



Universidade Federal de Santa Catarina
Centro Tecnológico
Departamento de Informática e Estatística
Ciência da Computação
INE5431-07208 - Sistemas Multimídia



Relatório - Prática 1

Guilherme Augusto Oliveira Pedrozo (22100621)
Julia Macedo de Castro (23250860)
Victor Henrique Labes de Figueiredo (22200378)

Florianópolis
2024

Questões:

A) Em relação ao arquivo original, indique no relatório:

- i) O tamanho teórico do áudio (parte de dados);
- ii) O tamanho do arquivo em bytes (ver propriedades do arquivo, ou Linux utilize o comando "ls -l audio20232.wav) e indique o motivo da diferença entre este tamanho e o calculado em i).
- iii) O tamanho do arquivo em disco em bytes, observando as propriedades do arquivo, ou no linux utilize "du -s -B1 audio.wav", e indique o motivo da diferença entre este tamanho e o tamanho do arquivo em ii).
- iv) Indique o tamanho deste arquivo em disco se o seu HD fosse formatado para um tamanho de bloco (unidade de alocação em disco) de 1024 B.
- v) Indique o maior componente de frequência (em Hertz) possível quando estes parâmetros de digitalização são utilizados.

B) Exporte o áudio de A) em um novo arquivo em formato WAV (Microsoft), usando como taxa de amostragem (Sample Rate) de 16000Hz e codificação Signed 16-bits PCM. Depois de exportar o arquivo, é necessário abri-lo no Audacity. Finalmente, indique no relatório:

- i) o tamanho da parte de dados do novo áudio;
- ii) o valor máximo de componente de frequência teórico para o novo formato do áudio;
- iii) Visualizar os canais do som no software Audacity e compare com o áudio original. Qual o efeito ocorreu no áudio gerado na etapa B) e explique porque isto ocorreu?

C) Exporte o áudio de B) em um novo arquivo em formato WAV (Microsoft), usando como taxa de amostragem (Sample Rate) de 8000Hz e codificação Unsigned 8-bits PCM. Depois de exportar o arquivo, é necessário abri-lo no Audacity. Finalmente, indique no relatório::

- i) o tamanho da parte de dados do novo áudio;
- ii) Indique o maior componente de frequência (em Hertz) possível quando estes parâmetros de digitalização são utilizados;
- iii) explique os efeitos no áudio ocasionado pela redução do número de bits por amostra. Ouça o áudio, principalmente nos períodos de silêncio.

Respostas:

QUESTÃO A:

i)

Tamanho em bits = taxa de amostragem * bits por amostra * nº de canais * duração em segundos

Tamanho em bytes = Tamanho em bits/8

- Nº de canais: 2;
- Taxa de amostragem: 44.1kHz = 44.100Hz;
- Bits por amostra: 16 bps;
- Duração em segundos: 30;
- Tamanho em bits (audio20242.wav) = $44.100 * 16 * 2 * 30 = 42.336.000$ bits;
- Tamanho em bytes = Tam. em bits (audio20242.wav)/8 = 5.292.000 bytes = 5,292MB (base decimal) \sim 5,04 MB (base binária).

ii) Tamanho teórico dos dados: 5.292.000 bytes (5.292 MB);

Tamanho real do arquivo (com o comando 'ls -l'): 5.292.044 bytes;

Esses 44 bytes há mais incluso no tamanho real do arquivo, se dá por conta do 'header (cabeçalho)' que contém informações do próprio arquivo, como a frequência de amostragem, formato do arquivo e afins.

iii) Tamanho real do arquivo: 5.292.044 bytes;

Tamanho do arquivo (com o comando 'du -s -B1 audio20242.wav'): 5.296.128 bytes;

O tamanho do arquivo com esse comando, mostra o espaço em disco efetivamente utilizado pelo arquivo, o que pode incluir a fragmentação e o preenchimento das alocações de blocos do sistema de arquivos, cujo blocos podem variar de 512 bytes a 4 KB ou mais. Mesmo que o arquivo não use todo o espaço de um bloco, o sistema reserva esse espaço.

iv) Número de blocos = $\lceil \text{tamanho real do arquivo} / 1024 \rceil = \lceil 5.292.044 / 1024 \rceil = \lceil 5.168,01171875 \text{ blocos} \rceil = 5.169 \text{ blocos}$

- Tamanho do arquivo em disco = número de blocos * 1024;
- Tamanho do arquivo em disco (blocos de 1024 bytes): 5.293.056 bytes.

v) Pelo Teorema de Nyquist, a frequência máxima pode ser representada metade da taxa de amostragem, ou seja:

- Frequência Máxima = Taxa de amostragem/2

Logo, o maior componente de frequência (arquivo original) é: $44.100 \text{ Hz} / 2 = 22.050 \text{ Hz} = 22,05 \text{ kHz}$

QUESTÃO B:

i)

Tamanho em bits = taxa de amostragem * bits por amostra * nº de canais * duração em segundos;

Tamanho em bytes = Tamanho em bits/8;

- Nº de canais: 2;
- Taxa de amostragem: 16kHz = 16.000Hz;
- Bits por amostra: 16 bps;
- Duração em segundos: 30;
- Tamanho em bits: $16.000 * 16 * 2 * 30 = 15.360.000$ bits;
- Tamanho em bytes (sem 'header'): $15.360.000/8 = 1.920.000$ bytes;
- Tamanho em bytes real (com 'header'): 1.920.044 bytes = 1,92 MB (base decimal) \approx 1,83 MB (base binária);

ii) Pelo Teorema de Nyquist, a frequência máxima pode ser representada metade da taxa de amostragem, ou seja:

- Frequência Máxima = Taxa de amostragem/2

Logo, o maior componente de frequência deste áudio é: $16000 \text{ Hz}/2 = 8000 \text{ Hz} = 8\text{kHz}$.

iii) Após a análise, nota-se a perda de sons agudos, devido à perda de qualidade e à redução da taxa da amostragem, que limita as frequências mais altas. Antes da conversão, o áudio original pode representar frequências até 22,05 kHz, após tal processo, o áudio pode representar somente frequências de até 8kHz. Contudo, qualquer frequência acima de 8kHz será excluída, resultando na perda significativa de agudos.

QUESTÃO C:

i)

Tamanho em bits = taxa de amostragem * bits por amostra * nº de canais * duração em segundos;

Tamanho em bytes = Tamanho em bits/8;

- Nº de canais: 2;
- Taxa de amostragem: 8kHz = 8.000Hz;
- Bits por amostra: 8 bps;
- Duração em segundos = 30;
- Tamanho em bits: $8000 * 8 * 2 * 30 = 3.840.000$ bits;
- Tamanho em bytes (sem 'header'): $3.840.000/8 = 480.000$ bytes;
- Tamanho em bytes real (com 'header'): 480.000 bytes = 480KB (base decimal) \approx 468KB (base binária);

ii) Pelo Teorema de Nyquist, a frequência máxima pode ser representada metade da taxa de amostragem, ou seja:

$$\text{- Frequência Máxima} = \text{Taxa de amostragem}/2$$

Logo, o maior componente de frequência deste áudio é: $8000\text{Hz}/2 = 4000\text{Hz} = 4\text{kHz}$.

iii) Por conta da redução do número de bits por amostra, acarretou na redução da qualidade do áudio e mais perdas de agudos, pois tal consegue representar frequências de até 4kHz.