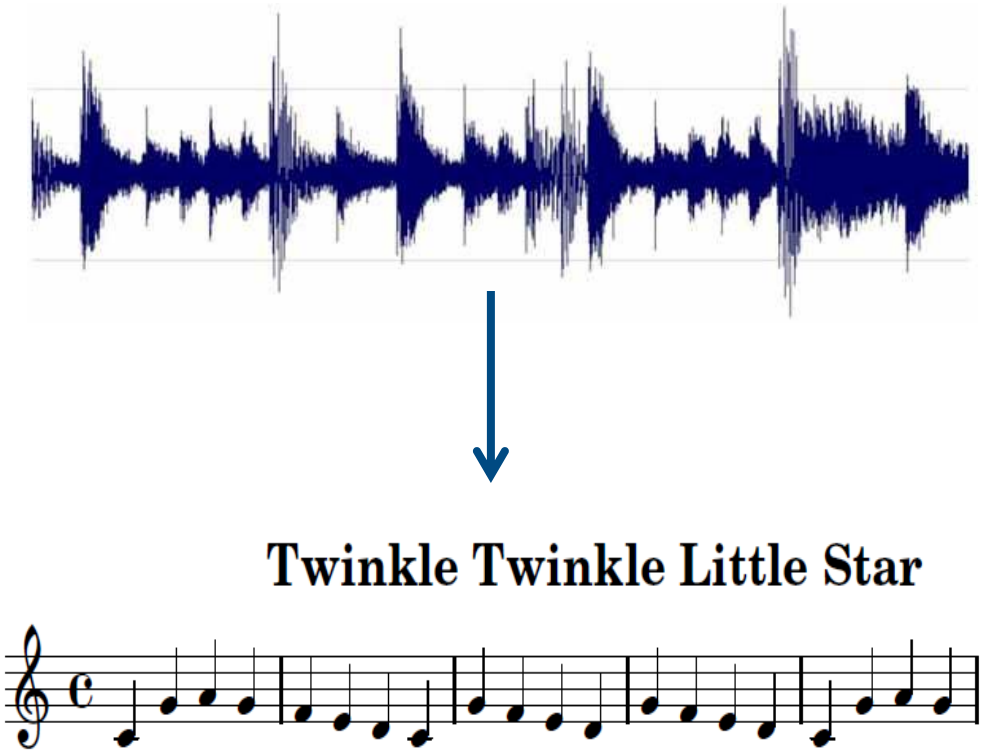
Polyphonic Music Transcription on FPGA

Shuyi Chen
Lizi George
Kelly Ran

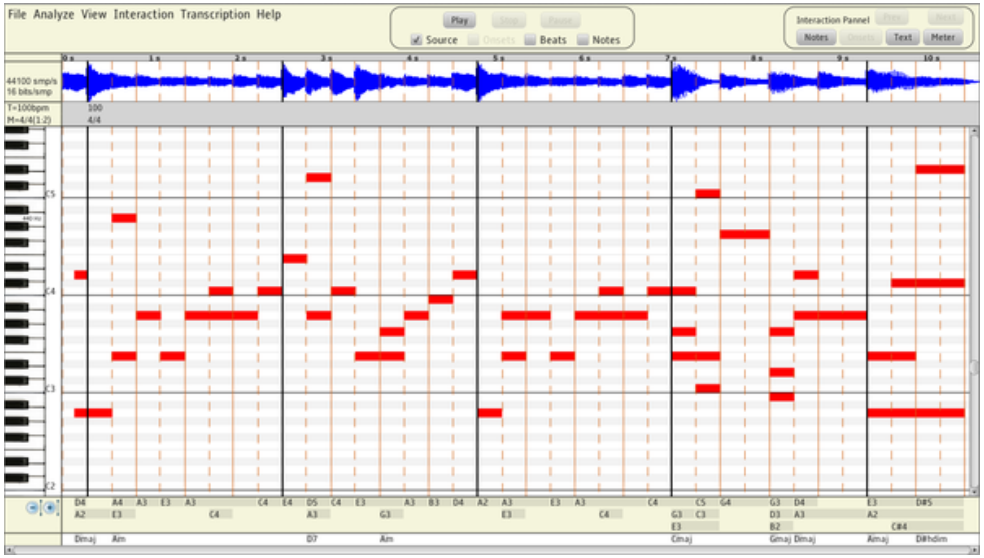
Outline

- Motivation
- System
- Architecture
- Evaluation
