Sound Synthesis

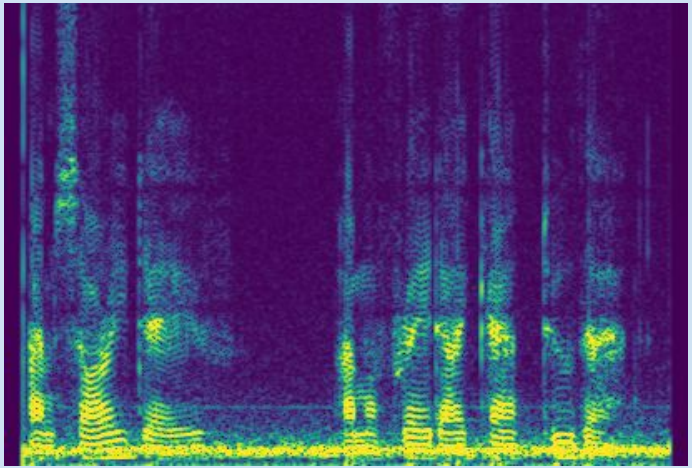
Joe Bumpus & Noah Miller

ECE 477 - Computer Audition

University of Rochester

Sound Synthesis

The generation of sounds using algorithms implemented in analogue or digital forms



“Home: Centre for Digital Music (C4DM): Queen Mary, University of London,” *Home | Centre for Digital Music (C4DM) | Queen Mary, University of London*. [Online]. Available: http://c4dm.eecs.qmul.ac.uk/. [Accessed: 03-Dec-2022].

Why is it important?

- There are many situations in which a full recording of some sound may not be the most cost-effective or efficient method of acquiring audio
- Recording high quality audio can be very expensive, and frequently requires specialized engineers, equipment, and environments
- Certain physical musical instruments are cumbersome, impractical, and can be extremely difficult to acquire
- There are certain scenarios, such as communications, where transmitting a full size audio recording is prohibitively inefficient

Smith, J.O. Physical Audio Signal Processing, http://ccrma.stanford.edu/~jos/pasp/, online book, 2010 edition. [Accessed: 01-Dec-2022].

Traditional Approaches

Physical Modeling

Based on mathematical descriptions of acoustic musical instruments

Describe the behavior of musical instrument elements (strings, reeds, tubes, membranes, etc.) with set of differential equations. Solve the set of equations to yield an output subject to some input excitation

Spectral Modeling

Based on the description of sound characteristics, perception of sound

Describe spectral characteristics (harmonics, timbre, etc.) of a sound. Generate target sound through diverse mathematical and computational procedures (method dependent)

Concatenative Synthesis

Data driven approach using actual recordings

Uses segments of sound from a large database, segmented into units. A unit selection algorithm finds the sequence of units that best matches the target sound. Selected units can then be transformed to fully match the target specification, and are concatenated.

J. O. Smith III, “Viewpoints on the history of Digital Synthesis - Stanford University,” 11-Aug-2002. [Online]. Available: https://ccrma.stanford.edu/~serafin/UVA/MUSI445/hw/kna.pdf. [Accessed: 02-Dec-2022].
X. Serra, “State of the Art and Future Directions in Musical Sound Synthesis,” Nov-2007. [Online]. Available: https://www.researchgate.net/publication/4303318_State_of_the_Art_and_Future_Directions_in_Musical_Sound_Synthesis. [Accessed: 01-Dec-2022].
S. Bilbao, Physical Modeling, 15-Nov-2006. [Online]. Available: https://ccrma.stanford.edu/~bilbao/booktop/node12.html. [Accessed: 03-Dec-2022].
D. A. Jaffe and J. Smith, “Extensions of the Karplus-Strong Plucked-String Algorithm,” Semantic Scholar, 01-Jan-1983. [Online]. Available: https://www.semanticscholar.org/paper/Extensions-of-the-Karplus-Strong-Plucked-String-Jaffe-Smith/995470417936f34e4a6f350a8b16eedbdc1b7d28. [Accessed: 03-Dec-2022].
J. Urrego, “Plucked String ‘Nothing Else Matters’ using Karplus-Strong Synthesis,” Semantic Scholar, 2013. [Online]. Available: https://www.semanticscholar.org/paper/Plucked-String-%E2%80%9C-Nothing-Else-Matters-%E2%80%9D-using-Urrego/f23e41b6bbae30fd3449a04e87691bd5e400a540. [Accessed: 03-Dec-2022].

