

Nome:  
Nº:



Departamento de Engenharia Informática  
Faculdade de Ciências e Tecnologia  
Universidade de Coimbra

## Multimédia (LEI)

20 de Junho de 2025

Exame Normal

Duração: 2h

### Notas prévias:

- 1) Consulta permitida: consulta fornecida ou slides das aulas teóricas (impressos e sem quaisquer notas, excepto marcadores de índice e highlights).
- 2) É permitido o uso de calculadora (não programável).
- 3) Qualquer tentativa de fraude conduzirá à anulação da prova para todos os intervenientes e activação do procedimento disciplinar da Universidade de Coimbra.
- 4) Escolha múltipla: as respostas **erradas subtraem 25%** da cotação da pergunta.
- 5) As cotações das questões poderão sofrer alterações ligeiras para beneficiar a maioria dos alunos.

1. (5%) Uma imagem fotográfica *true color*, sem canal alfa, gravada no formato PNG tem a dimensão de 1.4 MB. Qual a resolução provável da mesma?

☐ 640x480    ☐ 800x600    ☐ 1024x768    ☒ Standard HD    ☐ Full HD

2. (5%) Para representar a cor HSB,  $H = 57^\circ$ ,  $S = 65\%$ ,  $B = 55\%$ , no modelo de cor RGB (escala 0 a 255), que valores corresponderão a R, G e B?

☒ R = 140, G = 136, B = 49    ☐ R = 255, G = 49, B = 136    ☐ R = 255, G = 136, B = 53  
☐ R = 140, G = 102, B = 49    ☐ R = 140, G = 49, B = 136

3. (2.5%) Uma imagem foi comprimida através do codec GIF. O número máximo de cores distintas (**perceptuais**) que um utilizador pode observar na imagem é:

☐ 1    ☐ 8    ☐ 16    ☐ 256    ☒ >256

4. (5%) Considere uma imagem codificada através do codec JPEG. O conteúdo de um bloco 8x8 do canal Y é o apresentado abaixo. Qual o valor do coeficiente DC da DCT?

Y							
0	0	150	150	150	150	150	150
0	0	150	150	150	150	150	150
0	0	150	150	150	150	150	150
0	0	150	150	150	150	150	150
150	150	150	150	150	150	0	0
150	150	150	150	150	150	0	0
150	150	150	150	150	150	0	0
150	150	150	150	150	150	0	0

☐ 100    ☐ 300    ☐ 476.3    ☐ 600    ☒ 900

5. (2.5%) No codec **mp3**, a análise em frequência é feita em quantas bandas de frequência?