- Exploration



Motivation

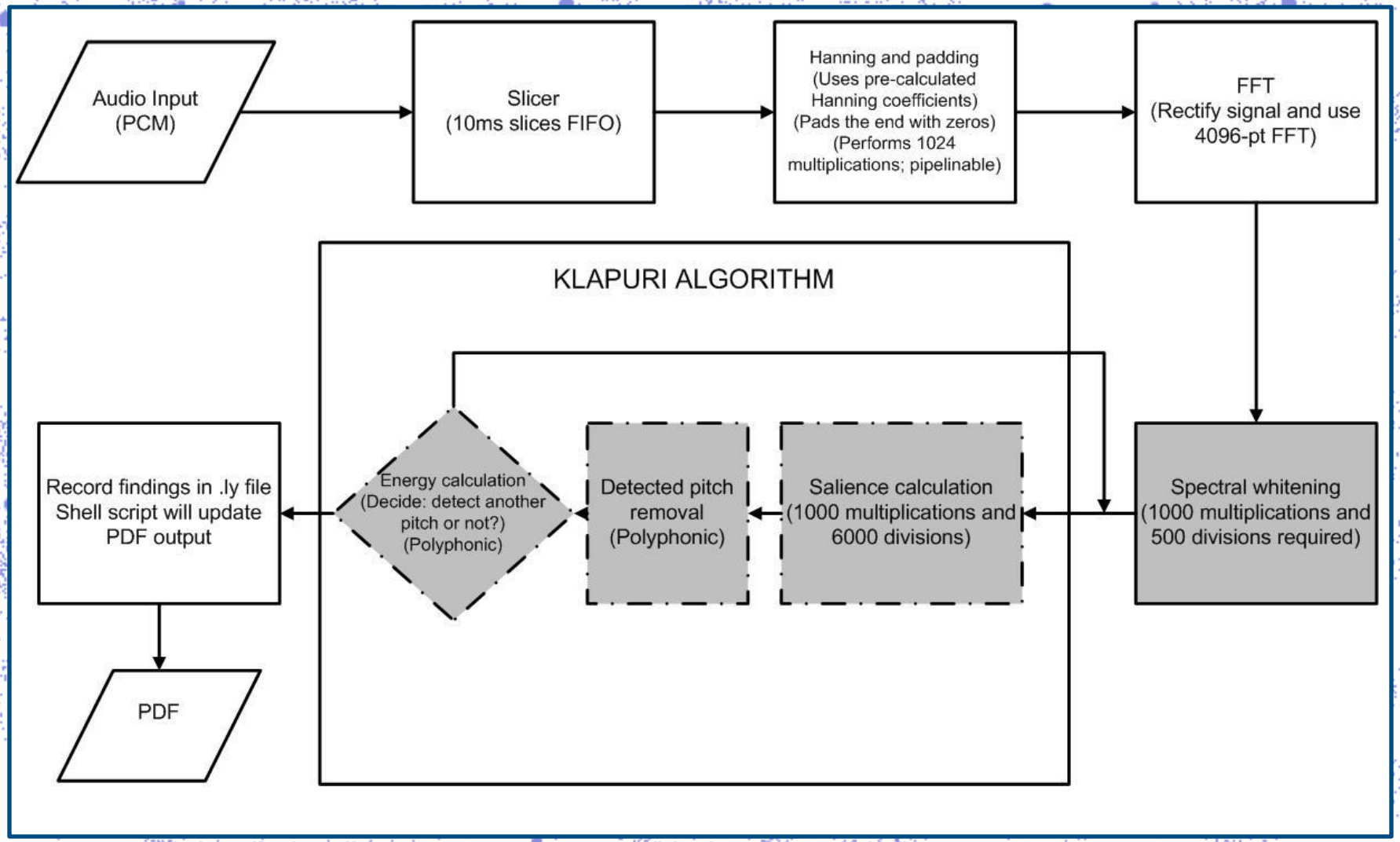
- Speed
- Polyphonic detection methods are still being investigated
  - Using techniques from growing field of speech recognition
- Real-time transcription