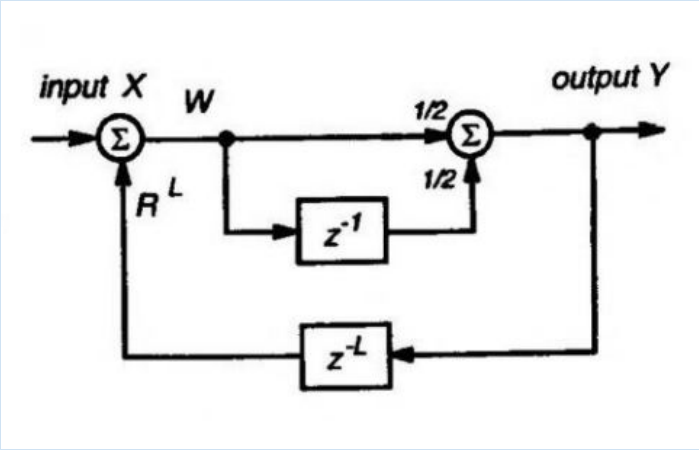
Physical Example: Karplus-Strong

Karplus-Strong String Synthesis

- Represent string with high-order digital filter
- Represent “pluck” with a short noise burst

The loop filter (right) determined by cascade of

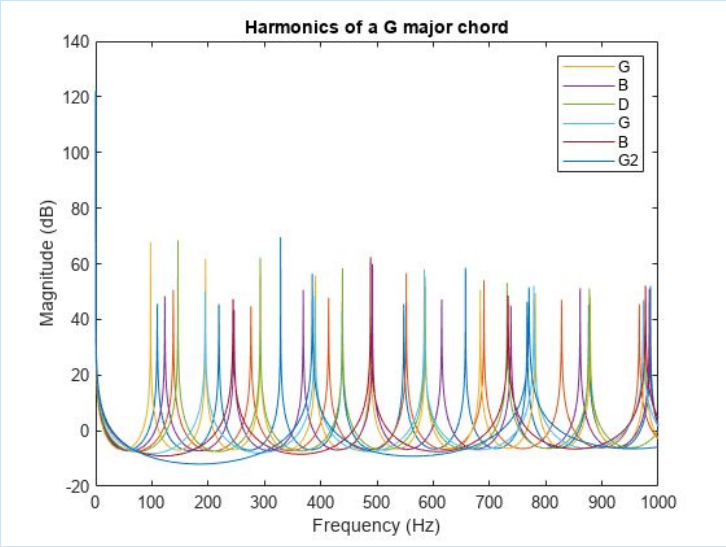
- the filtering experienced by a traveling wave in traversing the string twice
- the reflection transfer functions of the two terminations



D. A. Jaffe and J. Smith, “Extensions of the Karplus-Strong Plucked-String Algorithm,” Semantic Scholar, 01-Jan-1983. [Online]. Available: https://www.semanticscholar.org/paper/Extensions-of-the-Karplus-Strong-Plucked-String-Jaffe-Smith/995470417936f34e4a6f350a8b16eedbdc1b7d28. [Accessed: 03-Dec-2022].

J. Urrego, “Plucked String ‘Nothing Else Matters’ using Karplus-Strong Synthesis,” Semantic Scholar, 2013. [Online]. Available: https://www.semanticscholar.org/paper/Plucked-String-%E2%80%9C-Nothing-Else-Matters-%E2%80%9D-using-Urrego/f23e41b6bbae30fd3449a04e87691bd5e400a540. [Accessed: 03-Dec-2022].

