

EN 1023

$$\textcircled{1} \quad H_i = \left\lfloor \frac{30}{60} \right\rfloor = 0$$

$$f = 0,5 - 0 = 0,5$$

$$p = 0,25 (1 - 0,5) = 0,125$$

$$q = 0,25 (1 - 0,5 \times 0,5) = 0,1875$$

$$t = 0,25 (1 - ((1 - 0,5)0,5)) = 0,4875$$

$$R = 0,25 \times 255 \approx 64$$

[64, 48, 32]

$$G = 0,1875 \times 255 \approx 48$$

$$B = 0,125 \times 255 \approx 32$$

$$\textcircled{2} \quad \frac{4 \times 3840 \times 2160}{10} \approx 3,3 \text{ MB}$$

Mais próximo é 2,5 MB

③

$$DC = F(0,0) = \frac{8 \times 8 \times 128}{8} = 1024$$

$$AC = 0$$

④

PNG L277

⑤

$$\text{dur} = 220 \text{ s}$$

$$\text{size} = 2,64 \text{ MB} = 21,12$$

$$\text{size} = \text{br} \times \text{dur} \Rightarrow \text{br} = \frac{21,12}{220} = 0,096 \text{ Mbps} =$$

$$96 \text{ Kbps}$$

⑥ MDCT

⑦ Torna-se mais perceptível ruído silencioso

⑧ Codificação de resíduos

⑨ 2160p

⑩ P

⑪ Nenhum

⑫

I B B, P B B, P B B, P B B, P B B

↑
4

$$\frac{4}{15} = 0,267$$

26,7%

⑬ $f_s = 44.1 \text{ KHz}$

40 filtros

lower limit = $0 \text{ Hz} = 0 \text{ mel}$

$$\text{upper limit} = \frac{44100 \text{ Hz}}{2} = 22050 \text{ Hz} =$$

$$= 2595 \log_{10} \left(1 + \frac{22050}{700} \right) \text{ mel} = 3923,34 \text{ mel}$$

Os 40 filtros estão espaçados linearmente na escala de mel. A posição do centro do filtro K (mel)

é:

$$40 \left| \frac{3923,34}{40+1} \right| = 3827,649 \text{ mel}$$

← meio do first

$$f = 700 \left(10^{\frac{3827,649}{2595}} - 1 \right) \approx 20198$$

(14)

$$\text{sum} = 300 + 100 + 100 + 40 + 30 + 10 = 680$$

$$\text{sum} = 578$$

$$\text{cum sum/bin} = 600$$

$$\Delta f = \frac{f_s}{2 \#B} = 300 \text{ Hz}$$

$$\text{rolloff} = 2 \times 300 \text{ Hz} = 600 \text{ Hz}$$

(15)

9300 totais

145 relevantes totais

172 recuperadas

127 relevantes

$$\text{Precision} = \frac{127}{172} = 0,738$$

$$F = \frac{2 P_r R_e}{P_r + R_e} = 0,801 \approx 80,13 \%$$

$$\text{Recall} = \frac{127}{145} = 0,8758$$

(16) Sim, pode haver downsampling.

(17) Cada implementação de mp3 é diferente, muitos parâmetros podem mudar.

e.g. # pontos p/ cálculo da MDCT

Diferentes modelos psicoacústicos

(18) LZ77 é melhor para dados com alta correlação entre amostras vizinhas. Isso não se verifica no áudio.

Golomb-Rice é melhor para um alfabeto de resíduos quase-geométrico

① Experimentamente o melhor balanceamento entre compressão e qualidade

② Mais próximas da percepção humana mas mais complicadas de extrair e sequer definir computacionalmente
e.g. Artista, género, emoção

③ Cada membro avaliou as sugestões numa escala de 1 a 5, considerámos músicas cuja média era superior a 2,5 como relevantes.

$$p = \frac{\# \text{relevantes}}{\# \text{sugestões totais}}$$

A disparidade entre avaliadores aumentou o std.