

## Lista 1

1. Dê a sua própria definição de multimídia para a área computacional, com base no que foi discutido em aula.

É a combinação, controlada por computador, de pelo menos um tipo de mídia estática (texto, fotografia, gráfico), com pelo menos um tipo de mídia dinâmica (vídeo, áudio, animação)

2. Você classificaria um dispositivo tal como o iPod Classic como um sistema multimídia? Se sim, liste os diferentes tipos de mídia que ele suporta. Para cada mídia listada, classifique-a de acordo com a finalidade (percepção, representação, apresentação, armazenamento e transmissão) e dependência do tempo (discreta ou contínua).

O iPod Classic é um sistema multimedia pois apresenta 3 medias

- 1) Imagem: discreta e armazenamento
- 2) Audio: contínuo e de representação
- 3) Video: contínuo e apresentação.

3. Aponte qual a motivação em comprimir dados digitais.

Apesar do constante aumento da capacidade de armazenamento de dados digitais, a demanda por uma maior capacidade sempre cresce mais rápido, logo a compressão de dados provavelmente continuará muito importante.

4. É aconselhável realizar compressão lossy em textos? Por quê?

Não, pois com a compressão lossy algumas informações serão perdidas no processo de descompressão, o que não é aceitável para textos

5. Seja um determinado tipo de informação que é usualmente codificado por longas strings que alternam zeros e uns, como, por exemplo, 010110101010010101010101110101. É possível adaptar o algoritmo de codificação por carreira (run-length) para obter boas taxas de compressão para esse tipo de informação? Explique e mostre como seria o resultado da compressão para a string dada como exemplo. (Dica: pense em substrings de tamanho maior que 1)

Sim, basta tomar por exemplo 01=a e 10=b c=11 d=00

Então a string comprimida fica assim:

a,2 b,3 a,6 c,1 b,2 1,1

6. Dê exemplo de uma situação em que a codificação por diferenças pode ocasionar perdas.

Perdas ocorrem quando a quantidade de bits utilizados para armazenar as diferenças entre duas amplitudes sucessivas não é grande o suficiente para representar o valor correto

7. O que significa dizer que a árvore de Huffman tem a propriedade do prefixo? Qual é a vantagem dessa propriedade?

A propriedade do prefixo indica que nenhum código pode ser prefixo de outro, a vantagem é que ela impede que um símbolo seja decodificado erroneamente, posto que uma parte do seu código

representaria um outro símbolo.

8. Sabe-se que uma das vantagens da codificação aritmética é que ela sempre atinge o valor da entropia. Cite uma de suas principais desvantagens, intrinsicamente relacionada com o hardware em que ela é executada.

O problema com a codificação aritmética é que neste método, o número de dígitos no código cresce linearmente de acordo com o tamanho da string. Logo, o número máximo de caracteres em uma string é determinado pela precisão de ponto flutuante na máquina destino.

Solução: Strings grandes podem ser quebradas em duas ou mais substrings.

9. Decodifique a string representada pelo código 0,44220. Dado o alfabeto formado por A, C, O e S cujas probabilidades são:

$$P(A) = 0,4$$

$$P(C) = 0,3$$

$$P(O) = 0,1$$

$$P(S) = 0,1$$

$$P(.) = 0,1$$

10. Explique como uma palavra é decodificada utilizando o algoritmo LZW.

Este algoritmo é baseado em um dicionário, primeiramente contendo apenas os caracteres ASCII, então as palavras são acodificadas usando-se os índices dos caracteres que a compõe e, então a palavra é armazenada no dicionário, seguindo o esse processo até que todas as palavras sejam codificadas. Para decodificar um palavra basta procurar seu código no dicionário e recupera-la.

## Lista 2

1. A digitalização de áudio se baseia em transformar um sinal de áudio analógico, com valores expressos em Volts, para um sinal de áudio digital, com valores expressos em bits. Isto é necessário para que o computador possa operar sobre o sinal de áudio. A digitalização pode ser dividida em duas etapas: a amostragem e a quantização. Explique o que ocorre com o sinal de áudio em cada um delas.

Amostragem: são realizadas leituras periódicas e instantâneas da amplitude do sinal em espaços de tempo uniformes

Quantização: converte os valores analógicos amostrados em valores digitais

2. Para que fosse possível fazer o processo inverso da digitalização, isto é, através do sinal digital reconstruir o analógico, seriam necessárias amostras infinitas do sinal de modo que não houvesse distorções que causassem perda da qualidade. Felizmente, há um teorema que dita qual a taxa de amostragem mínima para que um sinal não sofra distorções ao ser reconstruído.

a. Qual é este teorema e qual a taxa de amostragem mínima a ser adotada?

Teorema de Nyquist: Para obter uma representação precisa de um sinal analógico, sua amplitude deve ser amostrada a uma taxa mínima igual ou superior ao dobro da componente de mais alta frequência presente no sinal

b. Qual o fenômeno que ocorre caso não seja obedecido o teorema? Explique.

Aliasing, o sinal não é representado corretamente

3. Diferente do que acontece na amostragem, a quantização do sinal de áudio não possui nenhum teorema.

a. Neste caso, qual a abordagem a ser seguida para obter um bom número de bits a ser usado por amostra?

Existe um trade-off entre o numero de amostras e o numero de níveis, deve-se utilizar o bom senso, levando-se o objetivo em consideração.

b. Explique o que acontece quando se utilizam poucos bits por amostra.

