Sound Analysis and `recognition of Musical notes`[☆]

Maunil Vyas, Deep Patel, Shreyas Patel, Jay Dangar

[a] *Digital Signal Processing*
[b] *SEAS, Ahmadabad University*

Abstract

The article represents a detail study of different sound analysis and synthesis, It talks about the different sound analysis and synthesis approaches which are used for Harmonium note with pros and cons, form an application perspective compression mechanism is also demonstrated using STFT. The article first talks about the analysis part and then it discuss about the synthesis approaches.

Keywords: Fast Fourier Transform, Short-time Fourier Transform, Waveguide synthesis

1. Analysis Methods Used

The harmonium and piano notes have been recorded in a proper environments and multiple notes been analyzed properly to understand the instrument behaviour, still we are at moderate stage to conclude anything about an instruments behaviour because the recording process is not as rich as it required to catch all the essential things from an instrument.

1.1. FFT

Fast Fourier Transform is an algorithm which computes the Discrete Fourier Transform (DFT) of a sequence. Fourier analysis converts a signal from its original domain (often time or space) to a representation in the frequency domain and vice versa. Here, we have analyzed the musical chord of middle C note of the Piano and Harmonium using FFT.

one can clearly see that though both notes have same pitch still due to overtones they produce different sound.

1.2. **STFT**

The spectral content of speech changes over time (non stationary),As an example, formants change as a function of the spoken phonemes or musical notes,Applying the DFT over a long window does not reveal transitions in spectral content
This problems can be solved using Short Time Fourier Transform because for short enough windows, speech can be considered to be stationary. STFT segments the signal into narrow

[☆] This is only an example

Dominant Frequencies of the Harmonium:

Frequency(X)	Amplitude(Y)
265.9	0.04351
533	0.03106
799.3	0.02501
1066	0.007601
1332	0.03025
1595	0.009108
1865	0.0162
2132	0.01177
2398	0.008996
2664	0.01075
2932	0.01091
3191	0.002015
3465	0.001889
3722	0.001304
3988	0.001014
4254	0.0003058
4531	0.0003311
4795	0.000171

Dominant Frequencies of the Piano:

Frequency(X)	Amplitude(Y)
261.8	0.05643
522.2	0.006164
785	0.01735
1047	0.01309
1310	0.004121
1574	0.001015
1840	0.001694
2105	0.00198
2372	0.0007315

Figure 1: Frequency domain analysis of Harmonium and Piano note

time intervals and then takes the fourier transform of this segments. Each FT provides the spectral information of a separate time-slice of the signal, providing simultaneous time and frequency information.
Here as part of analysis, hamming window is used. Here is the Spectrogram or STFT of the A5-E5-G5 chord of the Piano. Notice the Bright Yellow stripes, they are the dominant frequencies.

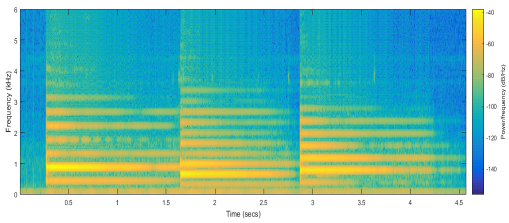

Figure 2: Spectrum Analysis of the Musical Chord

2. Synthesis Methods Used

2.1. Approach 1 : FFT Synthesis

It's about to analyze the Fourier spectrum of a given note and then using it's Fourier coefficients re back to the time domain characteristics. Using the frequencies of Fig 1 and Fig 2, we have successfully synthesized the harmonium and piano note.

Pros and Cons

- Method is simple enough.
- Useful for those musical instruments which works on harmonics.

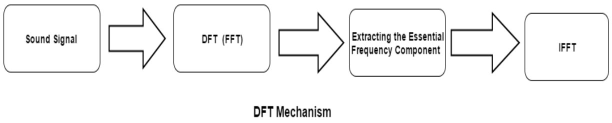

Figure 3: FFT Synthesis Approach

- Perception wise sound is not good enough
- Different musical instruments like guitar,flute etc are not able to synthesis easily with this method because they don't work on harmonics, so it's tough to simply use FFT coefficients for the overtones.
- Hard to determine the amplitude envelope for individual frequencies

2.2. Approach 2 : Waveguide Synthesis

Waveguide synthesis is as instrumental sound synthesis method. It is a technique to convert the actual device into the computational model and then implement it on the computers.

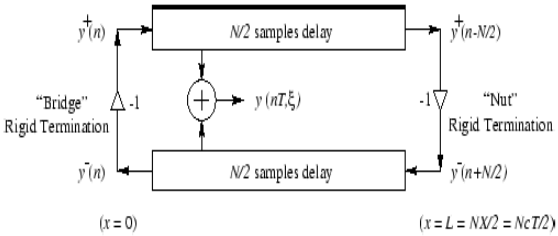

Figure 4: Waveguide General Model

At first, we have implemented the guitar. The original waveform and synthesized waveform of the F4(370 Hz).

