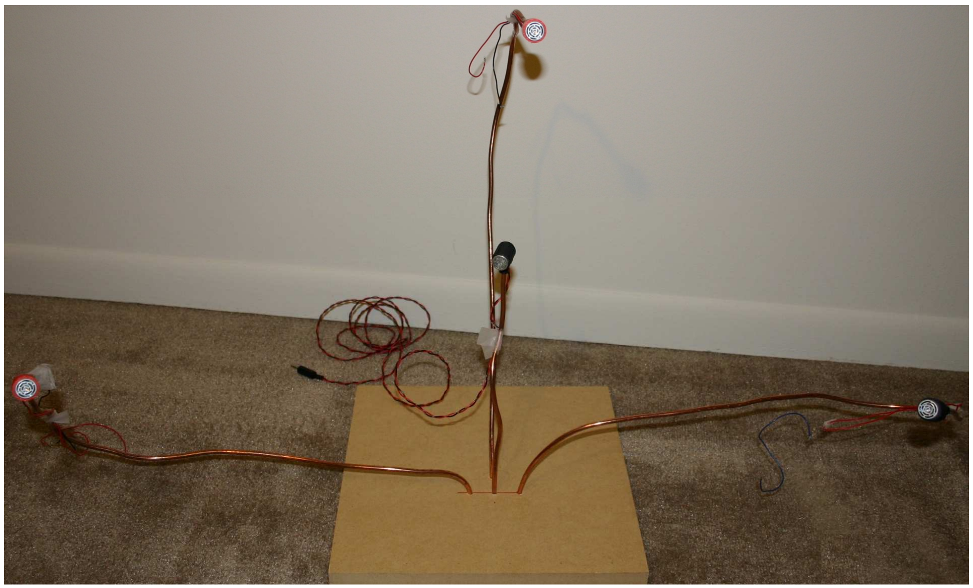


Identifying talking faces..



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

The Gesture Recognizer



Medusa: Our gesture recognizer

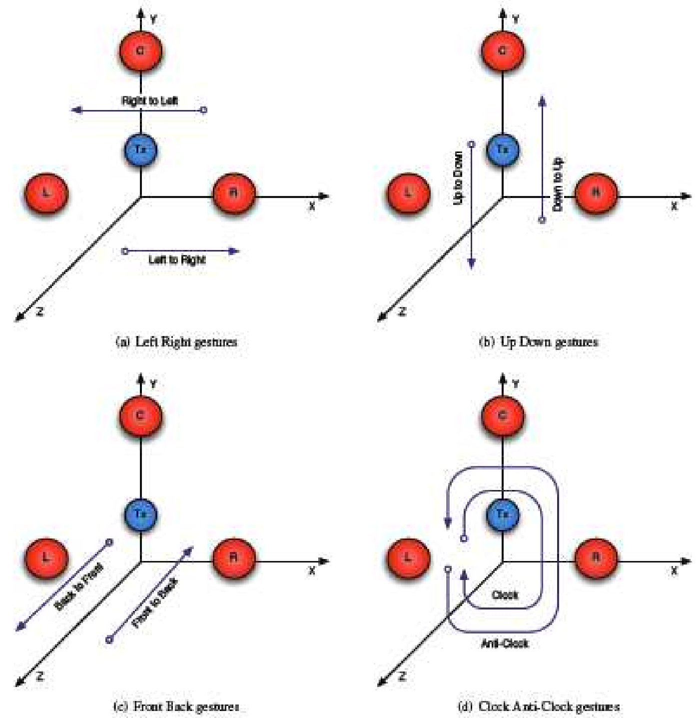
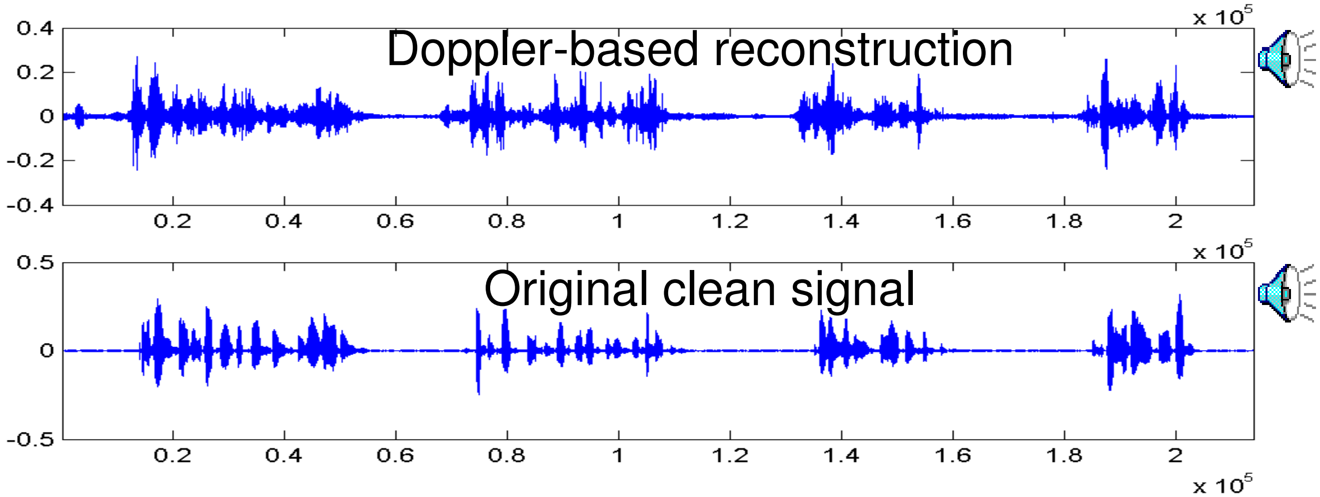


