Gesture Detection

A camera is used to detect gestures made in the camera's field of view using infrared LEDs. Detected gestures will be used to control a music player.

Andrew Zhong
Patrick Kuczera
Shahzeb Asif

Motivation

- Create a prototype for a virtual conductor program.
- Gestures are used to control music.
- The program should be able to seamlessly recognize one gesture leading into another.
- Multiple audio streams could be added as an extension that allow the user to conduct them as if they're different orchestra sections.

Functionality

- Gestures are made by moving an LED flashlight in front of a camera.
- A camera is used to track the movement of the flashlight in an image frame and find the location of the flashlight's illuminated point in an image frame.
- The stream of location information is compared to stored gestures to recognize the movement of the LED as a gesture.
- Recognized gestures are used to control music. Gestures are assigned to play, pause, etc.

Hardware

- Altera DE2
- Speakers
- SD Card
- NTSC Camera
- Monitor
- LED Flashlight
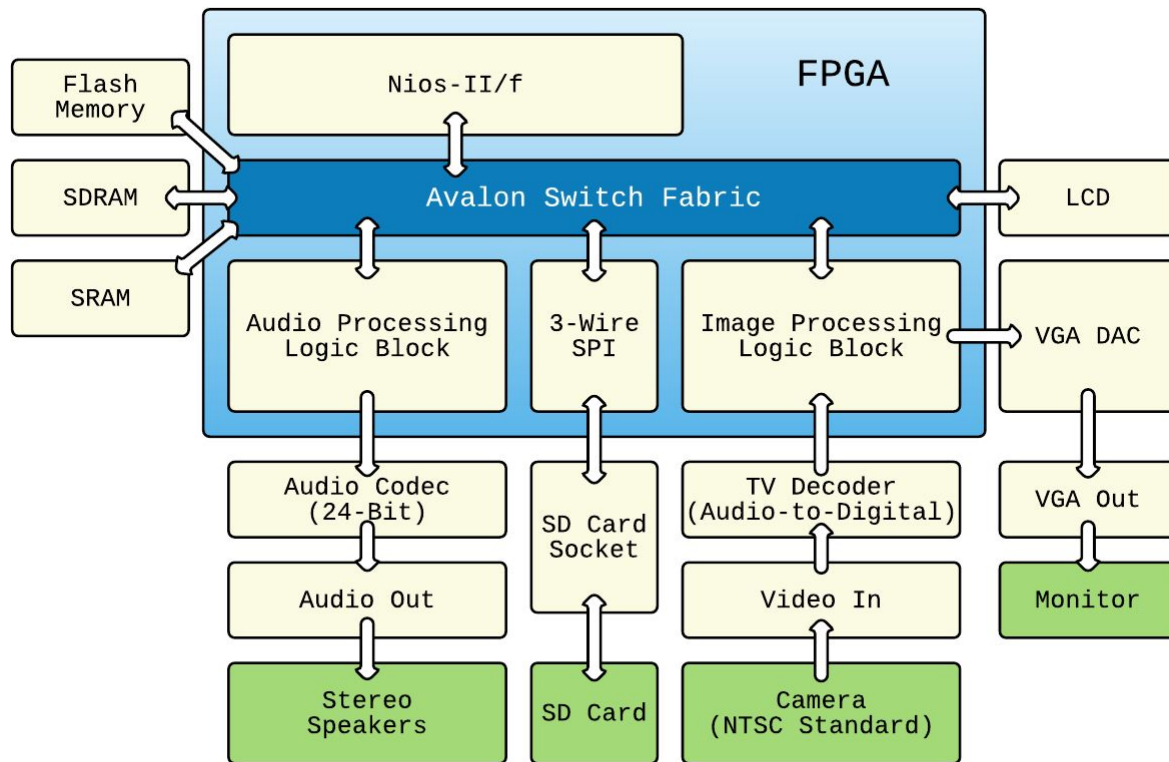
LED Detection

