Temporal Dissection Synthesis: Pitch Preservation in Time-Stretching Manipulations

1st Edition

Digital Down

September 18, 2024

Abstract

Temporal Dissection Synthesis is an audio processing technique that enables precise time-stretching of audio signals at the oscillation level. This method offers a unique approach to altering the temporal characteristics of sound while preserving pitch.

1 Introduction

Temporal Dissection Synthesis builds upon the principles of Dissection Synthesis [1], focusing on the temporal aspects of audio signals rather than harmonic transformations. By manipulating the sample lengths of the oscillations, this technique allows for the expansion or compression of time without altering the fundamental frequency of the sound.

2 Methodology

2.1 Oscillation Identification

The process of oscillation isolation and extraction follows the method outlined in Extraction Synthesis [2]. An oscillation is identified as a sequence of sample values that starts at zero, becomes either positive or negative, crosses zero, and then returns to zero.

2.2 Temporal Transformations

Once oscillations are extracted, they are manipulated to achieve the desired time-stretching effect. Two primary modes of operation are defined: TDSSlow for time expansion and TDSFast for time compression.

2.2.1 TDSSlow (Time Expansion)

To extend the duration of the audio, each oscillation is duplicated immediately after its occurrence. This process effectively doubles the time of the initial audio while maintaining its pitch.

<div align="center">Oscillation Duplication</div>

$$y[n] = \begin{cases} x[n] & \text{for the original oscillation} \\ x[n - L_i] & \text{for the duplicated oscillation} \end{cases}$$

Where $x[n]$ represents the original signal, $y[n]$ is the time-expanded signal, and L_i is the length of the i-th oscillation.

2.2.2 TDSFast (Time Compression)

To reduce the duration of the audio, every other oscillation is omitted from the output. This process approximately halves the time of the initial audio while preserving its pitch characteristics.

<div align="center">Oscillation Omittance</div>

$$y[n] = \begin{cases} x[m] & \text{if oscillation index } i \text{ is even} \\ \text{omitted} & \text{if oscillation index } i \text{ is odd} \end{cases}$$

Where m represents the sample index within the kept oscillations, and i is the index of the oscillation in the original signal.

Different channels have different oscillation sample lengths, therefore the right channel is not processed with Oscillation Omittance. The same sample lengths from the left channel are applied to the right channel to ensure temporal alignment. Mono signals are preferred in this method because of the absence of processing the oscillations in the right channel, leadind to intermittent artifacts.

3　Conclusion

Unlike traditional time-stretching methods that alter the pitch of the audio, Temporal Dissection Synthesis maintains the frequency information in each oscillation.

References

[1] Digital Down. (2024). Dissection Synthesis: Intermediate Harmonic Transformations at the Oscillation Level. 1st Edition.

[2] Digital Down. (2024). Extraction Synthesis: Audio Synthesis Through Oscillation Isolation. 1st Edition.

Further Research

Further research is currently being conducted on Dissection Scaling and Chordal Dissection Scaling, which aims to be pitch preservation tuning on an individual note level. Dissection Scaling defines a scale and adjusts the oscillations to be notes within the specific scale, likewise Chordal Dissection Scaling defines a chord and adjusts the oscillations to be notes within the specific chord.