The Constant-Q Transform Spectral Envelope Coefficients

I. Scope

THE sliding discrete Fourier transform (SDFT) ... [1], [2]

[3]: NSynth

- Aware that some methods are attempting to learn more adapted feature like that - The envelope component can additionally be refined, along with the pitch component - Incidentally, the pitch component can be used to identify the pitch/key

II. Relevance

III. Prerequisites

Basic knowledge of audio signal processing and music information retrieval is required to understand this article, in particular, concepts such as the Fourier transform, convolution, spectral envelope, pitch, CQT, and MFCC.

IV. Problem Statement

A. Observations

Assumption: A log-spectrum, such as the CQT-spectrum, can be represented as the convolution of a pitch-invariant log-specrtal envelope component (= timbre) and a envelope-normalized pitch component.

- A pitch change in the audio translates to a linear shift in the log-spectrum.
- The Fourier transform (FT) of a convolution of two functions is equal to the point-wise product of their FTs (convolution theorem).
- The magnitude FT is shift-invariant.

V. Solution

VI. Numerical Example

VII. What We Have Learned

We have shown that ...

VIII. Author

Zafar Rafii (zafarrafii@gmail.com) received a PhD in Electrical Engineering and Computer Science from Northwestern University in 2014, and an MS in Electrical Engineering from both Ecole Nationale Superieure de l'Electronique et de ses Applications in France and Illinois Institute of Technology in the US in 2006. He is currently a senior research engineer at Gracenote in the US. He also worked as a research engineer at Audionamix in France. His research interests are centered on audio analysis, somewhere between signal processing, machine learning, and cognitive science, with a predilection for source separation and audio identification.

References

[1] J. C. Brown, "Calculation of a constant Q spectral transform," *Journal of the Acoustical Society of America*, vol. 89, no. 1, pp. 425–434, 1991.

[2] J. C. Brown and M. S. Puckette, "An efficient algorithm for the calculation of a constant Q transform," *Journal of the Acoustical Society of America*, vol. 92, no. 5, pp. 2698–2701, 1992.

[3] J. Engel, C. Resnick, A. Roberts, S. Dieleman, D. Eck, K. Simonyan, and M. Norouzi, "Neural audio synthesis of musical notes with WaveNet autoencoders," in *34th International Conference on Machine Learning*, Sydney, NSW, Australia, August 6-11 2017.