Voice Controlled Video Console

Ben Horkley and Jonny Surick
6.111, Fall 2013
Voice Recognition is Everywhere

What's the meaning of life?

Try and be nice to people, avoid eating fat, read a good book every now and then, get some walking in, and try to live together in peace and harmony with people of all creeds and nations.

Dance!
Console features

- Camera automatically tracks voice of speaker
- Voice commands control both camera and video display
Video interface

- Displays video from camera with filters
  - Color filters: red, sepia, B/W
  - Blur
  - If time: more complex filters
    - Motion blur
    - Fade in/out

- Receives commands from voice control module
Video interface

- command (4)
- done_in
- raw_video
- Video Filter Logic
- enable
- video_data
- Video Filter 1
- ... (multiple)
- Video Filter N
- xvga_out
Audio tracking

- By default, camera follows audio source
  - Measures differential level on mics mounted on either side of camera
- Can be overridden by voice commands
- Camera is controlled by HS–311 Servo Motor
Audio tracking
Voice control

- Trained to recognize one voice command for each action
  - Feature extraction and comparison on ~1 second samples

- Commands control both video display and camera tracking
  - Select video filter
  - Override automatic tracking to manually move camera
Planned Voice Commands

- 4 Initial Commands For Video
  - “None”
  - “Red”
  - “Black”
  - “Blur”

- 4 Commands For Motor Control
  - “Right”
  - “Left”
  - “Go”
  - “Stop”
Voice control
Timeline

11/17
- Video: working NTSC–RGB conversion
- Tracking: manual servo setting

11/24
- Video: working filters, set from kit
- Tracking: working audio tracking, build assembly
- Voice: software prototype (MATLAB or Python)

By Thanksgiving
- Transfer voice to hardware and integrate

Post–Thanksgiving
- General debugging and finishing
Questions?