Background

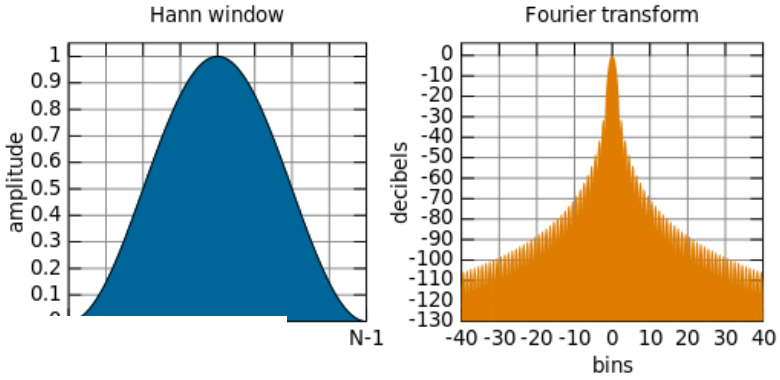
- Research is currently ongoing into best algorithm for polyphonic pitch detection with timbre rejection
- A paper in Spring 2012 describes Klapuri algorithm and several others
  - We chose to implement Klapuri because it had the lowest error rate of 6 algorithms discussed for 1-4 pitches per time-slice

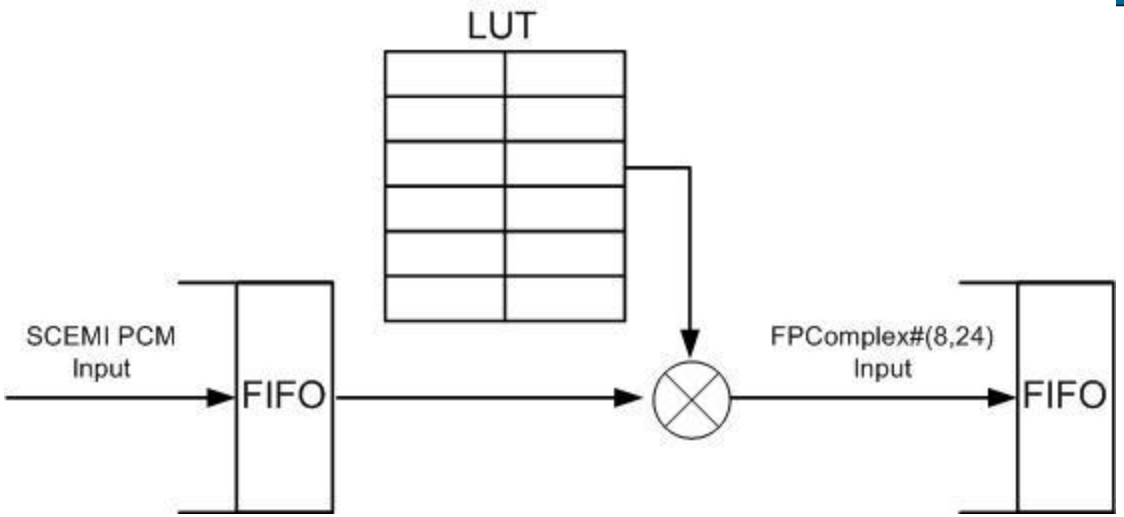
System Block Diagram



Input and Hanning Windows

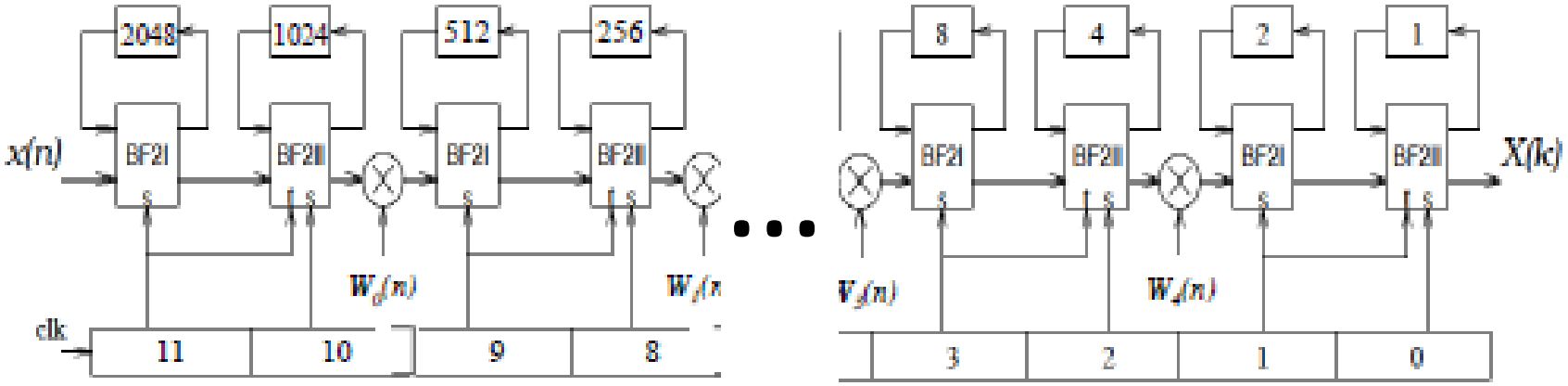
- Negate impact of sharp edges in time slices





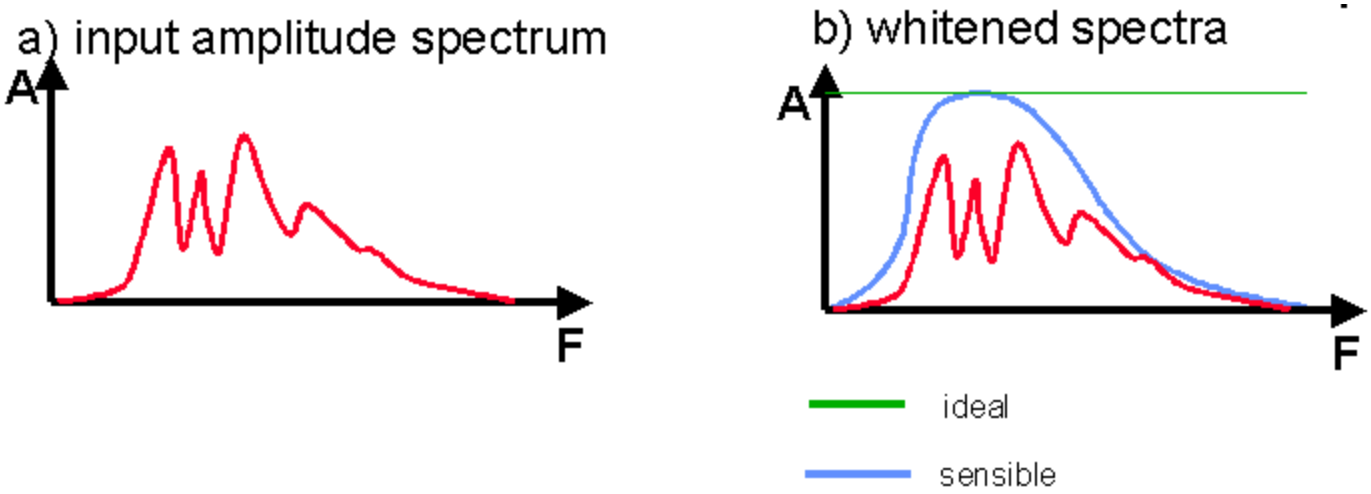
FFT and Absolute Value

- 4096-pt Streaming FFT
  - Minimal area requirements
  - ML605 Virtex-6 FPGA needed

