Extraction Synthesis: Audio Synthesis Through Oscillation Isolation

1st Edition

Digital Down

August 26, 2024

Abstract

Extraction Synthesis is an audio processing technique that isolates individual oscillations from audio inputs. This technique operates by analyzing the raw audio data at the sample level, identifying oscillations that correspond to musical notes, and reconstructing these isolated oscillations into new audio outputs.

1 Introduction

Extraction Synthesis begins by converting audio files into a sequence of 16 bit raw audio values. In stereo audio, this results in a list of pairs, where each pair contains the left and right channel values for the sequence of samples. Samples are then scanned for alternating positive and negative values. The sequence of positive and negative values summed together creates a single oscillation. The oscillations are transferred into a new list of sample values that correspond to a musical note.

2 Methodology

Audio Input → Conversion to Sample List → Processing → Conversion to Audio File → Audio Output

2.1 Oscillation Identification

Extraction Synthesis scans the list of sample values to identify complete oscillations. An oscillation is defined as a sequence of sample values that starts positive, becomes negative, and then returns to positive.

Let S be the set of sample values $\{s_1, s_2, ..., s_n\}$
An oscillation O is defined as a subsequence of S where:

$$O = \{s_i, s_{i+1}, ..., s_j\} \text{ where } s_i > 0, s_j > 0, \text{ and } \exists k : i < k < j, s_k < 0$$

The algorithm tracks two counters:

- Positive Count: The number of consecutive positive samples
- Negative Count: The number of consecutive negative samples

When a complete oscillation has been scanned, the algorithm checks if an accepted sample sequence has occurred.

2.2 Note Frequency Matching

For each identified oscillation, the total sample count (Positive Count + Negative Count) is compared to the expected sample count for each musical note. The expected sample count for a note is calculated as follows:

$$\text{Expected Sample Count} = \text{Round}(\frac{\text{Sample Rate}}{\text{Note Frequency}})$$

Where:

- Sample Rate = 44100 Samples Per Second
- Note Frequency is in Hz

For example, for the note A4:

- A4 Frequency = 440 Hz
- Expected Sample Count = Round(44100 / 440) = 100

The algorithm allows for a tolerance of ±1 sample to account for rounding errors and allow more information to pass through.

2.3 Amplitude Threshold

To reduce the impact of noise, dithering, and unwanted sounds, the algorithm implements an amplitude threshold. The average absolute value of the samples in an oscillation must exceed this threshold for the oscillation to be accepted.

$$\text{Average Absolute Value} = \frac{\sum |s_i|}{n}$$

(s_i are the samples in the oscillation and n is the number of samples)

2.4 Oscillation Extraction and Storage

When an oscillation meets both the frequency matching and amplitude threshold criteria, it is extracted and added to a list corresponding to its matched note.

2.5 Note Frequency Tables

The following tables list all the notes considered in the current implementation, along with their frequencies and expected oscillation counts, divided by octave:

Note	Frequency (Hz)	Expected Oscillation Count
	Octave 1	
C1	32.70	1348
C#1/D♭1	34.65	1273
D1	36.71	1201
D#1/E♭1	38.89	1134
E1	41.20	1070
F1	43.65	1010
F#1/G♭1	46.25	953
G1	49.00	900
G#1/A♭1	51.91	849
A1	55.00	802
A#1/B♭1	58.27	757
B1	61.74	714

Table 1: Note Frequencies and Expected Oscillation Counts - Octave 1

Note	Frequency (Hz)	Expected Oscillation Count
	Octave 2	
C2	65.41	674
C#2/D♭2	69.30	636
D2	73.42	601
D#2/E♭2	77.78	567
E2	82.41	535
F2	87.31	505
F#2/G♭2	92.50	477
G2	98.00	450
G#2/A♭2	103.83	425
A2	110.00	401
A#2/B♭2	116.54	378
B2	123.47	357
	Octave 3	
C3	130.81	337
C#3/D♭3	138.59	318
D3	146.83	300
D#3/E♭3	155.56	283
E3	164.81	268
F3	174.61	253
F#3/G♭3	185.00	238
G3	196.00	225
G#3/A♭3	207.65	212
A3	220.00	200
A#3/B♭3	233.08	189
B3	246.94	179
	Octave 4	
C4	261.63	169
C#4/D♭4	277.18	159
D4	293.66	150
D#4/E♭4	311.13	142
E4	329.63	134
F4	349.23	126
F#4/G♭4	369.99	119
G4	392.00	113
G#4/A♭4	415.30	106
A4	440.00	100
A#4/B♭4	466.16	95
B4	493.88	89

Table 2: Note Frequencies and Expected Oscillation Counts - Octaves 2, 3, and 4

Note	Frequency (Hz)	Expected Oscillation Count
	Octave 5	
C5	523.25	84
C#5/D♭5	554.37	80
D5	587.33	75
D#5/E♭5	622.25	71
E5	659.25	67
F5	698.46	63
F#5/G♭5	739.99	60
G5	783.99	56
G#5/A♭5	830.61	53
A5	880.00	50
A#5/B♭5	932.33	47
B5	987.77	45
	Octave 6	
C6	1046.50	42
C#6/D♭6	1108.73	40
D6	1174.66	38
D#6/E♭6	1244.51	35
E6	1318.51	33
F6	1396.91	32
F#6/G♭6	1479.98	30
G6	1567.98	28
G#6/A♭6	1661.22	27
A6	1760.00	25
A#6/B♭6	1864.66	24
B6	1975.53	22

Table 3: Note Frequencies and Expected Oscillation Counts - Octaves 5 and 6

2.6 Reconstruction and Output

After processing the entire input audio, the extracted oscillations for each note are concatenated and converted back into an audio file. This results in a set of audio files, each containing the isolated oscillations for a specific note.

3 Conclusion

The output files are individual notes designed to be used as instruments in an audio production environment, allowing musicians and producers to create custom sounds using the extracted oscillations.

Further Research

Further research is currently being done to extract percussive elements from a sound source by isolating formerly rejected oscillations and applying several similar amplitude threshold functions as a separate type of Extraction Synthesis tentatively named "Unfamiliar Extraction Synthesis". By furthering this research, potentially a more qualitative approach can be developed with more user control over the dynamics of the generated output leading to more desirable audio output.