Figure 2. Action Constituting a Gestures

- Gesture recognizer
  - and examples of actions constituting a gesture

Synthesizing speech from ultrasound observations of a talking face



- Subject *mimes* speech, but does not produce any sound
- Can we synthesize understandable speech?

Sound Classification: Identifying Cars / Automobiles from their sound

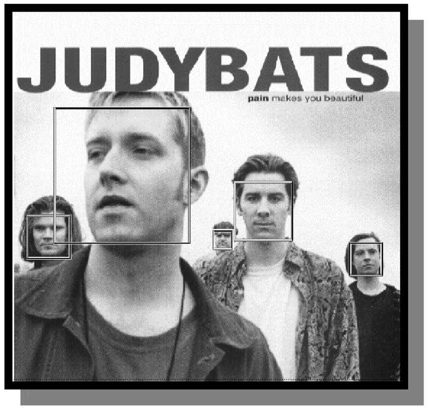
- Sounds are often signatures
- Simple problem: Can we build a system that can identify the make (and possibly model) of a car by listening to it?
  - Can you make out the difference between a V6 and a V8?
    - What do you know of the underlying design that can help?
- Issues:
  - Gathering Training Data
  - Signal Represenation
  - Modeling

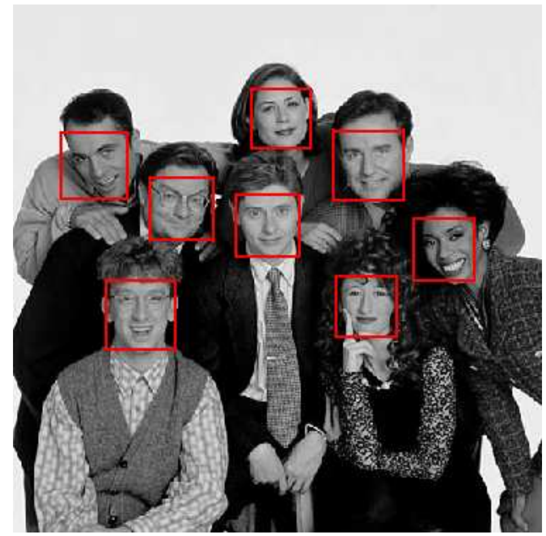




IMAGES

Viola Jones Face Detection





- Boosting-based face detection algorithm
  - State of the art
- Problem: Build a Viola-Jones detector that can detect faces in images
  - Can we also build a classifier that will detect the *pose* (profile or facing) of the face?
  - Can it work from Video?
  - Can we *track* face locations in continuous video

Face Recognition

- Similar to the face detector, but now we want to *recognize* the faces too
  - Who was it who walked by my camera?
- Can use a variety of techniques
  - Boosting, SVMs..
  - Can also combine evidence from an ultrasound sensor
  - Can be combined with face detection..

