Aurélien Antoine*[1], Eduardo R. Miranda*[2]

*Interdisciplinary Centre for Computer Music Research, University of Plymouth, United Kingdom

[1]aurelien.antoine@postgrad.plymouth.ac.uk, [2]eduardo.miranda@plymouth.ac.uk

Computer Generated Orchestration: Towards Using Musical Timbre in Generative Composition

ABSTRACT

Timbre is a musical attribute that has been largely discussed among the research community. However, there is still a lot to investigate, especially in regards to timbre and orchestration, which involves polyphonic timbre: a phenomenon that emerges from the mixture of instruments playing together. In this paper, we report on the development of a system capable of automatically analysing and classifying perceptual qualities of timbre within orchestral audio samples. This approach has been integrated in a computing system designed to combine string instruments matching specific timbral qualities. Our rationale for developing such a system is to create a means of incorporating musical timbre in combinatorial rules, which is often focused mainly on traditional Western music theory. Such developments could enrich creative music system.

Background

Timbre is an important property of music. Since the first experiments by von Helmholtz (2013), researchers have produced a large and diverse quantity of work on this complex attribute, both from an artistic and scientific viewpoint (Schouten 1968; Grey 1975; Barrière 1991; Siedenburg *et al*. 2016). Timbre is traditionally defined as 'that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar' (American National Standards Institute 1973). Several works have demonstrated the importance of acoustic features in defining musical timbre (Schouten 1968; Grey 1975; Zwicker *et al*. 2013).

Timbre is also an important characteristic in musical orchestration. Writing for an ensemble of instruments offers a large range of unique timbres, emerging from the mixture of instruments playing together. This phenomenon, sometimes named polyphonic timbre (Aucouturier 2006; Alluri *et al*. 2010), remains an important element to investigate, especially for computer-aided composition tools.

Aims and Repertoire Studied

This paper reports on the development of a system capable of automatically analysing and classifying specific timbral qualities within orchestral audio samples. Here, we have decided to use verbal descriptors, such as brightness or roughness, to represent the perception of musical timbres. The rationale for using words of the everyday language is to make the tool accessible to non-acoustics experts.

Furthermore, the automatic timbral classification method has been implemented in a computing system capable of generating combinations of string instruments based on timbral properties defined by the user. Our rationale for developing such a system is to create a means of incorporating timbre in the rules for combining instrument notes.

Methods

For our automatic classification system, we experimented with five verbal descriptors: breathiness, brightness, dullness, roughness, and warmth. Further information about our acoustic feature analysis method can be found in a previous publication (Antoine *et al*. 2016).

There is no agreed metrics for classifying audio samples according to perceived responses of timbre quality. Therefore, we established a comparative scale for each timbre attribute implemented in the system. We collected over 250 audio recordings of various orchestral pieces, such as Beethoven's *5th Symphony*, Vivaldi's *Four Seasons*, and Debussy's *Suite Bergamasque*. Each audio recording has been split into 1, 2, 3 and 4 second long audio samples, which resulted in analysing and retrieving timbral data for over 236,000 audio files. Combined with statistical information about the dataset, the analysis enabled us to establish a scale for each attribute, and, thus, be able to normalise the data among the five verbal descriptors. Scales are continually calibrated as new audio files are analysed.

The comparative scale enabled us to input the normalised data into a machine learning method designed to automatically classify audio files. We chose to implement a Support Vector Machine (SVM) algorithm (Cortes *et al*. 1995). We created the SVM's corpus training set with 250 samples for each attribute — 1250 labelled samples in total. After retrieving acoustic features from the audio file, values for each verbal descriptor are analysed and then classified by the SVM. In addition to the standard classification, users are able to train the machine learning algorithm by listening and rating a selection of audio files, and, therefore, personalise the classification according to their own musical perception.

The automatic timbral classification method has been implemented in a computing system designed to combine string instruments. This system integrates a knowledge-based algorithm, designed to create chords using string instruments. The generation of chords combination is based on rules from traditional Western music theory. Further combinatorial rules can be implemented using different compositional frameworks or be created by the user.

In order to guide the creation of chord combinations, the user can specify different musical parameters, such as type of chords, name and number of instruments. Then, the user can select the desired perceptual quality using the verbal descriptors. The system uses audio samples of instrumental notes to generate an audio file of the chord combination, which is subsequently analysed in order to retrieve its timbral qualities. Then, it evaluates if the combination fits the timbral criteria

before it is output. Once the user finds a solution that matches their criteria, they can export the musical excerpts generated by the system as an audio file or as a musical score.

Implications

The audio analysis of orchestral recordings enabled us to define a scale for each verbal descriptor. It also facilitated the creation of a normalisation algorithm, which enables us to compare the data between all timbral attributes. We used the normalised dataset to train a machine learning algorithm capable of automatically classify orchestral audio files according to their polyphonic timbre content. Furthermore, users can personalise the classification by listening and rating a selection of audio files.

We incorporated this automatic timbre classification method into a system designed to generate combinations of string instruments. The generation is based on chord combination rules taken from the traditional Western music theory framework. An option to extend the combinatorial rule is also available. Here, the important criteria to guide the generative process is the timbral quality, and not solely pitches or scales. Thus, solutions are output only if the resulted audio combination conveys the selected timbral attribute using the defined list of instruments.

Such developments could enrich computer-aided orchestration systems by harnessing perceptual aspects of polyphonic timbre within orchestral sound.

Keywords

Timbre, Musical Acoustics, Musical Perception, Orchestration, Computer-Aided Orchestration.

REFERENCES

Alluri, Vinoo, and Toiviainen, Petri, 2010. ‘Exploring Perceptual and Acoustical Correlates of Polyphonic Timbre’, in *Music Perception: An Interdisciplinary Journal 27*/3: 223–42.

American National Standards Institute, 1973. *Psychoacoustic Terminology S3:20*. New York (NY): American National Standards Institute.

Antoine, Aurélien, Williams, Duncan, and Miranda, Eduardo R., 2016. ‘Towards a Timbral Classification System for Musical Excerpts’, in *Proceedings of the 2nd AES Workshop on Intelligent Music Production*. London: Queen Mary University of London. (<http://c4dm.eecs.qmul.ac.uk/events/wimp2/>, accessed 21/03/2023.)

Aucouturier, Jean-Julien, 2006. *Dix expériences sur la modélisation du timbre polyphonique*. PhD diss. Paris: Université Paris 6.

Barrière, Jean-Baptiste, 1991. *Le timbre: métaphore pour la composition*. Paris: Christian Bourgois.

Cortes, Corinna, and Vapnik, Vladimir, 1995. ‘Exploring Perceptual and Acoustical Correlates of Polyphonic Timbre’, in *Machine Learning 2*0/3: 273–97.

Grey, John M., 1975. *An Exploration of Musical Timbre*. PhD diss. Stanford (CA): Stanford University.

Helmholtz, Hermann, 2013. *On the Sensations of Tone*. North Chelmsford (MA): Courier Corporation.

Schouten, Jan Frederik, 1968. ‘The Perception of Timbre’, in *Reports of the 6th International Congress on Acoustics* 76: 10.

Siedenburg, Kai, Fujinaga, Ichiro, and McAdams, Stephen, 2016. ‘A Comparison of Approaches to Timbre Descriptors in Music Information Retrieval and Music Psychology’, in *Journal of New Music Research* 45/1: 27–41.

Zwicker, Eberhard, and Fastl, Hugo, 2013. *Psychoacoustics: Facts and models*. Berlin: Springer.