Hand Motion Control of a Media Player

Doris Lin and Diana Cheng

1 Introduction

Recent technology has taken advantage of motion sensors to accomplish basic commands usually implemented through switches or remote controls. The benefits of this motion control allow the user to conveniently command the devices from a distance without the hassle of carrying around a remote control.

There are three aspects to designing the media player. One aspect consists of detecting the user's hand motion and analyzing the gesture to transform it into a command for the media player. Hand motions will be detected by an optical sensor via an LED on the user's hand. The sensor's signals will be processed by the controller, which will then implement such commands as volume control, audio or video playback, audio or video record, rewind, and fast forward according to the detected hand gesture.

The other two parts of the system include handling audio files and handling video files. Overall, the system will function as a media player that can play and record both audio and video files. MP3 audio files can be stored onto a large RAM when the FPGA itself is programmed, or an audio clip can be recorded through a microphone if the user so chooses. In addition, a video file can be stored onto the RAM upon programming the FPGA, or a video clip can be recorded by the user through a camera connected to the system. The basic commands of the media player, such as fast forward and rewind, will apply to both audio and video.

The user will be able to interact with the media player through several devices attached to the labkit, including the optical sensor, a camera, a microphone, a viewing monitor, an LED display, and speakers.

2 Implementation

2.1 User Interface

From the point of view of the user, he/she will see three instruments to change the outputs of the media player and three instruments to view the results. The three inputs to the system include an optical sensor, a camera, and a microphone. By motioning his/her hand in front of the optical sensor, the user has the ability to control volume and playback of the media player. The user will be wearing an LED on his/her hand, which will be detected and tracked by the optical sensor. The sensor then sends this signal to the system and the motion signal is analyzed.

The other two inputs—the camera and microphone are used for recording video and audio, respectively. Record mode for the media player can be signaled with a hand motion.

The three outputs of the media player include an audio output, visual output, and a command display. Attached to the system will be a speaker used to listen to audio clip playback and a monitor used to view stored video clips. The speaker and monitor display the audio or video data that is stored in the memory, but the third output, the LED display will show what command is being implemented by the user.

2.2 Analog Digital Converter

The ADC will interface from the outside world into our system by taking in an analog input from the optical sensor and converting it to a digital output sent to the motion mapping.

2.3 Motion Mapping

This block will take the digital input from the ADC and map the signal to a corresponding audio function, i.e. play, stop, etc. Upon first programming to the FPGA or upon a reset, the user is required to create the table lookup by training this block. Training is needed to teach the converter how to properly recognize each analog signal, since hand motions will vary from the user to user. By running through each of the motions in a predetermined sequence, the processor will map which motions correspond to which commands. Once this is calibrated, the user can mimic the motions to issue instructions to the FSM. The controller receives the output signal of the motion mapping letting the FSM know which command has been requested.

2.4 Controller

This module will contain the main FSM of the machine taking input from the motion mapping that allows the FSM to change states. The input should be a selector signal that indicates which state the FSM switches to based on the command that the user has requested. The FSM is divided into two main sections—audio and video. Each section has its own states of functionality controlling different media devices. The audio and video each contain states, including playback, record, stop, rewind, fast forward, skip forward, skip back, volume up and volume down control. Dividing into two groups will help micromanage the controller as the user can indicate which media type they plan to use. The controller will output a command to the video player/recorder when the user indicates the video section to be accessed. Thus, when the user motions a command, all the instructions are directed to changing the video's states and not the audio states. Similarly if the user specifies audio, then the FSM will only update the audio states and the video states will remain unchanged. The controller also outputs to a simple LED display which will show the sub-states of the system. This requires the hex display of the lab kit in which each command will be displayed in letters. This will help in debugging as we will be able to visually see which command is being requested.

2.5 Loading Audio and Video Files

While a user can record his/her own audio and video clips to be played in the media player, it is also an option to load pre-recorded audio and video files into the RAM upon programming the FPGA. When the audio or video files are passed into the system, they will be decoded by a provided MP3 or video decoder. The decoded files will be stored in the memory and will be called to the data out of the memory when the system is in playback mode.

