

Proposta de trabalho - MAC0499

Um estudo de formatos de codificação de audio e de seus respectivos codecs

Guilherme Simões Santos Marin

Orientador: Marcelo Queiroz

Introdução

Os formatos de codificação de audio (ou compressão de audio) são especificações técnicas que definem a forma como um sinal de audio pode ser representado digitalmente. Frequentemente, estes formatos são construídos com a intenção de reduzir a quantidade de bits necessária para armazenar e transmitir o sinal digital [1]. Já um codec de audio é uma implementação em software ou hardware dos algoritmos que efetuam a codificação e decodificação de um sinal digital de acordo com um dado formato de codificação [2]. Entre as aplicações dessas tecnologias, podemos citar o seu uso em streaming de audio e vídeo [3] e nas indústrias de radiodifusão, cinema e de produtos de multimídia [4].

Dentro desse contexto, o presente trabalho tem como proposta o estudo dos principais formatos de codificação de audio da atualidade, seguido da análise experimental de pelo menos um codec de audio para cada formato escolhido. Inicialmente, será feita uma análise da estrutura e dos algoritmos de compressão empregados em cada formato. Posteriormente, será feito um levantamento em relação aos codecs de audio mais importantes em cada formato e estes serão analisados empiricamente, por meio da sua aplicação em um conjunto de sinais

de audio representativos. Entre os formatos que serão analisados, estão o MP3 (formalmente “MPEG-1 Audio Layer III” ou “MPEG-2 Audio Layer III”), o AAC (“Advanced Audio Coding”) e o FLAC (“Free Lossless Audio Codec”), dentre outros.

Objetivos

Esse trabalho busca oferecer uma análise ampla de vários dos principais formatos de codificação de audio da atualidade. Dessa forma, espera-se que o estudo possa servir como um referência para a escolha da melhor técnica de codificação para um sinal de audio qualquer, levando em conta as características do sinal a ser codificado, os diferentes formatos e codecs disponíveis, e também considerando as situações concretas em que o conteúdo de audio será consumido por um ouvinte e/ou armazenado. A análise experimental do desempenho de cada codec pode ser particularmente útil, já que frequentemente ter acesso apenas à descrição técnica de cada codec ou de cada formato não é suficiente para se ter uma boa indicação sobre qual a melhor solução de codificação a ser utilizada, visto que, principalmente no caso de compressão com perda de dados, em que existe degradação da qualidade do audio original, a escolha da solução depende também de aspectos perceptuais do resultado final [5].

Metodologia

Inicialmente será realizada uma seleção dos formatos de codificação a serem estudados. Essa seleção será baseada nos seguintes critérios: a popularidade e importância do formato; as características do formato, como por exemplo se ele realiza compressão com ou sem perda de dados; e a disponibilidade pública de codecs de audio para o formato. Em seguida, será realizado um estudo de cada formato e dos seus pressupostos teóricos de forma a elucidar o seu funcionamento interno. Esse estudo teórico inicial é importante para elucidar quais são os pontos fortes e fracos de cada formato e também para se ter uma visão clara sobre quais

pressupostos embutidos nos modelos empregados podem afetar o resultado da codificação e de que forma esses pressupostos podem ser explicitados na subsequente análise empírica dos codecs. Após isso, o conteúdo de audio a ser analisado e os codecs a serem testados serão escolhidos. A escolha dos sinais de audio será baseada no estudo dos formatos feito na fase anterior, já que o objetivo principal dos testes é verificar empiricamente e perceptualmente as propriedades de cada formato que foram descritas nesse estudo. Além disso, como podem existir diferenças de resultado que são consequência de escolhas técnicas realizadas pelos implementadores de cada codec de audio, vamos tentar escolher pelo menos dois codecs diferentes para cada formato analisado, dependendo da disponibilidade de acesso público a eles em cada caso. Por último, os testes serão realizados, usando o conteúdo de audio e os codecs selecionados, e os resultados serão discutidos. Nessa última fase, também será necessário descrever claramente como podemos analisar a qualidade do audio resultante usando critérios científicos e objetivos, e quais parâmetros dos resultados serão comparados.

Cronograma

Atividade	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov
1	X	X							
2	X	X							
3		X	X	X					
4			X	X	X	X			
5					X	X			
6						X	X		
7							X	X	X

Legenda das atividades:

1. Levantamento acerca dos codecs de audio disponíveis para cada formato

2. Escolha dos formatos a serem estudados
3. Revisão da literatura existente acerca de compressão de audio
4. Estudo dos formatos, dos seus algoritmos e dos modelos que fundamentam o seu funcionamento
5. Escolha do conjunto de sinais de audio a serem testados
6. Realização dos testes
7. Análise e discussão dos resultados dos testes

Referências

- [1] T. Painter and A. Spanias, “A review of algorithms for perceptual coding of digital audio signals,” in *Proceedings of 13th International Conference on Digital Signal Processing*, vol. 1, pp. 179–208 vol.1, 1997.
- [2] Y.-C. Wu, I. D. Gebru, D. Marković, and A. Richard, “Audiodec: An open-source streaming high-fidelity neural audio codec,” in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5, 2023.
- [3] N. Zeghidour, J. Copet, Y. Adi, A. Mohamed, M. J. Romero, and D. Grangier, “High fidelity neural audio compression,” 2022. arXiv preprint arXiv:2210.13438.
- [4] A. Gersho, “Advances in speech and audio compression,” *Proceedings of the IEEE*, vol. 82, no. 6, pp. 900–918, 1994.
- [5] K. Sayood, *Introduction to data compression*. Morgan Kaufmann, 2017.