

Técnicas computacionais para sumarização de áudio

Felipe Silva Felix
Orientador: Marcelo Queiroz

15 de Agosto de 2016

1 Introdução

Esse projeto de Iniciação Científica pretende abordar o seguinte problema: encontrar o segmento mais representativo de uma gravação musical, tarefa frequentemente referida como “*audio thumbnailing*”.

2 Objetivos

Esse trabalho tem como objetivo estudar a teoria por trás das principais técnicas de sumarização de gravações musicais, isto é, encontrar o segmento de áudio que mais se repete numa dada gravação, o que também envolverá análises estruturais de gravações musicais. Também temos por objetivo implementar três técnicas e comparar seus resultados analisando suas diferenças e similaridades através de alguns experimentos comparativos.

3 Metodologia

Nesse trabalho iremos estudar a técnica desenvolvida por Müller *et al.* [6][7], entender as relações da estrutura musical com a matriz de similaridade, explorar espaços de características e sua relação com o resultado obtido. Também estudaremos a abordagem elaborada por Bartsch *et al.* [1][2], baseada num espaço de características derivado do cromagrama. Tais análises por similaridade serão recorrentes nesse trabalho, que também são discutidas no trabalho de Cooper e Foote [3][4]. Além disso, teremos como referência de conceitos que serão estudados no semestre um livro também de Müller [5].

3.1 Cronograma

Pretendemos cumprir o seguinte cronograma:

- **Agosto - Setembro:** Leitura da bibliografia indicada.
- **Setembro - Outubro:** Implementações das técnicas estudadas.
- **Outubro - Novembro:** Experimentos e análises dos resultados.

4 Acompanhamento

O progresso do projeto poderá ser acompanhado através do link <https://fsfelix.github.io> e apresentará relatórios semanais contendo os avanços do projeto.

Referências

- [1] Mark A Bartsch and Gregory H Wakefield, *To catch a chorus: Using chroma-based representations for audio thumbnailing*, Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the, IEEE, 2001, pp. 15–18.
- [2] ———, *Audio thumbnailing of popular music using chroma-based representations*, IEEE Transactions on multimedia **7** (2005), no. 1, 96–104.
- [3] Matthew Cooper and Jonathan Foote, *Summarizing popular music via structural similarity analysis*, Applications of Signal Processing to Audio and Acoustics, 2003 IEEE Workshop on., IEEE, 2003, pp. 127–130.
- [4] Matthew L Cooper and Jonathan Foote, *Automatic music summarization via similarity analysis.*, ISMIR, 2002.
- [5] Meinard Müller, *Fundamentals of music processing: Audio, analysis, algorithms, applications*, Springer, 2015.
- [6] Meinard Müller, Peter Grosche, and Nanzhu Jiang, *A segment-based fitness measure for capturing repetitive structures of music recordings.*, ISMIR, Citeseer, 2011, pp. 615–620.
- [7] Meinard Muller, Nanzhu Jiang, and Peter Grosche, *A robust fitness measure for capturing repetitions in music recordings with applications to audio thumbnailing*, IEEE Transactions on audio, speech, and language processing **21** (2013), no. 3, 531–543.