Recognizing Gender of a Face

- A tough problem
- Similar to face recognition
- How can we detect the gender of a face from the picture?
  - Even humans are bad at this

Image Manipulation: Seam Carving

- See video

- Project
  - Implement Seam Carving
  - Experiment with different ways of eliminating objects without affecting the rest of the image

Image Manipulation: Filling in

- Some objects are often occluded by other objects in an image
- Goal: Search a database of images to find the one that best fills in the occluded region

Image Manipulation: Filling in

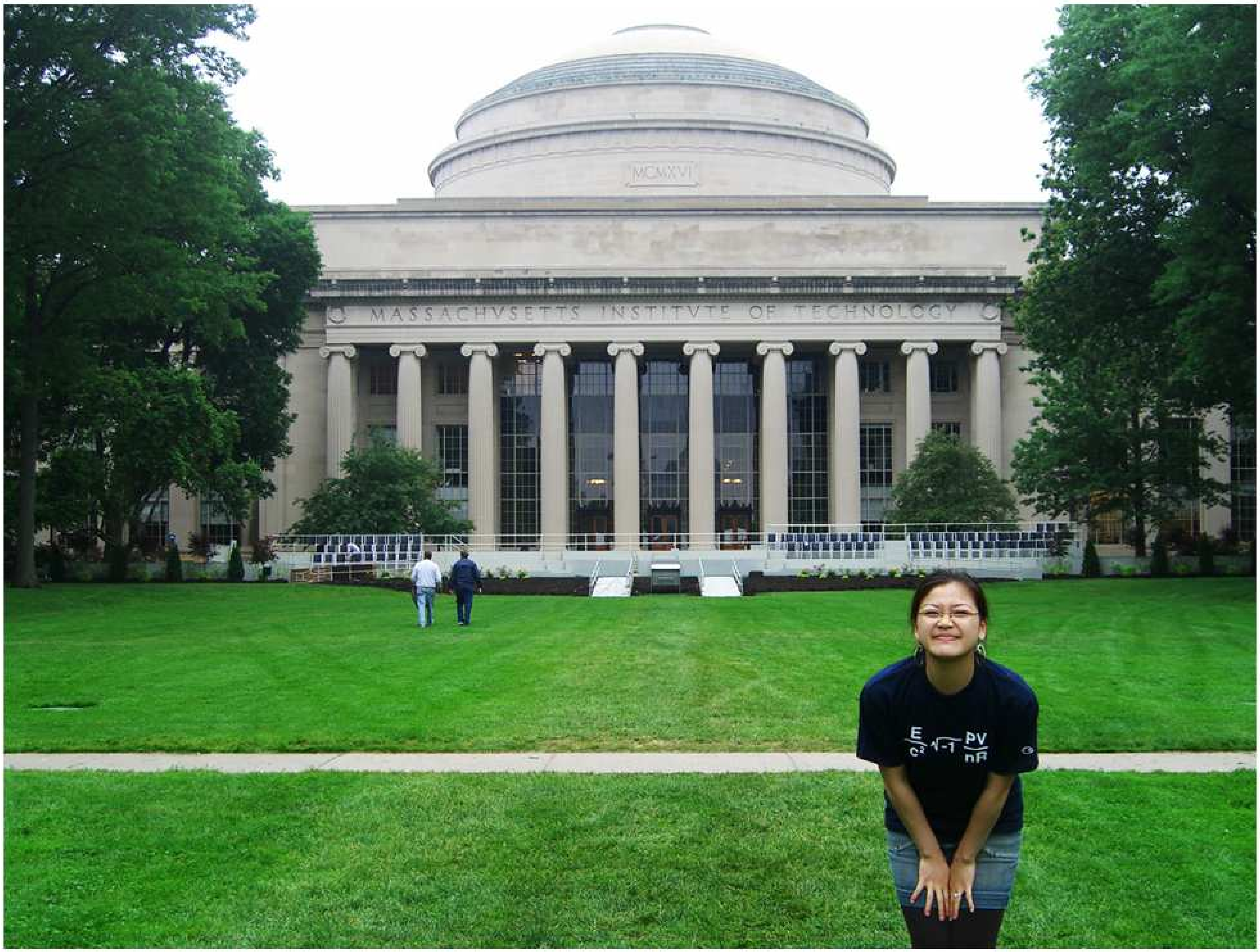
- Some objects are often occluded by other objects in an image
- Goal: Search a database of images to find the one that best fills in the occluded region

