Homework #5: Audio Processing

HCI / CprE / ComS 575: Computational Perception
Out: April 13, 2010
Due: May   2, 2010

This homework has two goals: 1) to expose you to audio processing using both C and Matlab; and 2) to expose you to multimodal perception, which, in this case, boils down to combining audio and video processing to achieve a task.

AUDIO

Task 1:   Use Matlab to calculate the spectrogram of the sound snippet stored in the WAV file popcan_drop.wav. Experiment with the number of frequency bins that you use to calculate the spectrogram. Sample Matlab files and the WAV file are posted here:
 http://www.ece.iastate.edu/~alexs/classes/2010_Spring_575/HW/HW5/Matlab/

Post your code and images of the spectrograms on the wiki. Use the provided template.

Task 2: Repeat Task 1, but this time use C to calculate the spectrogram. You will have to download and install one or more of these libraries: OpenAL, ALUT, fftw. Sample C code is posted here (Note: Most likely this works only on Unix/Linux)
http://www.ece.iastate.edu/~alexs/classes/2010_Spring_575/HW/HW5/OpenAL/

Task 3: Write a C (or Matlab) code that processes an audio file and outputs the timestamps of buzzer events relative to the start of the file. You can detect when the buzzer goes off using an algorithm of your choice. A sample audio file is posted here:
http://www.ece.iastate.edu/~alexs/classes/2010_Spring_575/HW/HW5/buzzer/

VIDEO

Task 4: Write an OpenCV (or Matlab) code that uses color tracking to track the tip of the robot's finger as it presses a doorbell button.
http://www.ece.iastate.edu/~alexs/classes/2010_Spring_575/HW/HW5/fingertip/

Your result should be a video file that clearly shows the part of the image that is being tracked during any given frame.

Task 5: Repeat task 4 for all 9 video sequences stored in this dataset:
http://www.ece.iastate.edu/~alexs/classes/2010_Spring_575/HW/HW5/dataset/

MULTIMODAL

Task 6: Combine your solutions to Task 3 and Task 4 and output an image that shows the locations of the robot's finger when the buzzer went off. Plot these locations in red on top of the clear image of the button that is part of the dataset.
http://www.ece.iastate.edu/~alexs/classes/2010_Spring_575/HW/HW5/fingertip/

Task 7: Repeat task 6 for all 9 audio and video sequences stored in this dataset:
http://www.ece.iastate.edu/~alexs/classes/2010_Spring_575/HW/HW5/dataset/