Real-Time Time-Domain Pitch Tracking Using Wavelets

Ross Maddox and Eric Larson

August 3rd, 2005
NSF REU Program
UIUC Physics
Prof. Steve Errede, Advisor
Physics of Music 2005 Projects

- Development of a PC/104-based single board computer DAQ system
- Investigation of non-linear properties of second sound in superfluid LHe
- Measurement of the mechanical resonances of differently shaped ‘ukuleles
- Real-time time-domain pitch tracking of monophonic musical signals using wavelets
PC/104-based SBC DAQ system

Using the Diamond Systems “Athena” single-board computer:
PC/104-based SBC DAQ system

Goals:
- Develop a reliable system for intense, long-term data acquisition
- Apply system to several different existing experiments to verify proper operation, as well as improve their functionality and reliability
- Pave the way for brave new experiments!
PC/104-based SBC DAQ system

Athena DAQ hardware:
- 16 16-bit analog-digital converters
- 4 12-bit digital-analog converters
- 24 TTL (digital) I/O lines
- No fan!
PC/104-based SBC DAQ system

Windows XP Embedded allows us to

- Select individual OS components
- Cut the crap
- Run existing DAQ programs, with limited adaptations for Athena
Physics of Music 2005 Projects

- Development of a PC/104-based single board computer DAQ system
- Investigation of non-linear properties of second sound in superfluid LHe
- Measurement of the mechanical resonances of differently shaped ‘ukuleles
- Real-time time-domain pitch tracking of monophonic musical signals using wavelets
Mechanical Resonances in ‘Ukuleles

- Different ‘ukuleles have different sounds
- This is a result of varying body
  - Shapes
  - Sizes
  - Materials
  - Finishes
Kanile’a Koa Concert ‘Ukulele

Kanile’a Concert Ukelele
|V| vs. Frequency

Frequency (Hz)

Open Strings
Muted Strings
Kanile’a Koa Concert ‘Ukulele

Im(V) vs. Re(V)

Open Strings
Muted Strings
X-Axis
Y-Axis
Painted Chinese Soprano ‘Ukulele

Soprano Ukelele

|V| vs. Frequency

Frequency (Hz)

Open Strings  Muted Strings
Painted Chinese Soprano ‘Ukulele

Soprano Ukulele
Im(V) vs. Re(V)
Don Tomás Cigar Box ‘Ukulele

Kanile’a Concert Ukulele
|V| vs. Frequency

Frequencies (Hz)
Don Tomás Cigar Box ‘Ukulele

Kanile’a Concert Ukelele

Im(V) vs. Re(V)
Why measure?

- Useful for synthetically modeling the sounds
- Gives a deeper understanding of why the instruments sound the way they do and allows for prediction of how an instrument might sound based on its construction
Physics of Music 2005 Projects

- Development of a PC/104-based single board computer DAQ system
- Investigation of non-linear properties of second sound in superfluid LHe
- Measurement of the mechanical resonances of differently shaped ‘ukuleles
- Real-time time-domain pitch tracking of monophonic musical signals using wavelets
Why Pitch Tracking?

- Pitch information is used in music and speech analysis
- Fourier analysis / Spectrogram is vague
- Pitch tracking is accurate and easy to read
Project Goals

Develop software to track the pitch of a monophonic (one note at a time) audio signal that is:

- Accurate (within a quarter-tone, or ~1.5%)
- Low Latency (~25 ms)
- Capable of high frequency and time resolution
- Accurate voiced/unvoiced detection with no global amplitude thresholds
- Robust to:
  - Noise
  - Weak fundamentals
  - Sliding pitches within a window
Our Approach

- Our eyes can readily see periodicity (or lack thereof), so should the computer...
- Implementation of fast, intelligent peak (and valley) finder coupled with simple wavelet signal manipulation
- Use found peaks to calculate period and frequency
Fast Lifting Wavelets

- Transform that separates a signal into a tree of approximations and details
- Separation can be based on several different wavelet shapes
- Our implementation uses the Haar wavelet
Wavelet Transform Levels

Original Signal

Approximation Level 1

Approximation Level 2

Approximation Level 3
Our Algorithm

Window the signal into $N$ 25-ms segments. $\forall n \in [0, N)$:

1. At level $i$ of a Haar wavelet transform of the data:
   a) Remove DC component
   b) Find all zero-crossings
   c) Find first local max after each zero-crossing that is above a certain percentage of the window max
   d) Calculate the distances (in samples) between local maxima
   e) Repeat b), c), d) for local minima
   f) Determine the averaged mode distance

2. If the mode distance at level $i \approx$ the mode distance at level $(i - 1)$, assume the distance $\approx$ period, yielding frequency;
   If not, go to the next wavelet level (up to $i = 6$) and return to step 1
Algorithm

First, the signal is windowed:
Next, the DC offset is accounted for (used in thresholding)
Algorithm

- The first wavelet transform is applied to generate an approximation.
Algorithm

- The first local maximum (minimum) between zero crossings above (below) a threshold is located.
- The mode distance between extrema is calculated.

![Graph showing a data approximation with labeled axes and values.](attachment:graph.png)
Algorithm

The next wavelet approximation is generated
Algorithm

- The maxima and minima are located
- The mode distance between extrema is compared to that of the previous level
Algorithm

- Since the mode distances matched, the averaged mode distance of the previous level is taken to be the period, yielding frequency
Mode Averaging Scheme

On sine wave signals (see right), averaging the mode yields far better accuracy than simply taking the statistical mode.
Results: Latency

- Using a window of 1024 samples at 44100 Hz implies ~22 ms minimum latency
- Computation time averaged only 4 ms in MATLAB
- Total pitch tracking latency: < 26 ms
- C++ implementation will be even faster, approaching the minimum of 22 ms
Results: Voiced/Unvoiced

- A voiced/unvoiced error rate of approximately 0.4% in a female vocal sample
- Comparable to established methods, but leaves room for improvement
Results: Accuracy

- A sinusoid-based test yielded favorable results
- Almost constant, but completely inaudible, error in terms of musical intervals
Conclusions

- Our algorithm compares favorably to established methods, and allows for real-time pitch tracking with low latency and high accuracy
- Further work could be done to improve voiced/unvoiced detection
Acknowledgments

We would like to thank:

- Dr. Steve Errede
- Matthew Winkler
- Jack Boparai
- Dr. James Beauchamp and Mert Bay
- UIUC Physics Dept.
Physics of Music 2005 Projects

- Development of a PC/104-based single board computer DAQ system
- Investigation of non-linear properties of second sound in superfluid LHe
- Measurement of the mechanical resonances of differently shaped ‘ukuleles
- Real-time time-domain pitch tracking of monophonic musical signals using wavelets
Questions?

- Development of a PC/104-based single board computer DAQ system
- Investigation of non-linear properties of second sound in superfluid LHe
- Measurement of the mechanical resonances of differently shaped `ukuleles
- Real-time Time-Domain Pitch-Tracking Using Wavelets