UNIVERSITY OF FLORIDA

Preliminary Design Report

Theremania - Theremin Vocoder

1/25/2022
Eric Barkuloo, Jackson Cornell

Abstract

The *Theremania* is a modern take on the theremin, an instrument popularized in music throughout the 20th century. The *Theremania* will involve a classic theremin, with the pitch of the instrument being controlled by the hands using an antenna. The output of this theremin will be modified by a digital signal processing stage, in which the pitch can be modified by vocal input through a microphone into a phase vocoder. Additionally, in this stage there will be options to change the signal note output of the theremin into more complex musical outputs such as chords and harmonies. In addition to this functionality, the *Theremania* will also feature conversion of the theremin output to a MIDI signal, allowing the theremin to be used as a virtual instrument controlled in any audio software allowing MIDI as an input.

The *Theremania* presents unique technical challenges, particularly regarding the signal processing required in order to provide all the planned project features. The modification of the signal via a phase vocoder, the conversion of the control voltage signal to MIDI, and the inclusion of other pitch shifting audio effects create a technically challenging project, in addition to the task of designing a working version of the original theremin instrument. The *Theremania* promises to create a flexible instrument with many creative capabilities, and provide an update to a classic instrument to modernize it and increase its usability.

The *Theremania* finds applications in the consumer musical instrument market, targeted towards professional musicians and hobbyists alike. The *Theremania* tackles two unique challenges in this market: a theremin MIDI controller for virtual instruments, and a unique vocoding effect on the human voice. There are currently no products that tackle either of these challenges.

Due to the unique nature of this effect, the project has a great potential to be utilized in both professional and hobbyist communities. The synth community in particular is a large consumer of similar instruments and effects, and can be attached to both analog and digital rigs. The *Theremania* will also be able to be utilized in a live setting, which will make it a prime candidate for performers looking for a unique sound.

Project Features and Objectives

The core of the *Theremania* will involve the design of an analog Theremin circuit that allows for the user to control pitch using the distance of their hand from a metal plate. Additionally, the volume of the Theremin should be controlled using a potentiometer. For this analog component, we will use a Minimum Theremin design that allows for the creation of a cheap, low-noise system. The user controlled analog Theremin output signal will then be passed into a TLV320AIC23BPW audio codec to handle the analog to digital conversion for the DSP chip, as well as being passed into the FPGA’s analog to digital converter to handle the pitch detection.

We will use a TMS320F28379DZWTT 2 core DSP chip to handle the majority of the digital audio processing for the *Theremania*. This processing will involve running the phase vocoder algorithm on the first processor to accomplish a harmonizer effect on the monophonic theremin input, as well as an auto tuner effect that will match vocal input to the Theremin pitch input. These effects will be accomplished in real time, and keeping effect quality while maintaining real time deadlines is a key challenge of this project.

The DSP will also use its second processor to continuously receive pitch information from the FPGA via a SPI communication line. This pitch information will then be transferred to the first processor and used by the phase vocoder algorithm to achieve the vocal auto tuning effect. Additionally, the second processor will handle taking in the 4 different control signals. This processor will also use the pitch information and volume information to generate a MIDI output.

We will use a 10M50DAF484C7G FPGA to handle the pitch detection algorithm. The FPGA will sample the incoming Theremin and audio signals, calculates the pitches using the Circular Average Magnitude Difference Function (CAMDF), finds the stretch factor between the two pitches, and sends this information to the DSP via SPI.

Technology

Phase Vocoder Algorithm

The Phase Vocoder Algorithm is a type of signal processing algorithm that uses frequency domain phase information extraction and modification of a signal combined with interpolation in the time domain to reconstruct a signal to its original size, while maintaining its modified features. In the algorithm a signal is first broken into frames of a set size. These frames are each processed by first applying a window function (a hanning window in our case), then applying an FFT to bring the frames into the frequency domain. In the frequency domain, the phase is extracted and modified, and the frame is brought back into the time domain using the IFFT. In the time domain, interpolation is used to change the frame length so the modified frames can be overlapped and added together to reconstruct the now modified signal. This algorithm can be used to accomplish a high quality low noise pitch shifting effect, which is how we will be using it in our project.

