AN EVALUATION FRAMEWORK AND CASE STUDY FOR RHYTHMIC CONCATENATIVE SYNTHESIS

Cárthach Ó Nuanáin, Perfecto Herrera, Sergi Jordà
Music Technology Group
Universitat Pompeu Fabra
Barcelona
{carthach.onuanain, perfecto.herrera, sergi.jorda}@upf.edu

ABSTRACT

In this paper we present and report on a methodology for evaluating a creative MIR-based application of concatenative synthesis. After reviewing many existing applications of concatenative synthesis we have developed an application that specifically addresses loop-based rhythmic pattern generation. We describe how such a system could be evaluated with respect to its its objective retrieval performance and subjective responses of humans in a listener survey. Applying this evaluation strategy produced positive findings to help verify and validate the objectives of our system. We discuss the results of the evaluation and draw conclusions by contrasting the objective analysis with the subjective impressions of the users.

1. INTRODUCTION

MIR-based applications are becoming increasingly widespread in creative scenarios such as composition and performance [14] [7] [8]. This is commensurate with the prevalence of sampling-based approaches to sound generation, thus the desire is to develop more rich and descriptive understanding of the underlying content being used.

One of the primary difficulties faced with designing instruments for creative and compositional tasks remains the elaboration of an appropriate evaluation methodology. Indeed, this is a trending challenge facing many researchers [2], and numerous papers address this directly with various proposals for methodological frameworks, some drawing from the closely related field of HCI (Human Computer Interaction) [13], [16], [11]. More generally the evaluation of computer composition systems has also been the subject of much discussion in the literature. One frequent benchmark for evaluating algorithmic music systems is a type of Turing test where the success criterion is determined by the inability of human listener to discern between human and computer-generated music. As Hiraga [11] notes, however, these kind of tests can be problematic for two reasons. Firstly, it makes the assumption that the music generated by the algorithm is intended to sound like music produced by humans, rather than something to be treated differently. Secondly it ignores other facets of the system that imperatively needs evaluation, such as the interface and the *experience*. Pachet also finds issue with simplistic Turing test approaches to music evaluation [18]. He repeats, for instance, the view that unlike the traditional Turing test which evaluated the ability to synthesis believable natural language, no such “common-sense” knowledge exists for aspects of music.

We have designed and developed an MIR-driven instrument that uses concatenative synthesis to generate looped rhythmic material from existing content. In terms of evaluation we face the challenge of evaluating an MIR driven software system, thus subject to the same scrutiny facing any information retrieval system that needs to be appraised. We also face the challenge of evaluating the system as a musical composition system that needs to serve the composer and listener alike.

In the next section we will give the reader brief familiarity with the instrument in terms of its implementation and functionality. Subsequently, existing concatenative systems will be reported on in terms of their evaluation methodologies (if any). Section 3 will propose the evaluation framework in questions and the results will be reported. We will conclude the paper with our impressions on what we have learnt and scope for improvement in terms of the system itself and the evaluation methodology.

2. INSTRUMENT DESCRIPTION

Concatenative synthesis builds new sounds by combining together existing ones from a corpus. It is similar to granular synthesis differing only in the order of size of the grains: granular synthesis operates on microsound scales of 20-200ms whereas concatenative synthesis uses musically relevant unit sizes such as notes or even phrases. The process by which these sounds are selected for resynthesis is a fundamentally MIR-driven task. The corpus is defined by selecting sound samples, optionally segmenting them into onsets and extracting a chosen feature set to build descriptions of those sounds. New sounds can finally be synthesised by selecting sounds from the corpus according to