Physical Example: Karplus-Strong



MATLAB Demo:

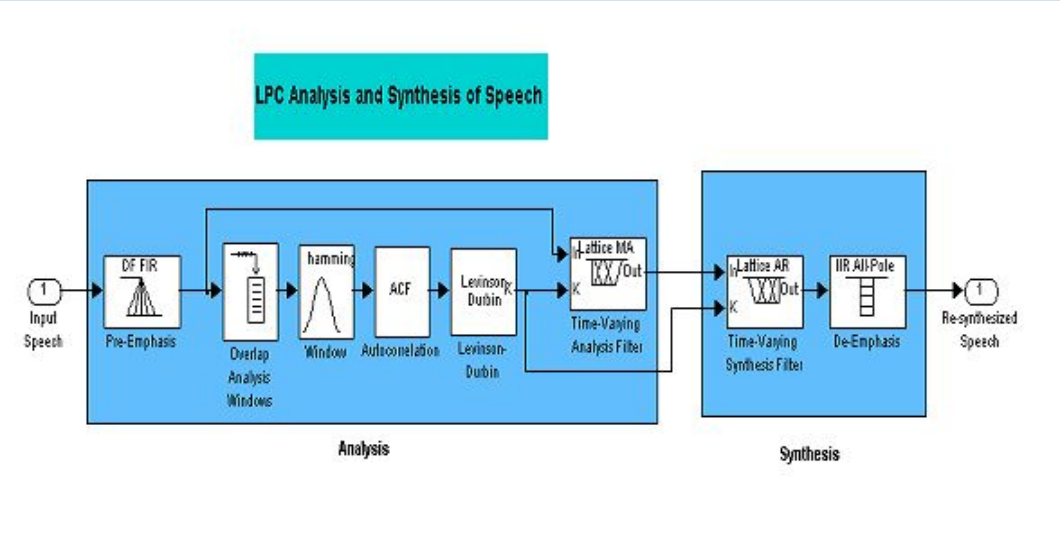
- Define fret positions and position dependent delays
- Build a cell array of numerator and denominator coefficients of an IIR filter for each string
- Filter random noise vector
- Add strumming effect with simple 50ms offset

Synthesized Guitar Output:

"Generating Guitar Chords Using the Karplus-Strong Algorithm," MATLAB & Simulink Example. [Online]. Available: https://www.mathworks.com/help/signal/ug/generating-guitar-chords-using-the-karplus-strong-algorithm.html. [Accessed: 03-Dec-2022].

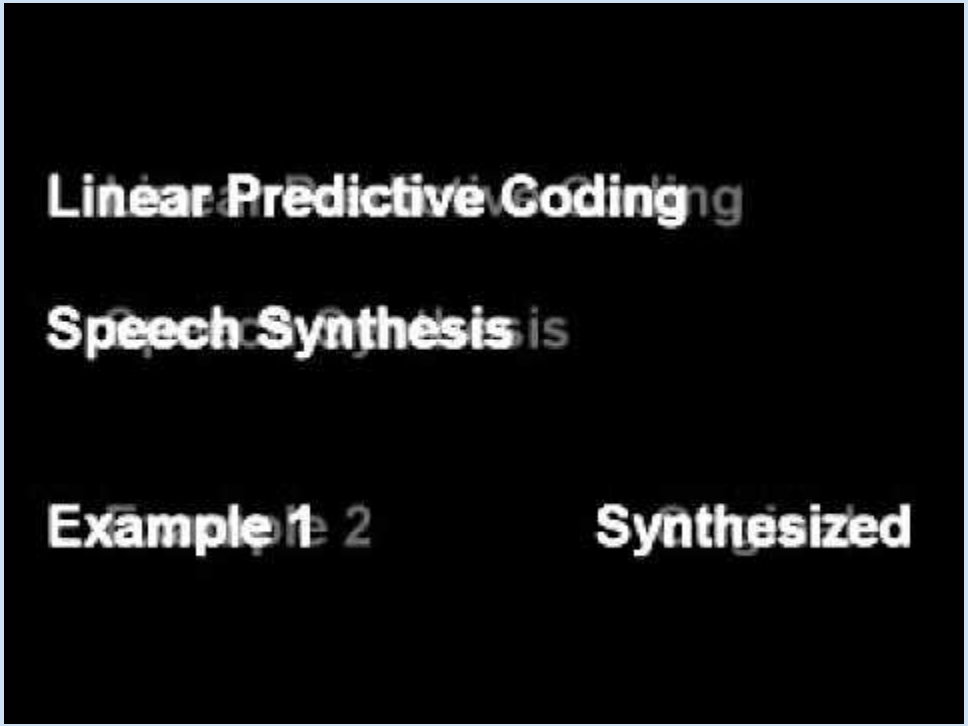
Spectral Example: LPC

- Linear Predictive Coding is a method for representing the spectral information and characteristics of an audio signal, and can be used to resynthesize audio, particularly speech
- LPC has been used widely in early communication and compression techniques for cellular networks, as it is much more efficient than communicating the raw audio signal