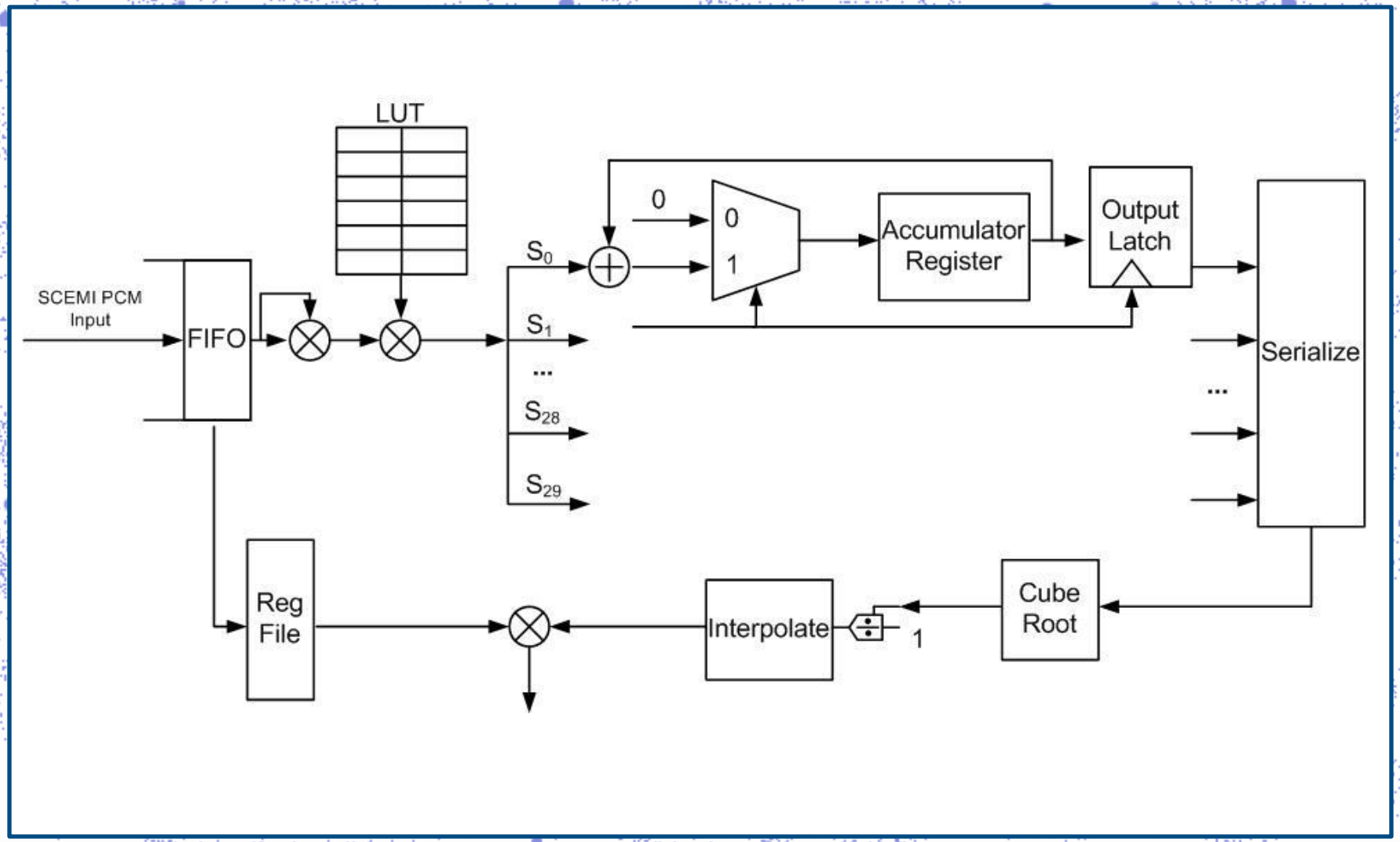


Spectral Whitening

- Smooth out spectrum of computed sub-bands without boosting noise
  - Uses inverse filtering to flatten spectrum energy distribution to mitigate timbre effects

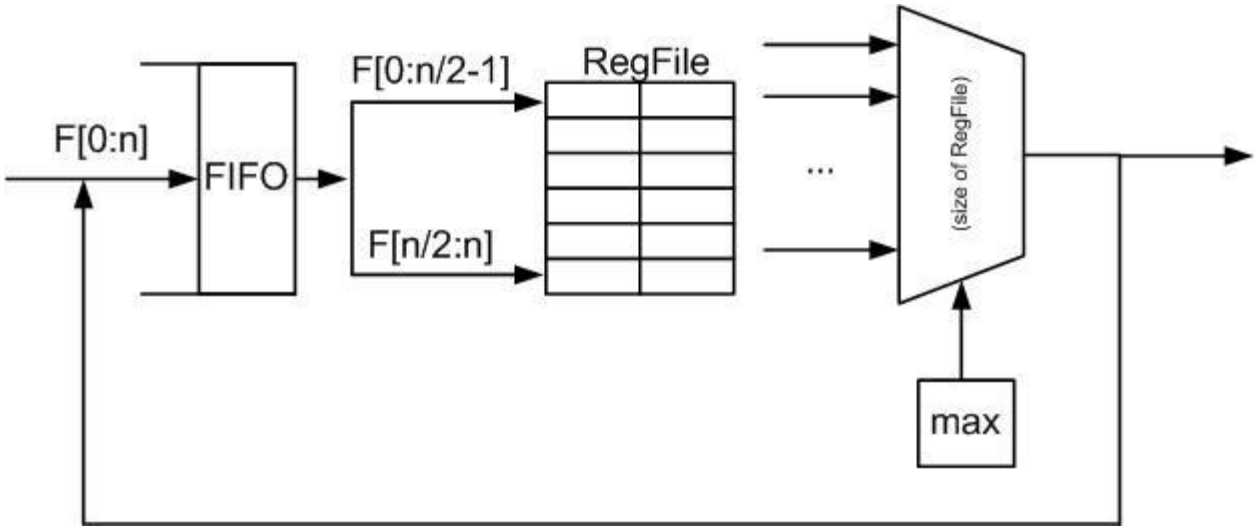


Spectral Whitening



Salience

- Uses variation of binary search to find pitches present during time-slice
  - Includes harmonic contributions



Test Bench

- MATLAB implementation of Klapuri algorithm
  - Compare to this result for algorithm checking
- Lilypond engraving software output compared to known
  - Shell scripts written to take output file and generate lilypond code and refresh pdf in realtime

Evaluation

- FFT output is in bit-reversed order
  - Implemented a sorter
- Cube root implementation in binary
  - More efficient
- SquareRoot and StreamFFT were reused IP blocks
- Realtime?

Future Work

- Increase FFT resolution in response to smaller differences between pitches as lower frequencies
- Decreasing window size
- Adding beat detection
- Attempt to improve algorithm for 4-part piano harmonies
- User-friendly real-time transcription system (front-end GUI, user input)

Synthesis Report

- 4096-pt FFT has difficulties synthesizing on FPGA
  - Using Virtex-6 ML605
- We plan to complete synthesis and run on FPGA in the next few days

Questions?