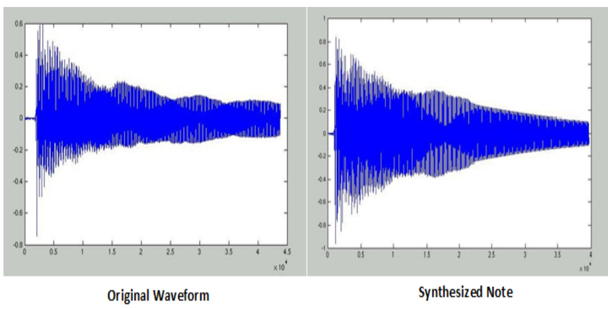

Figure 5: Waveguide General Model

We next tried to synthesis the harmonium by the waveguide model following the procedure as described in [1] and [2] , but we failed to do so as later we found out that synthesizing harmonium requires different approach than that applied to guitar synthesis. We need to emulate the organ pipes behavior for each note and then model it mathematically. But due to lack of resources and time we were unable to do so.

Pros and Cons

- Any Musical instrument is been synthesized using this technique.
- Too Complex to implement.
- Computation wise also it's high.

2.3. Approach 3 : STFT with Amplitude Envelop

The proposed approach is the best in terms of sound perception compare to the above two proposed approaches, the basic idea behind this approach is to analyze the time and frequency characteristics of signal together and then emulate the actual behaviour of the note which take care of the time and frequency specifications.

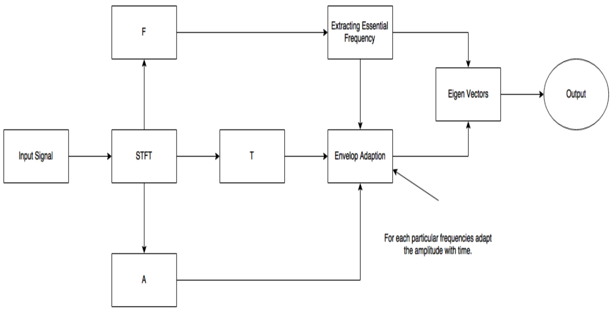

Figure 6: STFT with Amplitude Envelop

From fig 5 we can see that using STFT first we extract out the essential frequencies with associated time and magnitudes and then apply the envelope estimator for individual frequencies, the fact that harmonium follows the same behaviour for all notes the envelope estimation for one note is sufficient to incorporate the same behaviour for all notes.

Pros and Cons

- Method is comparatively simple than the waveguide synthesis.
- Promising results for instruments which follows the same behaviour for different notes.
- Hard to determine the amplitude envelope for individual frequencies
- Memory requirement is high.
- Different musical instruments like guitar,flute etc are not able to synthesis easily because they have different behaviour for high notes.

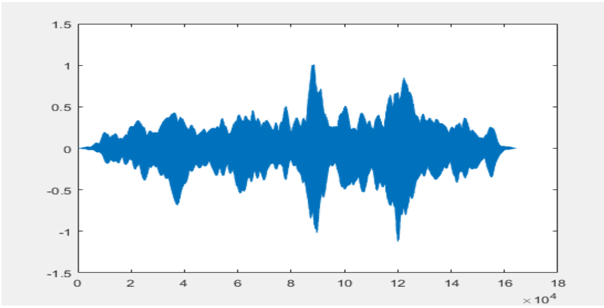

Figure 7: STFT with Amplitude Envelop

2.4. Compression mechanism using STFT

We have successfully implemented an audio compression mechanism using STFT and the results are as follows.

Size	Length	Window Size	No. of Frequency Components taken	Compression	Quality
0.815 MB	00:01:15	256 Samples	32	87.19%	Ok
2.48 MB	00:01:15	256 Samples	100	61.01%	BEST
0.407 MB	00:01:15	512 Samples	32	93.601%	BAD
1.24 MB	00:01:15	512 Samples	100	80.51%	GOOD
0.203 MB	00:01:15	1024 Samples	32	96.81%	WORST
0.636 MB	00:01:15	1024 Samples	100	90 %	WORST

Figure 8: STFT Compression Mechanism

3. Conclusion

The synthesized sound is good enough in terms of perception, the proposed method(Approach 3) suits well when it comes to synthesis of harmonium kind of musical instrument, still there are lot of scope for the improvements.

References

[1] Synthesizing a Guitar Using Physical Modeling Techniques, by Steven Sanders and Ron Weiss. www.ee.columbia.edu/ ronw/dsp/

[2] Towards High-Quality Sound Synthesis of the Guitar and String Instruments, by Matti Karjalainen , Vesa Vslimski , and Zoltan Janosy. International Computer Music Conference, September 10-15, 1993, Tokyo, Japan

[3] J. F. Alm, and J. S. Walker, Time-Frequency Analysis of Musical Instruments, Siam Review, vol. 44, no. 3, pp. 457-476, Aug. 2002

[4] J. O. Smith III, Physical Modeling using Digital Waveguides, Computer Music Journal, vol. 16, no. 4, pp. 74-91, Winter 1992.

[5] S. Kumar, and M. N. Mohanty, Synthesizing musical notes of Harmonium using Spectral Domain Modeling, International Journal of Scientific Engineering and Applied Science, vol. 1, issue 9, pp. 44-49, Dec. 2015.