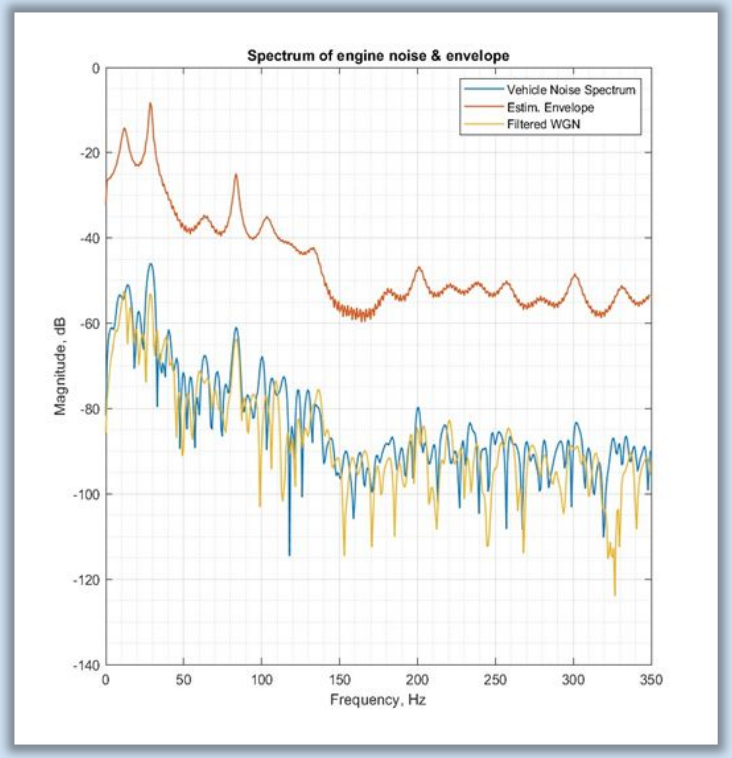
"LPC analysis and synthesis of speech," LPC Analysis and Synthesis of Speech - MATLAB & Simulink. [Online]. Available: https://www.mathworks.com/help/dsp/ug/lpc-analysis-and-synthesis-of-speech.html. [Accessed: 03-Dec-2022].