Image Manipulation: Modifying images

- Moving objects around
  - "Patch transforms", Cho, Butman, Avidan and Freeman
  - Markov Random Fields with complicated a priori probability models

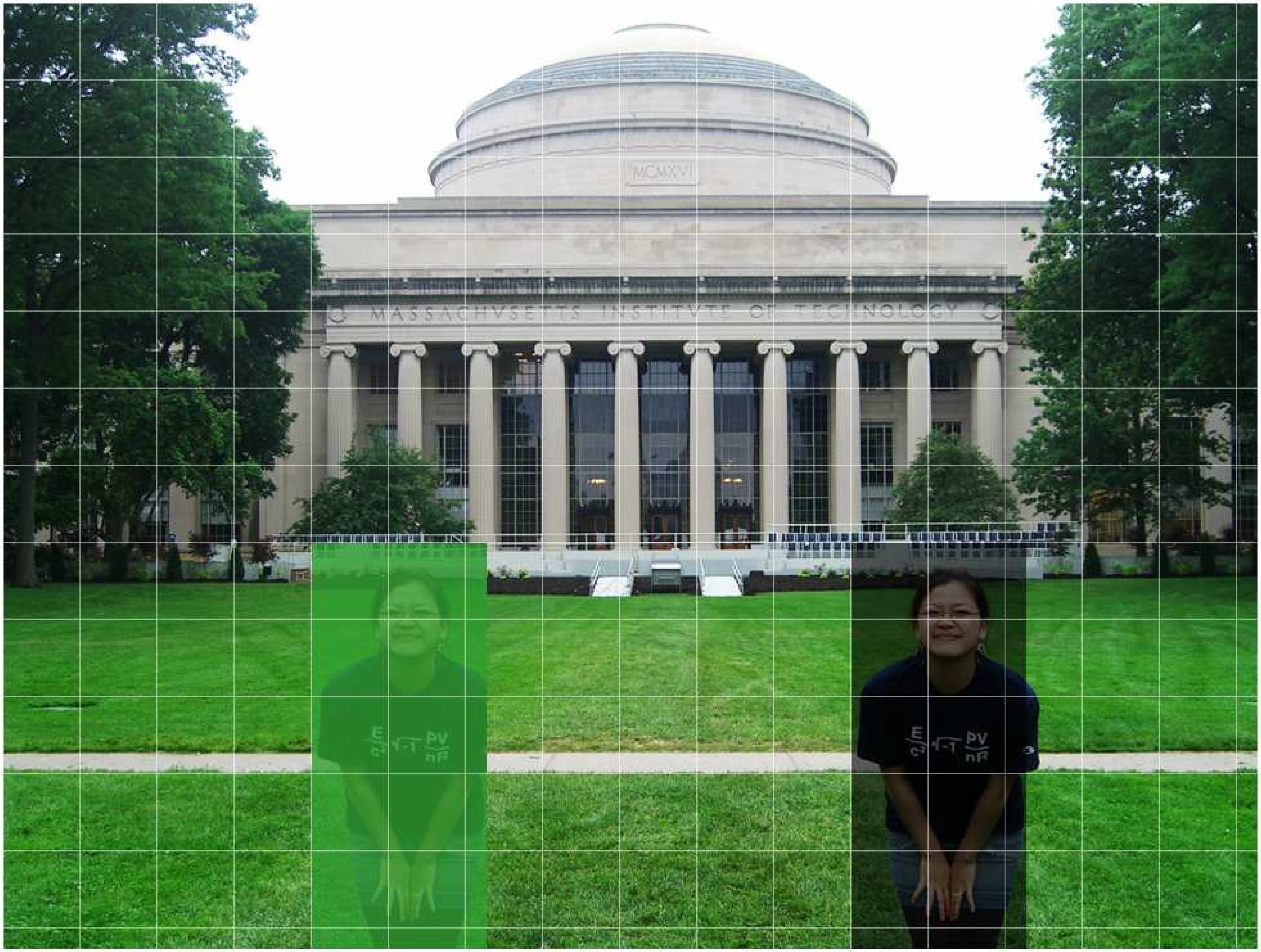
Applications – Subject reorganization

Input image



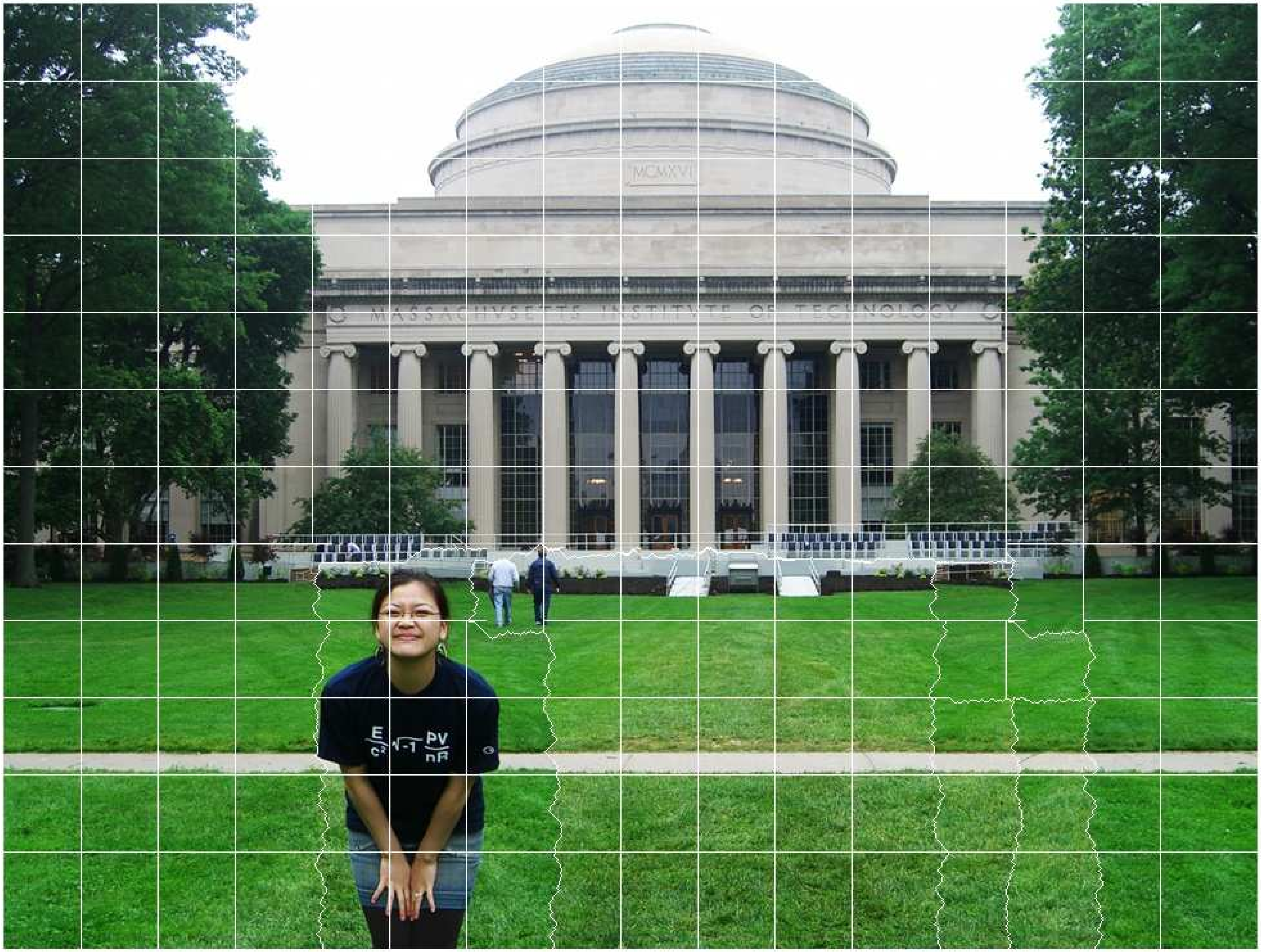
Applications – Subject reorganization