The DSP chip will interface with the codec using the Multichannel Buffer Serial Port (McBSP) protocol, and will read audio samples into two rotating sample buffers. Our software design approach will allow one buffer to be continuously fed new samples from the codec while the other buffer is being processed by the phase vocoder algorithm. This schema will minimize samples being dropped and allow for the most efficient real time processing. The phase vocoder code will utilize the Fast Fourier Transform and Inverse Fast Fourier Transform (FFT and IFFT) algorithms supplied by TI in order to maximize code efficiency.

Pitch Detection Algorithm

Our design utilizes the Circular Average Magnitude Difference Function (CAMDF) for calculating the pitch of the theremin and vocal audio signals. The CAMDF is a modified version of the autocorrelation function (ACF). The ACF is a useful algorithm for calculating pitch (among other properties) due to the fact that it has the same period as the signal it is being computed on. The CAMDF is useful in this context for a variety of reasons:

- It is symmetric, meaning only the first half of samples has to be computed
- It uses only subtraction and addition, rather than the ACF which uses multiplication
- It is robust against the "Double Pitch" error, where integer multiples of the actual pitch are computed, which can occur when in the presence of noise

The CAMDF can be described with the equation :

$$R(\tau) \; = \; \sum_{n=0}^{N-1} |x(mod(n \; + \; \tau, N)) \; - \; x(n)|$$

From there, the minimum value of $R(\tau)$ is found, which corresponds to the period of the signal's pitch.

The main downside of the CAMDF is its time complexity of O(N^2). A naive approach to pitch detection may use the FFT, which has a time complexity of O(Nlog(N)). However, this approach is not robust to noise or even harmonics, and thus would perform poorly in our application.

FPGA

Due to the large time complexity of the CAMDF algorithm, dedicated digital hardware is necessary to run it in real-time. Altera's 10M50DAF484C7G FPGA was chosen due to its internal ADC, space for hardware, and availability. The FPGA will use its internal ADC to sample the theremin and audio signals. These samples will be written to a buffer, which will feed into a pipelined datapath which will compute the CAMDF output $R(\tau)$ for the first N/2 values. The minimum value for $R(\tau)$ is found for both signals, computed into a pitch, and is then sent over SPI when requested from the DSP.

DSP

The TMS320F28379DZWTT DSP chip was selected for use due to its processing speed, dual core architecture, and compatibility with the TLV320AIC23BPW codec. The first core will be used exclusively for the phase vocoder algorithm described above. This design decision was made to maximize the processing power reserved for the audio processing. The second core will be used for several purposes. First, the processor will be used to take in several analog user control parameters (Volume, Harmonizer pitch control, and Mix) via 3 different ADC's. Second, the processor will interface with a digital rotary encoder to select the effect. Third, the processor will use SPI to communicate with the FPGA to receive pitch information in real time. The processor will communicate the user control parameters and the pitch information with the first processor for use in the phase vocoder algorithm. Finally, this processor will convert the pitch and volume information into a MIDI output via UART. This will be output to a standard MIDI connection, but if time allows we would like to send this MIDI output to USB as this is a more common modern MIDI connection.

Theremin

Being the oldest electronic instrument, there are many different implementations of an analog theremin. The simplest ones have only one antenna for pitch control, and instead uses a potentiometer to adjust the volume. The design we ended up using is based on the 2006 Minimum Theremin design by Harrison Instruments. This design was chosen because:

- It is a popular hobby build, and so there are many examples on the web to demonstrate the quality of audio it produces
- The oscillator circuit is fairly simple
- It uses relatively few parts

A theremin works by using the capacitance between an antenna and a player's hand to drive an oscillator circuit. The 2006 Minimum Theremin uses a transistor-based oscillator circuit, which has a reduced complexity compared to many other theremin oscillator circuits. The 2006 Minimum Theremin has two potentiometers for adjusting the pitch of the theremin, one to be tuned by the manufacturer and one that can be tuned by the player. The circuit was modified by us to include an additional volume control.