- A frame is streamed into a custom component pixel by pixel from the top left pixel to the bottom right moving right and down in a frame.
- The component thresholds the image using a threshold defined by the hardware switches.
- The longest consecutive sequence of bright pixels found is assumed to be created by the flashlight.
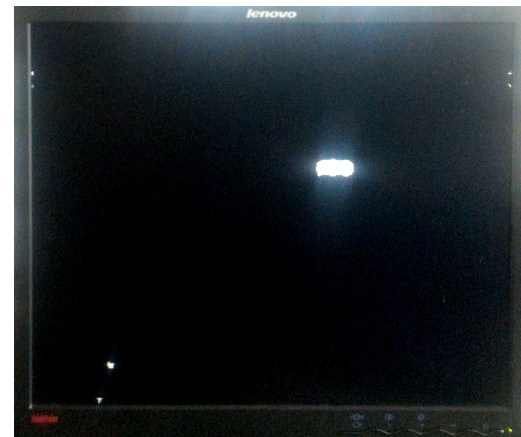
Gesture Recognition

- Gesture recognition will be done using a hybrid trie/graph-like data structure.
- The image frame is converted to a grid and incoming position information is converted to a grid number.
- Comparisons are done between the incoming grid numbers and the stored grid numbers.
- A pointer will be moved from node to node as new position information is streamed in.
  - The **red** path represents the path taken if a gesture is recognized.
  - The **blue** path represents the path taken if a gesture is not recognized.
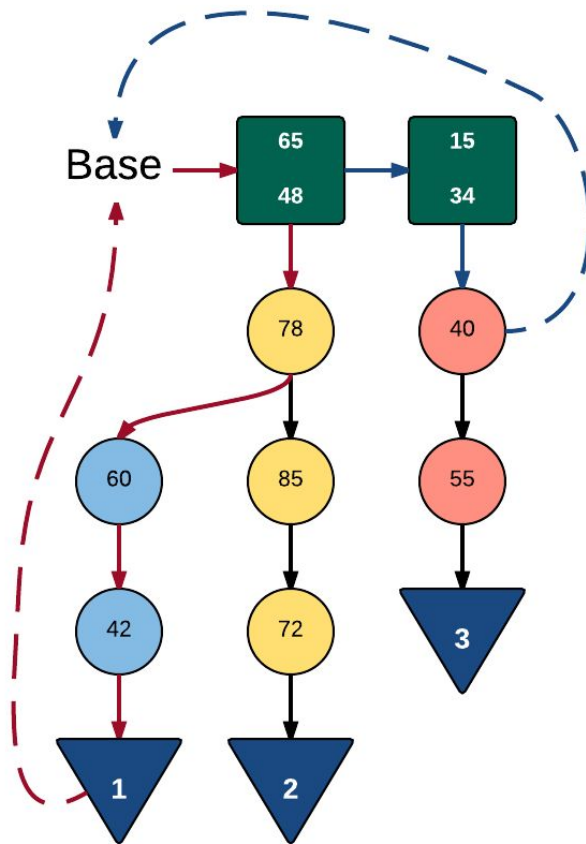
Audio Playback

- Stereo Wav files with a sampling rate of 44.1 kHz are stored on the SD card
- Audio files are found on the SD card and read while being played on the DE2 board
- Audio can be chosen to play, pause, and volume can be increased or decreased
- Pausing and playing audio is done by sending or withholding data from the speakers
- Volume changes are done by the provided Altera functions

Demo

Performance

- The biggest bottleneck was processing input frames.
- To fix this, processing input frames is done in hardware on a stream of data rather than storing anything to memory.
- Frames are processed at around 48 frames per second including gesture recognition.
- Gestures are *compressed* to remove redundant information which allows the frame rate to drop significantly before gesture detection is affected.

Possible Extensions

- Multiple audio streams
  - Add the ability to control multiple audio streams seamlessly
- Custom gestures
  - Add the ability to record and store custom gestures that can be assigned to certain actions.
- Color or object tracking
  - Track color or object shape instead of just the brightness of the LED flashlight.

Questions?