

Compressão de Áudio Digital

Aluno: Marcio Masaki Tomiyoshi

Supervisor: Marcelo Gomes de Queiroz

Introdução

O trabalho consistiu do estudo de diversas técnicas de compressão de áudio, tanto com perdas (*lossy*) quanto sem perdas (*lossless*).

O exemplo mais popular de compressão de áudio é o padrão MP3 (*lossy*), que, ao descartar informações inaudíveis aos seres humanos, é capaz de manter uma qualidade próxima a de um CD com um arquivo de cerca de 10% do tamanho original [1].

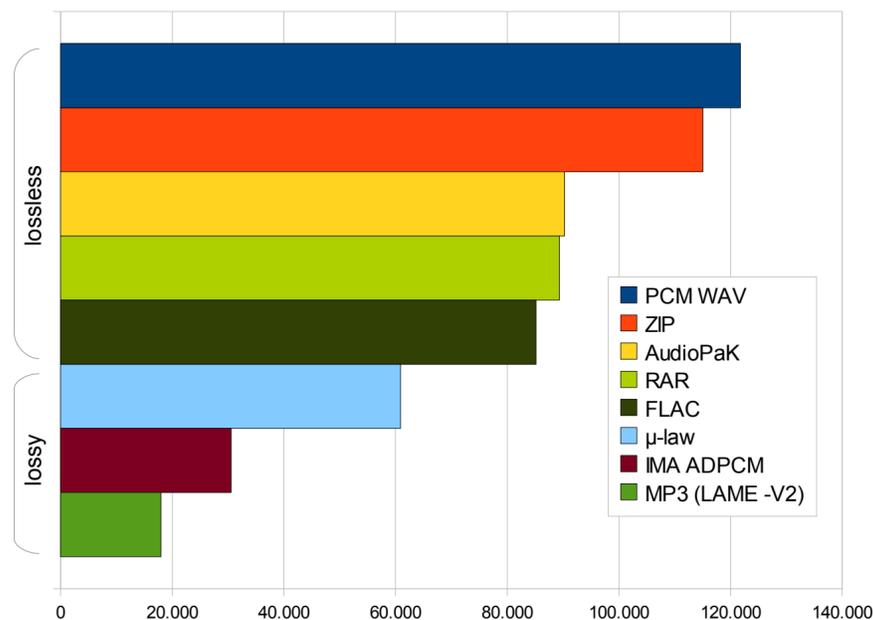
Porém, dependendo do uso, a informação original precisa ser mantida para não resultar em uma degradação exageradamente grande na qualidade. Nestes casos, compressores *lossless* são mais indicados. Um dos formatos mais populares é o FLAC, que atinge, em média, cerca de 55% do tamanho original [2].

Compressão Sem Perdas

Neste caso, tanto o codificador quanto o decodificador foram desenvolvidos [3] baseando-se no formato AudioPaK [4].

Ao invés de armazenar os valores individuais das amostras, como faz o padrão PCM, o AudioPaK armazena a diferença entre o valor amostrado e um valor estimado a partir das amostras anteriores.

Se o estimador utilizado para prever o comportamento da onda sonora desejada for eficiente, o valor armazenado pelo AudioPaK é menor que o valor individual da amostra, possibilitando que sejam utilizados menos bits



Resultado das implementações do AudioPaK, μ -law e IMA ADPCM comparado a outros formatos (em kilobytes).

em sua representação.

O código usado para representar as amostras com uma taxa variável de bits é a codificação de Golomb [4].

Compressão Com Perdas

Foram implementados os padrões μ -law e IMA ADPCM, e foi estudado o padrão MPEG-1.

• μ -law

Foi desenvolvido um codificador que se adequa ao formato WAVE. O formato usa uma escala logarítmica na representação das amostras, ao invés de linear como no PCM. Para cada amostra são utilizados 8 bits.

• IMA ADPCM

Utiliza-se de uma idéia semelhante à aplicada no compressor *lossless*, armazenando apenas a diferença entre as amostras, mas neste formato é permitido que a diferença não seja representada fielmente, ocasionando as perdas em relação ao áudio original. Cada amostra utiliza 4 bits.

• MPEG-1

O padrão MPEG-1 para áudio tem três formatos definidos, *layers I, II e III*. O mais popular de todos é o último, que corresponde ao MP3.

A cada *layer*, técnicas mais avançadas de análise e representação vão sendo usadas. Todas são baseadas em dividir o espectro de frequência do som em subbandas e analisar, através de um modelo psicoacústico, quais delas podem ter mais informação descartada de forma que o ouvido humano não consiga perceber a diferença em relação ao sinal original.

Referências

- [1] P. Noll (1997). MPEG digital audio coding.
- [2] FLAC – comparison. <http://flac.sourceforge.net/comparison.html> (acessado em 08/11/2008)
- [3] <http://www.linux.ime.usp.br/~masaki/mac499/>
- [4] M. Hans, R. W. Schafer (2001). Lossless compression of digital audio.