A Review of Modern Music Computational Analysis Methods

Sean Bisset – 1905262

*Music* is a universal language that can be interpreted in many ways. Computational music analysis has been around for many years and is still a growing field. However, the analysis still needs to include many chords and concepts. For instance, most analysis methods do not consider chords other than major or minor, limiting the development of more complex Audio Chord Estimation (ACE) models, therein further limiting deeper analysis of concepts such as emotion or quality of music. Some approaches also suffer from noise in the audio samples. This literature review aims to provide an overview of music analysis and shed some light on current analysis methods in music technology, then critically analyse those methods and discuss the applicability of the research for computational musical analysis.

Distinguishing key/chord/bass within a piece should be split into two stages: chromagram extraction from a piece and estimating the information based on the chromagrams.

*Short-Time Fourier Transform (STFT)*In [1], Fujishima proposed that the Chromagram could be extracted from the audio signal by sorting each musical frequency into nearby frequency bins. These are then mapped onto one octave. This method is simple yet effective and has been used in many key/chord estimation models since the early 2000s (e.g., [4], [5] ). This method, however, suffers when noise is introduced into an audio sample as the frequency counts become distorted by the noise. Noise is not the only problem; instruments that do not contribute to the harmony but instead infer the rhythm (percussive track) pose a problem. A different method,

*Constant Q Transform (CQT).*In [2], Brown proposes that the Chromagram can be extracted by calculating the pitch energy class within the log-frequency domain. This method is derived from the knowledge that musical notes are equidistant in log-frequency. This method could then be used for instrument recognition using a pattern recognition algorithm. However, this method only works for music where the frequencies are spread equidistantly, as in Western music. Many non-western pieces would not work using the CQT method where they would in the STFT method, provided there is no noise or percussive track in the audio sample. Both the methods above also have the problem that chords are only sometimes beat synchronous. This would mean that the chromagram outputs are sensitive to noise and local transients [6]. A problem that a few methods can resolve,

One of which is *Reducing Local Variations of Chords* [6]. The researchers proposed that repetition, something found continuously throughout music, can smooth the chromagrams and reduce noise. Smoothing the chromagrams is done by taking multiple similar sample chromagrams and averaging them out to smooth out noise and unsystematic deviations. This method of smoothing out noise is excellent. However, there would be a worry that it may remove interesting “one off” parts that add to a piece. Many pieces add slight variations to the main thematic elements to create interest, and this method, if tuned incorrectly, could average these elements out. A simple example of this would be a piece where the V chord, G, is used in almost every bar, then to jazz up a section, the piece may make the G chord a G7 for a bar, which contains three of the same notes as the G chord and may therefore be averaged out. Instead, this could be fixed by ensuring averages are only taken locally instead of within the whole piece of music.

Another popular and simple method is to simply pass the

The problem percussive tracks create when using these methods can be negated by separating the audio into separate tracks or chromagrams. Some methods for achieving this are listed below:

*Harmonic/Percussive Sound Separation (HPSS),* proposed in [3], that the harmonic and percussive tracks could be separated easily since harmony usually has a well-defined frequency content. In contrast, percussive parts are usually spaced in time to create the piece's rhythm. This method is excellent at separating music with a well-defined rhythm but would struggle if the rhythm were inconsistent. The researchers then continued to show that the output chromagrams from this method only yielded better chord estimation performance.

An Overview of Music Perception and Computational Analysis

Citations/References

1. T. Fujishima, "Real time chord recognition of musical sound: A system using common lisp music," online. Available at <https://quod.lib.umich.edu/i/icmc/bbp2372.1999.446/--realtime-chord-recognition-of-musical-sound-a-system-using?view=image>. Accessed [02/11/2022]
2. J. Brown, "Calculation of a constant q spectral transform," J. Acoust. Soc. Amer., vol. 89, no. 1, pp. 425–434, 1991.
3. N. Ono, K. Miyamoto, J. Roux, H. Kameeoka, and S. Sagayama, "Separation of a monaural audio signal into harmonic/percussive components by complimentary diffusion on spectrogram," in Proc. EUSIPCO, 2008.
4. Thomas, Mathew & Jothish, Mintu & Thomas, Navin & Koolagudi, Shashidhar & Srinivasa Murthy, Y.V.. (2016). Detection of similarity in music files using signal level analysis. 1650-1654. 10.1109/TENCON.2016.7848297.
5. Turrubiates, José de Jesús & Ledesma, Sergio & Gonzalez-Reyna, Sheila & Avina-Cervantes, Juan Gabriel & Ilunga Mbuyamba, Elisee. (2016). Guitar audio signal classification by collapsed Pitch Class Profile. 1-6. 10.1109/ROPEC.2016.7830637.
6. T. Cho and J. P. Bello, “A feature smoothing method for chord recognition

using recurrence plots,” in Proc. ISMIR, 2011. Online. Available at <https://ismir2011.ismir.net/papers/OS8-4.pdf> Accessed [02/11/2022]