User input



Applications – Subject reorganization

Output with corresponding seams



Applications – Subject reorganization

Output image after Poisson blending

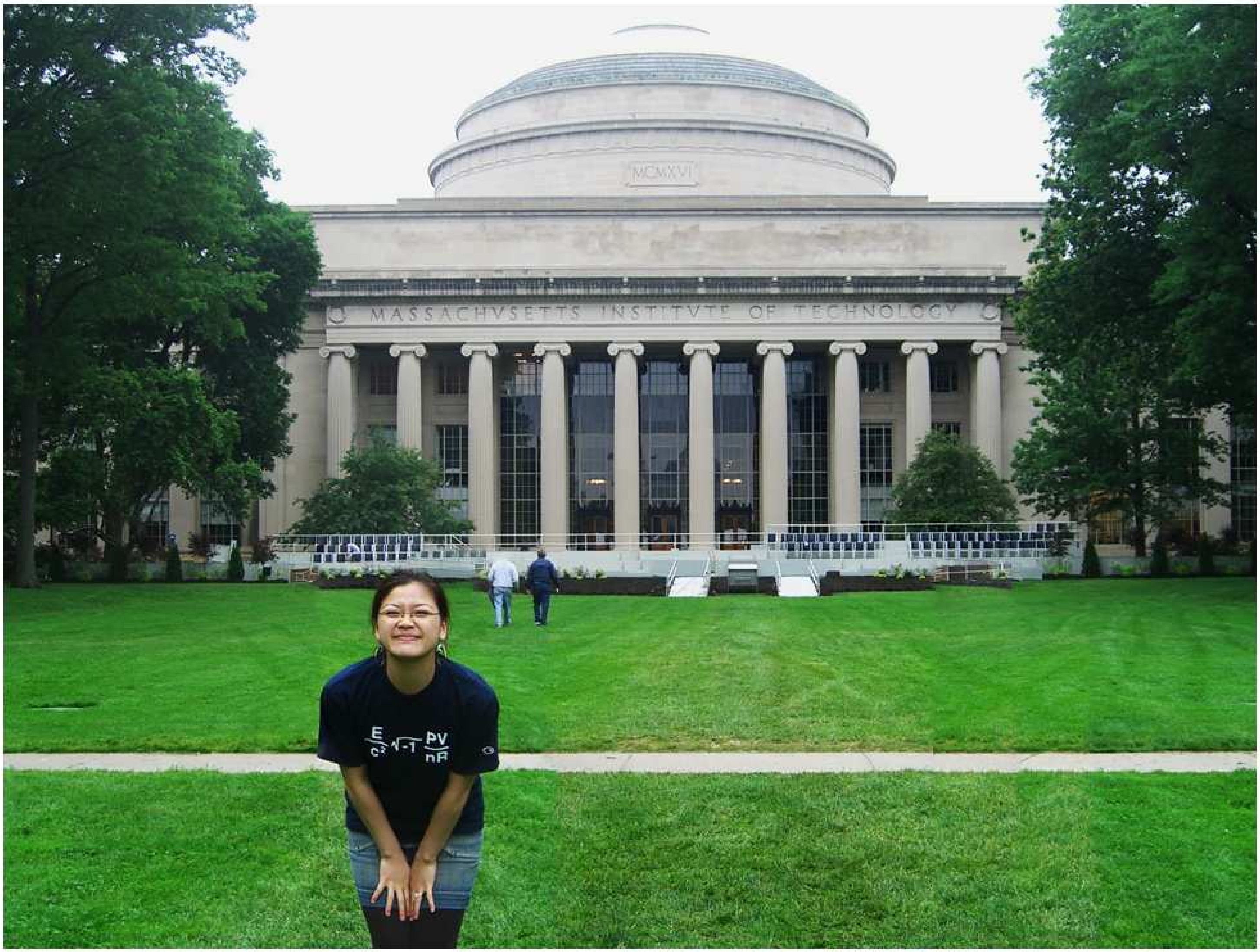


Image Composition

- Structure from Motion:
  - Given several images of the same person under different pose changes build a 3D face model.

Image Composition

- Solving for correspondence across view-point:
  - Given several faces images of the same person across different pose, expression and illumination conditions solve for the correspondence across facial features.
  - The frontal image will be labeled with 66 landmarks.
- Similar to patch models
  - Finding correspondences that match