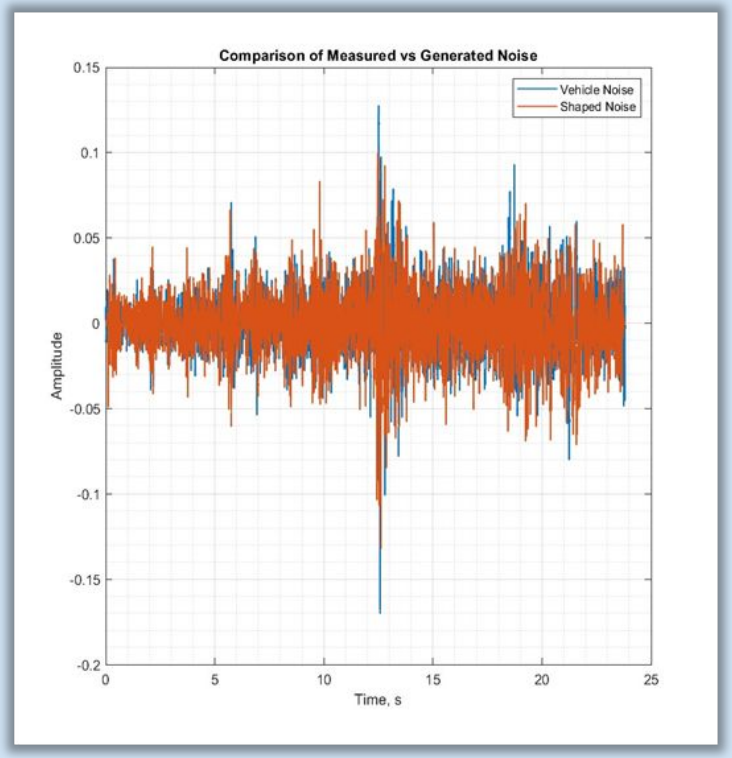
Spectral Example: LPC



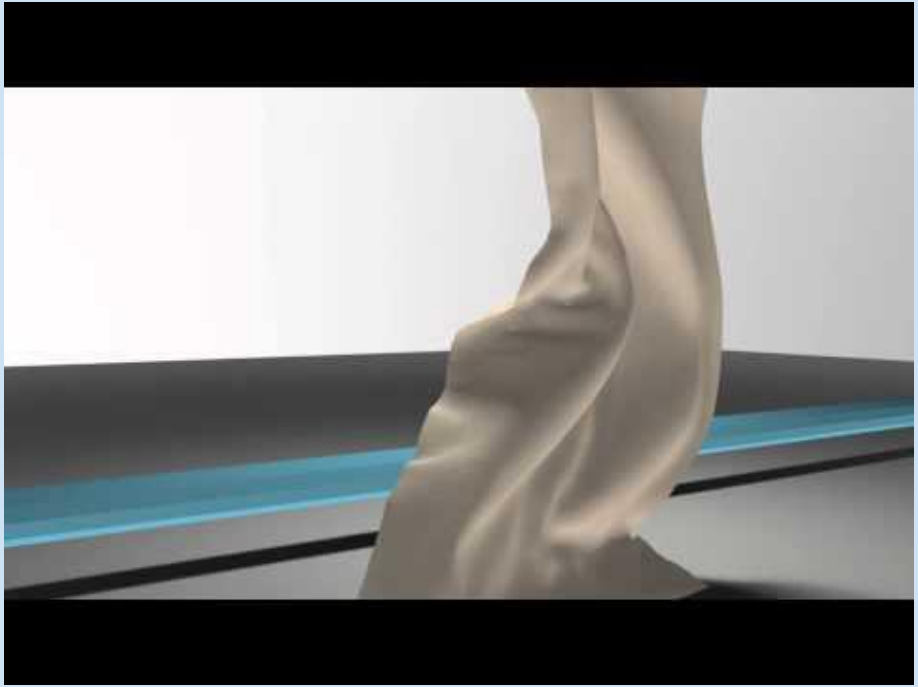
Linear Predictive Coding Speech Synthesis Samples. YouTube, 2014.

Spectral Example: LPC





Concatenative Synthesis Example

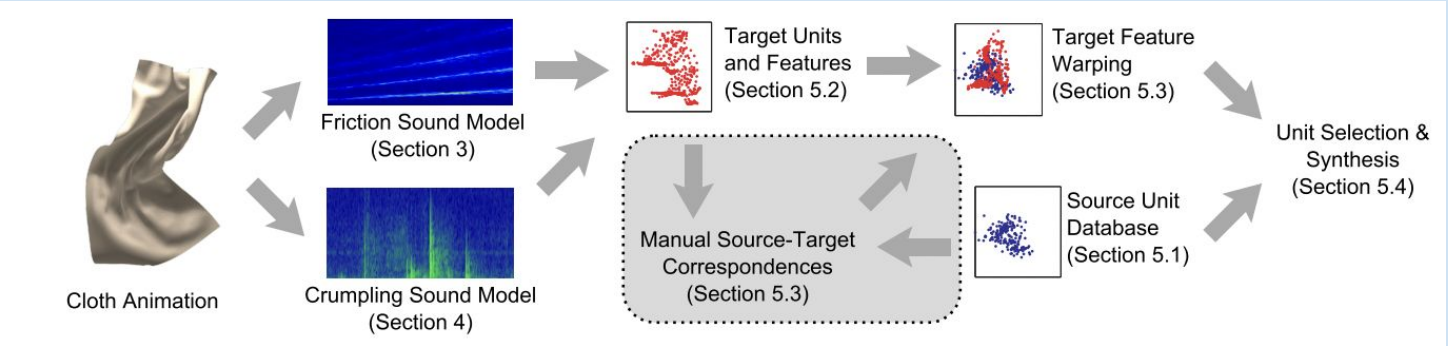


https://www.youtube.com/watch?v=56lq84Fx47g

Steven S. An, Doug L. James, and Steve Marschner, Motion-driven Concatenative Synthesis of Cloth Sounds, ACM Transaction on Graphics (SIGGRAPH 2012), August, 2012.