Pre-Amps

Pre-Amps are used to filter and amplify the incoming theremin and audio signals before being sampled by the FPGA's ADC. The ADC will be sampling at 4 kHz, so frequencies above 2 kHz must be filtered from the signals. Additionally, the signals must be offset by 2.5V and amplified so that the ADC gets a fine enough reading. This was accomplished by both signals having a two stage analog filter, where one stage amplifies and offsets the signal and the other stage being a low pass filter with a cutoff at 2 kHz. The LT1632 op amp chips were used due to being cheap, available, and low noise.

Speaker

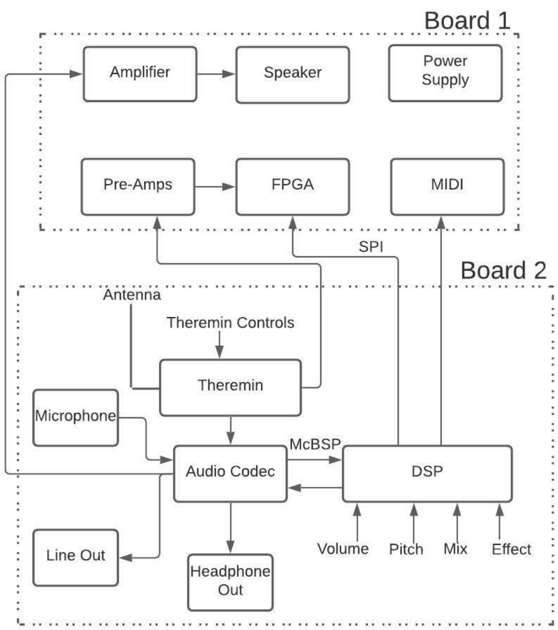
The line-out output from the codec will be amplified using an LM386 audio power amplifier to drive an 8 ohm speaker.

Power Supply

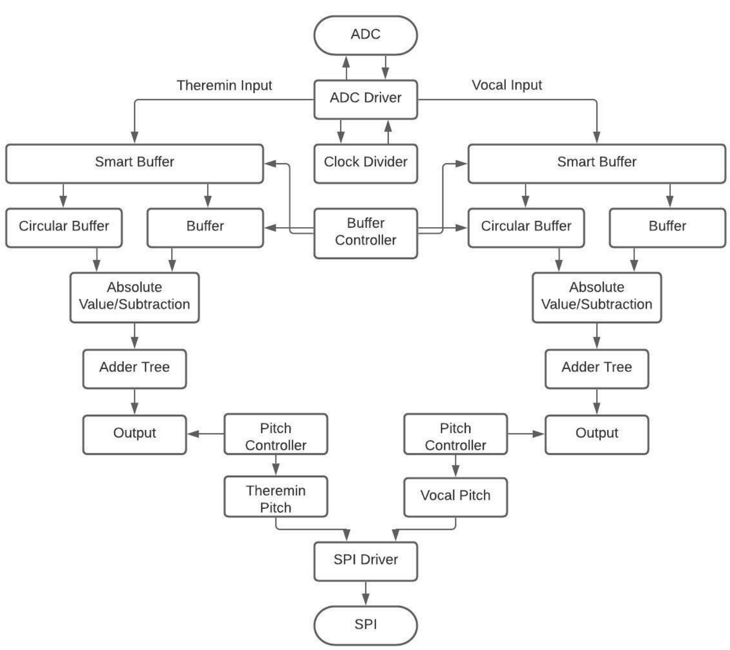
Power will be supplied by a 9V wall wart adapter standard with guitar pedals and other effects modules. Voltage regulators will be used to step the voltage down to 3.3 V and 5 V. Ferrite beads will be used to filter and isolate analog power sources from digital ones. All ICs will have bypass capacitors to stop power ripples from affecting their performance.

