Rhythmic Similarity – a quick paper review

Shi Yong
yong.shi2@mail.mcgill.ca

1. Introduction

Rhythm is a very important aspect of music. It is very easy for a human to tell the difference of two different rhythmic patterns, say, the Cha Cha and Rumba. However, it is not trivial to model the rhythmic patterns and to quantitatively measure the similarity or dissimilarity. The following works attempted different ways to automatically calculate the rhythmic similarity.

2. Foote's work

J. Foote and S. Uchihashi (2001) presented an approach to measure the rhythm and tempo of a piece of music, the beat spectrum. The idea of beat spectrum is very like the frequency spectrum, and the counterpart of frequency spectrogram is beat spectrogram, which is the successive beat spectrum over time. The extraction of beat spectrum is done by the following steps. First, the audio stream is windowed by overlapped 256 samples frames. Then, the power spectrum of each frame is calculated and used as feature vectors. The distances of all pairwise combinations of feature vectors are then calculated. A 2-D similarity matrix is constructed based on these feature vectors. The beat spectrum can then be estimate by two ways: either by summing the similarity matrix along the diagonal, or by calculating the auto-correlation of the similarity matrix.

Once the beat spectra have been measured for two different rhythm patterns, a similarity value can be obtained, as presented in J. Foote, M. Cooper and U. Nam (2002). The beat spectra, as 1-dimensional functions of lag time, are truncated to L-dimensional vectors, and then distance can be measured between any two beat spectra. Three different distance functions are evaluated, namely Euclidean Distance, Cosine Distance, and Fourier Beat Spectral Coefficients. The result shows the latter two functions outperform the simple Euclidean Distance.

Also, some possible applications were proposed:

- Automatic ranking songs by similar tempo and rhythm
- Music segmentation by rhythm
- Automatic tempo extraction
- Automatic DJ for personal music collections
- Hybrid music retrieval engine by pitch and/or timbre similarity

3. Paulus's work

Jouni Paulus and Anssi Klapuri (2002) presented an automatic rhythmic pattern segmentation and rhythmic similarity measuring system. The amplitude envelope signal at each frequency band was first extracted from the audio stream by a set of processing methods, and then the periodicity was analyzed by a periodicity detection algorithm, which was originally proposed for fundamental frequency estimation. An intermediary signal $s(\tau)$ was then obtained and used for further musical meter estimation. Pattern boundaries were estimated at three different levels, namely tatum, tactus, and musical measure. By using such pattern segmentation algorithm, a 67% correct rate of 365 pieces for tactus periods and a 77% correct rate of 141 pieces for musical measure lengths were reported. For each pattern, acoustic features were extracted from a series of consecutive time frames. Three features, namely loudness, spectral centroid, and MFCCs were tested. Dynamic Time Warping was used to match feature vectors with different lengths, and a similarity measure was obtained by comparing feature vectors of two patterns.

In an experiment with 41 different rhythmic patterns, the result shows this system works pretty good in that a high similarity was assigned to the same rhythms performed by different drum sets. The spectral centroid weighted with the loudness was used in the experiment, and it was proved to have a nice performance.

4. Dixon's work

Simon Dixon et al. (2004) presented a new way to characterize music by "typical bar-length rhythmic patterns", and evaluated its usefulness in dance music genre classification. The feature is based on the fact that different genres of ballroom dance music can be distinguished by the temporal sequence. For a piece of music, the amplitude envelope of each bar was extracted by RMS filters, and then a k-means clustering was used to find the most prominent rhythmic pattern, which was defined as the "typical rhythmic pattern". The usefulness of rhythmic pattern in dance music genre classification was tested in experiments. It was reported that 50% classification rate (the baseline is 16%) was achieved by using the rhythmic pattern feature alone, 84% by using together with other automatically computed features (derived from rhythmic patterns or audio data), and 96% by using measured tempo. It turns out that the rhythmic pattern can be a very useful feature in genre classification. It also turns out that the tempo is a very helpful feature.

5. Conclusion

In the abovementioned works, various approaches have been tested in similarity measuring. Features extracted from either frequency domain or time domain have been successfully tested. Typically, the similarity can be measured by means of distance functions, such as Euclidean distance, or cosine distance, etc. Feature vectors with different length can be matched by dynamic time warping algorithm. Pattern segmentation is not easy, while it is essential for rhythmic similarity measuring.

References

Foote, J. and S. Uchihashi. 2001. The Beat Spectrum: A New Approach to Rhythm Analysis. In *Proceedings of the International Conference on Multimedia and Expo*.

Foote, J., M. Cooper and U. Nam. 2002. Audio Retrieval by Rhythmic Similarity. In *Proceedings of the 3rd International Symposium on Musical Information Retrieval*.

Paulus, J., and A. Klapuri. 2002. Measuring the Similarity of Rhythmic Patterns. In *Proceedings of the 3rd International Symposium on Musical Information Retrieval*.

Dixon, S., F. Gouyon, and G.Widmer. 2004. Towards Characterisation of Music via Rhythmic Patterns. In *Proceedings of the International Conference on Music Information Retrieval.*