Concatenative Synthesis Example

- Given a cloth animation, analyze the deformations and use motion events to drive crumpling and friction sound models estimated from cloth measurements
- Synthesize a low-quality target sound
- Select a sequence of microsound units (4.2ms found to be typically long enough to maintain the characteristics of the recorded sounds) from a database of recorded cloth sounds which best match the synthesized target
- Concatenate units together to produce the final cloth sound

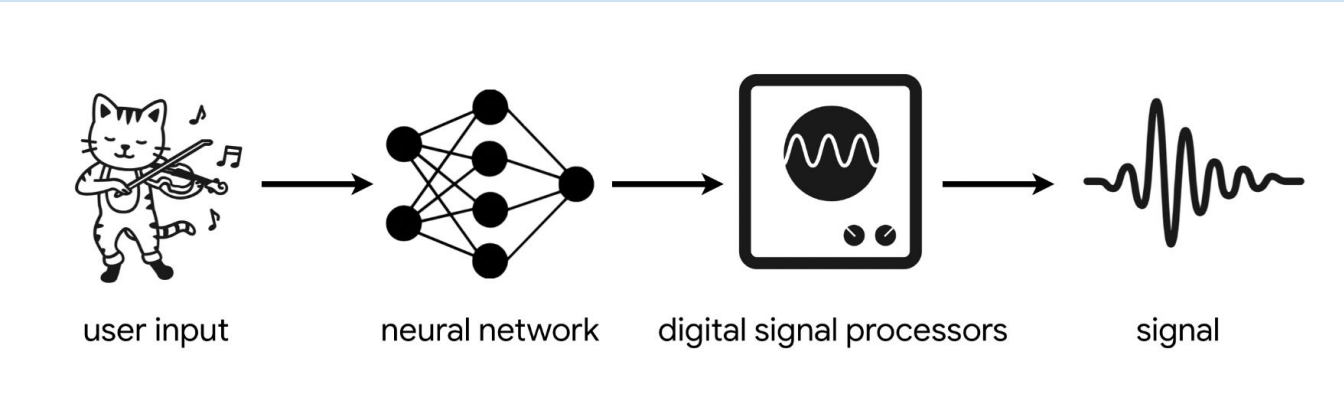


Steven S. An, Doug L. James, and Steve Marschner, Motion-driven Concatenative Synthesis of Cloth Sounds, ACM Transaction on Graphics (SIGGRAPH 2012), August, 2012.

Deep Learning Methods: DDSP

Differentiable Digital Signal Processing (DDSP)

- Python library that uses neural networks to accomplish a variety of sound synthesis tasks, including the transfer of an audio signal to a different instrument's timbre
- Combination of the interpretable structure of classical DSP elements (filters, oscillators, reverberation, etc.) with the expressivity of deep learning
- Use a neural network to convert a user's input into complex DSP controls that can produce more realistic signals