☐ 4    ☐ 8    ☐ 16    ☒ 32    ☐ 1152

6. (5%) Imagine que tem uma música PCM com qualidade de CD. Essa música é comprimida através do codec AAC com taxa de compressão 15:1, resultando um ficheiro com 3.6 MB. Qual a duração da música em causa? Considere 1K = 1000.
- ☐ 1m38s      ☐ 2m54s      ☐ 4m17s      ☒ 5m06s      ☐ 5m54s
7. (5%) Um ficheiro áudio PCM apresenta a seguinte sequência de amostras:  $x[n] = \{0, 1, 2, 4, 2, 1\}$ ,  $n = 0, 1, \dots, 5$ . Sabendo que: i) o ficheiro será codificado com o codec **FLAC**; ii) e o modelo de previsão utilizado é dado pela equação  $x_p[n+1] = \frac{1}{2}x[n] + \frac{1}{4}x[n-1] + \frac{1}{4}x[n-2]$  ( $x_p[n+1]$  é o valor previsto para a amostra  $x[n+1]$ ), quais os valores dos resíduos  $e[n]$ ,  $n = 3, 4, 5$  (por essa ordem)?
- ☒  $\frac{11}{4}, -\frac{3}{4}, -\frac{3}{2}$       ☐  $\frac{1}{2}, \frac{3}{4}, 2$       ☐  $\frac{11}{4}, -\frac{3}{4}, -\frac{3}{4}$       ☐  $0, -1, 1$       ☐  $\frac{1}{4}, \frac{3}{2}, 2$
8. (2.5%) Nas frames I do codec H.261, qual o esquema de sub-amostragem utilizado?
- ☐ 4:4:4      ☐ 4:2:2      ☒ 4:2:0
- ☐ Todas as anteriores      ☐ Nenhuma das anteriores
9. (2.5%) As frames I podem servir de referência a outras frames em que codec(s)? Escolha a melhor opção.
- ☐ MJPEG      ☐ MPEG-2 e MPEG-4      ☒ Em toda a família MPEG
- ☐ MPEG-1      ☐ MPEG-4 AVC
10. (2.5%) Na família de codecs MPEG, a análise para compensação de movimento é efectuada em que canal?
- ☐ P      ☐ Cb e Cr      ☐ B      ☒ Y      ☐ Y, Cb e Cr
11. (5%) Um vídeo a 30 fps, cuja estrutura do GOP (Group of Pictures) é determinada pelos parâmetros  $M = 1$  e  $N = 15$ , é comprimido através do codec MPEG-1. Qual o rácio de compressão esperado?
- ☐ 9.2 : 1      ☐ 13.4 : 1      ☒ 17.8 : 1      ☐ 20 : 1      ☐ 27 : 1
12. (5%) Uma base de dados (BD) contém um total de 10000 vídeos. Numa pesquisa baseada em exemplo, sabe-se que existem 340 vídeos relevantes para a query em causa. Na pesquisa realizada, foram recuperadas 254 vídeos, 202 dos quais relevantes. Qual o desempenho da pesquisa realizada em termos da métrica F-measure?
- ☐ 59.4%      ☒ 68.0%      ☐ 73.8%      ☐ 79.5%      ☐ 83.6%
13. (5%) Uma imagem apresenta as seguintes médias e desvio-padrões, (m, dp), para 3 features normalizadas no intervalo [0, 1]:  $F1 = (0.6, 0.34)$ ,  $F2 = (0.8, 0.22)$ ,  $F3 = (0.26, 0.46)$ . Assumindo que é definido um peso mínimo de 0.25 para utilização de features num sistema de MMIR, que features serão utilizadas usando o algoritmo de atribuição de peso pelo desvio padrão?
- ☐ F3      ☐ F1 e F2      ☐ F2 e F3      ☒ F1 e F3      ☐ todas
14. (5%) A magnitude útil do espectro de um sinal áudio, numa dada frame apresenta os seguintes valores: [0, 200, 100, 150, 70]. Assumindo  $f_s = 1200$  Hz, qual o spectral rolloff, em Hz?
- ☐ 240      ☒ 360      ☐ 480      ☐ 600      ☐ 720
15. (2.5%) mp3 está para WAV assim como:
- ☐ CorePNG está para MOV      ☐ MPEG-1 está para CorePNG      ☐ MPEG-4 AVC está para AVI
- ☒ MPEG-1 está para AVI      ☐ AVI está para CorePNG

16. (2.5%) Numa imagem JPEG, todos os píxeis de um dado bloco 8x8 são pretos. Sem fazer cálculos, indique o valor do coeficiente DC e do 1º coeficiente AC (varrimento em zig-zag) desse bloco. Justifique.

(10%)  $RGB = 0$  (tudo preto)  $\rightarrow Y=0$

(45%)  $DC = k \times \text{média da tonalidade} = k \times 0 = 0$

(45%)  $AC = 0$ , uma vez que não há variações de tonalidades

17. (5%) Devido a um número elevado de queixas, a empresa PhotoLQ, retirou do mercado o seu último produto: uma máquina fotográfica digital com profundidade de cor true color que armazena as fotografias no formato GIF. Analise a causa do insucesso, referindo o que poderia ter sido feito para o minorar (mas **mantendo o formato GIF**).