2.6 Memory

A new memory will be interfaced to the labkit rather than using the onboard ZBT RAM. In order to store video clips, a larger memory is needed. A memory chip that can store up to 10 seconds of video would be ideal. The memory can be partitioned to store both video and audio data if not all the memory required to store 10 seconds of video is used.

In order to interface the memory to communicate with the other elements on the labkit, it will be necessary to consult the memory chip's data sheets. The data sheets will provide timing specifications necessary to interface the labkit to the off board RAM. The data sheets will also provide the framework of the memory chip, such as inputs, outputs, and the purpose of each.

2.7 Video Player and Recorder

The video player and recorder will receive inputs from the controller block and video in from the system's video camera. With these inputs, the video player can either read from the memory and playback the data out from the RAM, or it can send data into the RAM during record mode when write enable is high. The video player will also send the address to the RAM that will specify which location to read from or to write to. In playback mode, the data read from the RAM will be shown on the video display—the computer monitor. In record mode, the video in from the camera rather than data from the memory will be fed to the video output.

The command input from the controller block, besides deciding playback or record mode, will also determine how the video signal is outputted to the monitor. For instance, if in playback mode the controller orders the player to fast forward, the video player will increase the speed of its output in the forward direction. If rewind, the video player will increase the speed of its output in a backward direction.

2.8 Audio Player and Recorder

The audio player module is analogous to the video player except used to play audio clips. It, too, receives a command input from the controller block and an audio input from the microphone attached to the system. In playback mode, the audio player reads in audio data from the memory chip and plays it through the speakers. In record mode, the audio player asserts write enable to write new audio data to the memory. While recording, the speakers will output the audio feed from the microphone rather than from the memory.

The command input from the controller block will further specify the behavior of this module. For instance, the command input may signal volume to increase or decrease. In this case, the audio output of the module is multiplied by a factor greater than one to increase the volume or a factor less than one to lower the volume. The user will then hear a difference in the output volume through the speaker.

3 Testing

During the construction of our system, the media player will be tested segment by segment. When all the modules have been wired together, the system can be tested and debugged by analyzing its outputs since outputs are either audio or visual.

Before putting all the parts together, the separate aspects of the media player will be tested individually. The audio portion of the system can first be tested by loading an audio file and executing a basic playback of the file. This test ensures that the audio player can correctly read from the memory and output correctly to the speakers. The same check can be done for the video player. The next step will be to test the record mode of first the audio player and then the video player. Tests can thus be divided into those for the audio path and those for the video path, and within each of those paths, there will be tests to check both playback and record modes.

The third part of the media player to test is the motion tracker. Simple motions can first be tested to investigate if the system is correctly processing the hand gesture. Once the motion tracker has been created and motions mapped to commands, the LED display can aid in debugging by displaying what command the system thinks it is carrying out. If the LED display shows the correct command, the hand motion is correctly being processed and mapped by the system.

Besides the LED display, the monitor and speakers provide a sensory check of the media player. The monitor allows a visual check while the speakers provide an auditory check. When all parts of the system have been put together, each of the calibrated hand motions will be carried out in front of the sensor. If the hand motion is correctly mapped to its corresponding command, besides seeing the command on the LED display, we should also be able to hear or see the command being implemented (i.e. volume goes up or down, displayed media skips to a different one). If the audio or video playback does not properly respond to the hand motion, there may be communication problems between the controller and the player/recorder blocks. Another possible problem could be switching between the audio and video players.

By testing each of the possible hand motions and analyzing the outputs, we will be able to see if the system is functioning properly.

4 Work Plan

Both partners will contribute to each of the parts of the system. However, Diana will focus primarily on the audio aspects of the system in addition to processing the motion detected by the sensor. This includes the audio player/recorder, analog-to-digital converter, controller, and motion mapping blocks. Doris will focus more on the video portions of the system and interfacing a larger memory to the labkit. These portions include the video player/recorder and memory blocks. In order to connect all the parts, some work on the controller and memory will be required of both people.

Hand-Controlled Media Player Block Diagram