ABSTRACT

In the field of sound and music computing, only a handful of studies are concerned with the pursuit of new musical knowledge. There is a substantial body of *corpus analysis* research focused on new musical insight, but almost all of it deals with symbolic data: scores, chords or manual annotations. In contrast, and despite the wide availability of audio data and tools for audio content analysis, very little work has been done on the corpus analysis of audio data.

This thesis presents a number of contributions to the scientific study of music, based on *audio corpus analysis*. We focus on three themes: audio description, corpus analysis methodology, and the application of these description and analysis techniques to the study of music similarity and 'hooks'.

On the theme of audio description, we first present, in part i, an overview of the audio description methods that have been proposed in the music information retrieval literature, focusing on timbre, harmony and melody. We critically review current practices in terms of their relevancy to audio corpus analysis. Throughout part ii and iii, we then propose new feature sets and audio description strategies. Contributions include the introduction of *audio bigram* features, pitch descriptors that can be used for retrieval as well as corpus analysis, and *second-order audio features*, which quantify distinctiveness and recurrence of feature values given a reference corpus.

On the theme of audio corpus analysis methodology, we first situate corpus analysis in the disciplinary context of music information retrieval, empirical musicology and music cognition. In part i, we then present a review of audio corpus analysis, and a case study comparing two influential corpus-based investigations into the evolution of popular music [122, 175]. Based on this analysis, we formulate a set of nine recommendations for audio corpus analysis research. In part ii and iii, we present, alongside the new audio description techniques,

new analysis methods for the study of song sections and within-song variation in a large corpus. Contributions on this theme include the first use of a probabilistic graphical model for the analysis of audio features.

Finally, we apply new audio description and corpus analysis techniques to address two research problems of the COGITCH project of which our research was a part: improving audio-based models of music similarity, and the analysis of hooks in popular music. In parts i and ii, we introduce *soft audio fingerprinting*, an umbrella MIR task that includes any efficient audio-based content identification. We then focus on the problem of scalable cover song detection, and evaluate several solutions based on audio bigram features. In part iii, we review the prevailing perspectives on musical catchiness, recognisability and hooks. We describe Hooked, a game we designed to collect data on the recognisability of a set of song fragments. We then present a corpus analysis of hooks, and new findings on what makes music catchy.

Across the three themes above, we present several contributions to the available methods and technologies for audio description and audio corpus analysis. Along the way, we present new insights into choruses, catchiness, recognisability and hooks. By applying the proposed technologies, following the proposed methods, we show that rigorous audio corpus analysis is possible and that the technologies to engage in it are available.