(50%) Causa: GIF apenas suporta 256 cores (35%), o que é muito inferior aos 16M cores possíveis com true color  $\rightarrow$  perda notória de qualidade, sobretudo se se usarem paletes estáticas (15%), como aparenta ser o caso.

(50%) Solução para minorar: Utilizar GIF com paletes adaptativas (15%), com as cores estatisticamente mais relevantes (10%), e dithering (15%), criando a ilusão de cores não presentes na paleta (10%)

Outros: espaço ocupado (10%)

18. (5%) A máquina fotográfica Panasonic Lumix DMC-ZX1 permite a gravação de pequenos filmes no formato MOV. Esses vídeos são comprimidos com compensação de movimento?

(50%) Não se sabe (15%), dado que o formato MOV pode utilizar vários codecs (35%).

(25%) Se for um codec da família MPEG, suporta.

(25%) Se for o codec MJPEG, não suporta (na verdade é este o codec, pelo que não suporta).

19. (7.5%) O codec mp3 tira partido do mecanismo perceptual de mascaramento simultânea de som. Explique sucintamente em que consiste, qual o modo de operação dessa funcionalidade no mp3 e qual a sua utilidade.

(40%) Mascaramento simultânea: um som de baixa intensidade na mesma banda crítica que um som de alta intensidade tende a não ser audível (ou audível com menos realce).

(30%) O mp3 tira partido desta funcionalidade reduzindo o número de bits de quantização dos coeficientes da MDCT.

(30%) Caso o ruído de quantização resultante esteja abaixo do limiar de mascaramento, o ruído não será ouvido. Assim, consegue-se comprimir o ficheiro, mantendo uma qualidade perceptualmente boa.

20. (10%) O Maestro Vítor de Almeida criou um ficheiro de 2 seg com a nota Lá tocada num diapasão (tom quase puro e com amplitude aproximadamente constante), com as seguintes características: som com qualidade de CD, monoaural e sem compressão. Em seguida, gravou o ficheiro em formato mp3 e AAC, com a mesma frequência de amostragem do original. Qual o impacto das características específicas dessa gravação na taxa de compressão alcançável, tendo em consideração os mecanismos de compressão possíveis em cada um dos codecs?

(35%) Tom puro:

- > Uma só frequência fundamental (não há harmónicos, ou têm amplitude quase nula)
- > não há mascaramento simultânea (por só haver um componente de frequência)

(30%) -> em cada janela, a MDCT apenas vai ter um pico (os restantes coeficientes serão nulos)  
-> muita compressão (praticamente não destrutiva), em qualquer dos casos

(35%) AAC: o método de Long Term Prediction é aplicado, pelo facto do sinal ser repetitivo (periódico e com amplitude quase constante) -> só é necessário codificar um período compressão ainda maior (só esta parte 40%)

Outros: Joint Stereo não aplicado por ser monoaural (5%)

21. (10%) Em que consiste o vácuo semântico e qual o seu impacto (semantic gap) em sistemas de Multimedia Information Retrieval? De que maneira se pode estreitar esse vácuo?

(50%) Vácuo semântico: vazio de significado entre as features de baixo-nível (tipicamente utilizadas) e a cognição/compreensão humana (25%)

Só low-level -Z não se conseguem captar conteúdos de alto-nível (12.5%)

Pesquisas com base em critérios de alto-nível podem, no entanto, ocasionar dificuldades, uma vez que:

-É difícil criar features computacionais que representem com precisão de conceitos de alto-nível (ex emoção) (12.5%)

(50%) Estreitar:

-Avançar a investigação e desenvolvimento de features de alto-nível mais rprecisas, e.g, através de técnicas de machine learning

-Aplicar técnicas de fusão de informação (e.g., áudio e letra, áudio e imagem, etc.) -> a combinação de várias fontes é mais informativa

-...