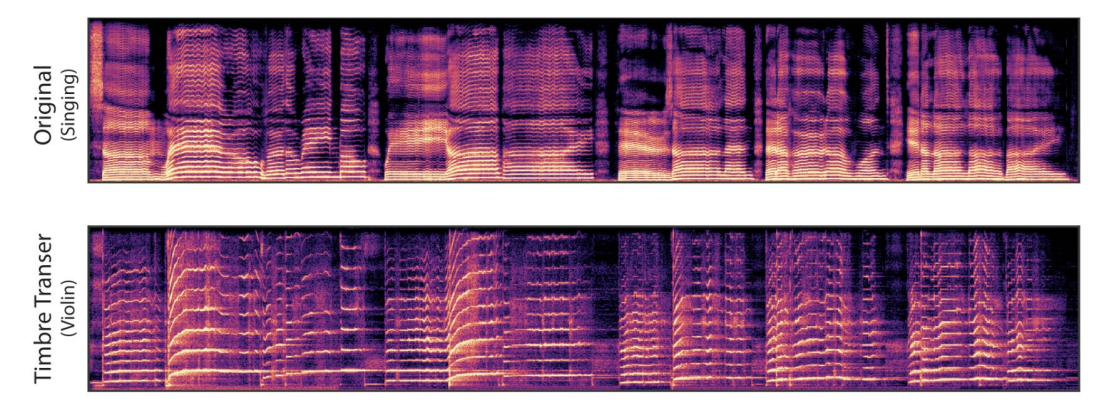


J. Engel, L. (H. Hantrakul, C. Gu, and A. Roberts, "DDSP: Differentiable Digital Signal Processing," OpenReview, 25-Sep-2019. [Online]. Available: https://openreview.net/forum?id=B1x1ma4tDr. [Accessed: 06-Dec-2022].

Deep Learning Methods: DDSP

DDSP Example: Timbrel Transfer

- Using differentiable oscillators, filters, and reverberation from DDSP library enables training of high-quality audio synthesis models with less data and fewer parameters as they do not need to learn to generate audio from scratch
- All models trained on less than 13 minutes of audio
- Timbrel transfer achieved by keeping frequency and loudness the same; change character of the instrument's sound

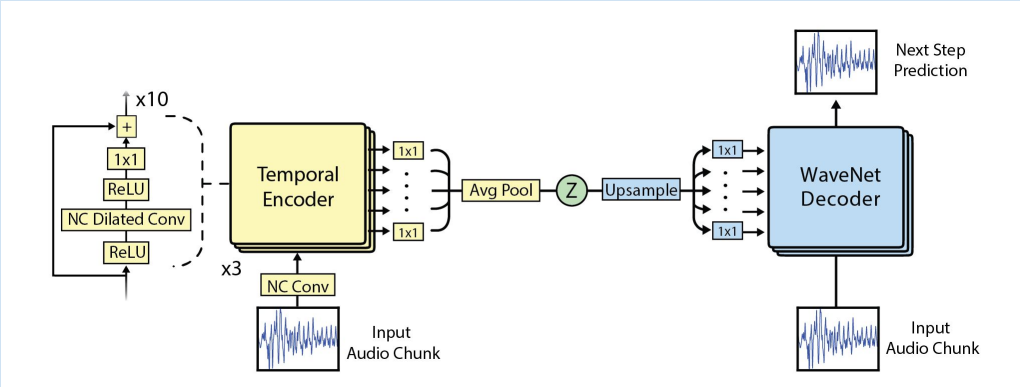




J. Engel, L. (H. Hantrakul, C. Gu, and A. Roberts, "DDSP: Differentiable Digital Signal Processing," OpenReview, 25-Sep-2019. [Online]. Available: https://openreview.net/forum?id=B1x1ma4tDr. [Accessed: 06-Dec-2022].

Deep Learning Methods: NSynth

- Nsynth: A Neural Network Based Synthesizer
- Modeled after Computer Vision algorithms and image processing, it utilizes a vast dataset of individual notes from 1000s of instruments to create a novel generative synthesizer
- Takes advantage of a WaveNet based decoder to effectively model and learn all of the spectral and harmonic features of the instruments in the datasets



Original Audio

NSynth

"NSYNTH: Neural Audio Synthesis," Magenta, 06-Apr-2017. [Online]. Available: https://magenta.tensorflow.org/nsynth. [Accessed: 04-Dec-2022].

Future Research

- These two examples are just some of the ways in which Neural Networks have been integrated into Sound Synthesis research
- As it is an emerging field, there are many possible innovations, although there are already current developments and examples of the methods used in sound synthesis
- Some areas of focus include using machine learning to more accurately model analog synthesizers, reverberations and filters, as well as examining techniques to add more human elements to a digital performance

J. Engel, L. (H. Hantrakul, C. Gu, and A. Roberts, “DDSP: Differentiable Digital Signal Processing,” OpenReview, 25-Sep-2019. [Online]. Available: https://openreview.net/forum?id=B1x1ma4tDr. [Accessed: 06-Dec-2022].

Questions?