Flowcharts

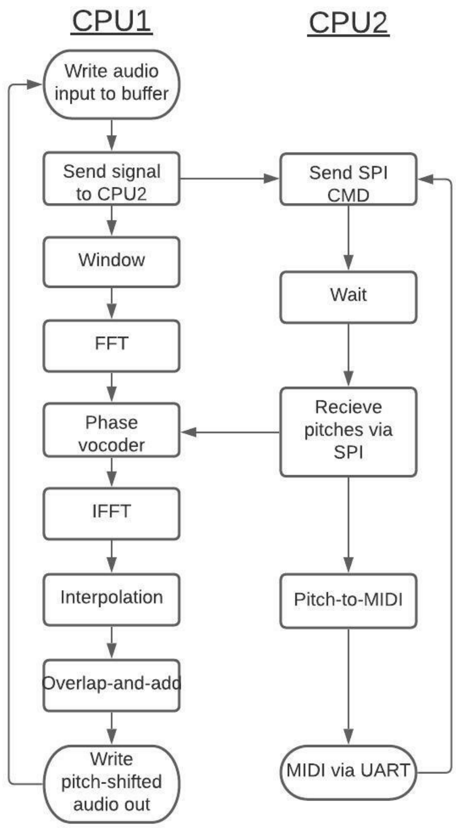
Hardware Block Diagram



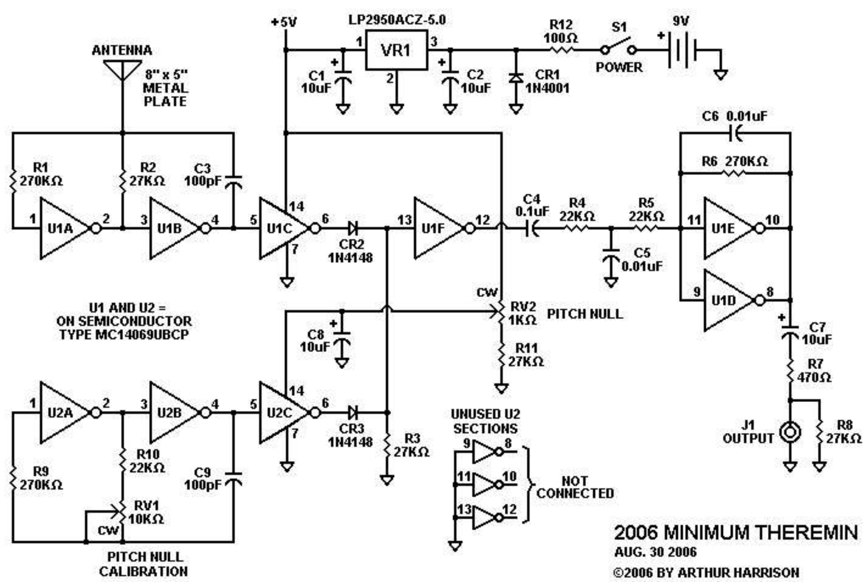
FPGA Block Diagram



Software Block Diagram



Original 2006 Minimum Theremin Circuit



Project Plan

Roles

Domain	Jackson	Eric
Hardware	- PCB 1 (theremin, codec, DSP, peripheral controls) - FPGA Design	- PCB 2 (FPGA, pre-Amp, speaker, power supply) - DSP
Software	- Algorithm prototyping - VHDL (pitch detection)	- Firmware Design - C (Phase Vocoder, MIDI output, User inputs)
Assembly	- PCB 1 - Enclosure	- PCB 2 - Enclosure

Gantt Chart