Ocorrem distorções no sinal reconstruído

4. Podemos considerar o PCM (Pulse Code Modulation) como um padrão para codificação de áudio. Uma versão melhorada desse padrão, o ADPCM (Adaptive Differential Pulse Code Modulation), possibilita a codificação do áudio com boa compressão utilizando-se a codificação por diferenças modificada para fazer lookahead. Explique como essa codificação é feita e o porquê da utilização de um lookahead.

Amostras adjacentes de áudio são parecidas. ADPCM faz previsão da amostra seguinte e codifica apenas a diferença. O lookahead é utilizado para adaptar a escala da diferenças de acordo com o tamanho da mudança.

5. O MP3 (MPEG-1 Audio Layer 3) é um dos padrões mais conhecido e utilizado para codificação de áudio. A boa taxa de compressão atingida envolve a mistura de diferentes técnicas de compressão lossy e lossless, bem como explora características do modelo psico-acústico humano, sendo elas: a sensibilidade do ouvido, o mascaramento de frequências e o mascaramento temporal.

a. Explique cada uma das características do modelo psico-acústico humano.

Sensibilidade da audição

Mascaramento de frequência (uma frequência de som pode diminuir a capacidade do ouvido humano de perceber frequências posteriores simultaneamente).

Mascaramento temporal (após ouvir um som alto, demorará alguns instantes até que o ouvido possa perceber um som mais baixo)

b. Como o algoritmo do MP3 explora cada uma dessas características para obter a compressão do áudio?

Sensibilidade do ouvido: elimina as frequências que o ouvido humano não consegue captar;

Mascaramento Temporal: após codificar um som intenso, não codifica sons menos intensos que seriam codificados posteriormente; Mascaramento de Frequência: após codificar determinada frequência não codifica sons no intervalo de frequências posteriores.

### Lista 3

1. Cones e bastonetes são fotorreceptores importantes para o sistema visual humano. Cada um está presente em diferentes regiões e é sensível a características diferentes. Por nós possuímos quantidades bem diferentes de cada um, somos mais sensíveis a uma característica do que a outra, o que é explorado na compressão de vídeos e imagens. Defina a função dos cones e dos bastonetes e a qual característica o olho humano é mais sensível.

Cones: fotorreceptores para cor.

Bastonetes: fotorreceptores para intensidade luminosa

O olho é mais sensível a intensidade luminosa

2. O CCD é o dispositivo responsável pela captura de imagens.

a. Explique seu funcionamento.

É composto por uma malha de material fotossensível. Cada célula da malha corresponde a um ponto da imagem. A intensidade da luz incidente em cada célula é convertida em sinal elétrico.

b. No que difere a captura pelo CCD de imagens em tons de cinza em relação a imagens coloridas?

Utilizam um CCD para cada primária (RGB). Um filtro separa a luz incidente direcionando as componentes para o CCD correto.

3. Quanto tempo seria necessário para transmitir uma imagem truecolor sem compressão com resolução 1024x768 em uma rede a 1.5 Mbps? (Dica: pesquisa quantos bits por canal existem para uma imagem truecolor)

Ao realizar os cálculos do exercício 3, você consegue perceber a real necessidade de compressão em imagens (e qualquer outro tipo de mídia digital). Felizmente, imagens possuem muitos dados redundantes que podem ser explorados visando compressão. Os próximos exercícios são relacionados às técnicas que permitem explorar os três tipos de redundância encontrados em imagens: redundância estatística, redundância espacial e redundância psicovisual.

São necessários 12.6 segundos para transmitir a imagem.

4. Explique o que é a redundância estatística e quais técnicas já estudadas anteriormente podem ser utilizadas para explorá-la?

Redundância Estatística é a redundância que se consegue quando se comprime uma fonte de dados usando códigos menores para os valores mais frequentes, ou seja, usar menos bits para representar as cores mais frequentes na imagem, pois os valores dos pixels em uma imagem tem função de densidade de probabilidade não uniforme.

Pode-se utilizar as codificações estatísticas vistas antes, como Huffman.

5. Qual a diferença entre redundância estatística e redundância espacial? Dê uma técnica sem perdas e uma técnica com perdas que pode ser utilizada para remoção da redundância espacial.

Redundância Espacial é o nome dado a correlação entre pixels vizinhos, ou seja, o valor de um vizinho pode ser estimado pelos valores de seus vizinhos. Para remover-se a redundância espacial

pode-se utilizar codificação por diferença (lossless) ou codificação por transformada (lossy).

6. A utilização de transformadas DCT é um passo importante na remoção de redundâncias psicovisuais. Apenas sua utilização, porém, não comprime os dados. Explique então o porquê de ela ser importante.

Porque ela transforma a matriz (imagem) em matriz de frequências espaciais. É a partir das frequências espaciais que o processo de compressão é feito

7. Explique o processo de compressão JPEG, destacando em que momentos ocorrem as remoções de redundâncias espacial, estatística e psicovisual.

O processo de compressão JPEG possui 5 etapas principais:

- 1) A imagem é dividida em blocos (ou matrizes) de 8x8
- 2) Os blocos passam pela Transformada Discreta dos Cossenos (DCT)
- 3) Os blocos passam pelo processo de Quantização, no qual a matriz é “dividida” por uma matriz default, nessa etapa removem-se as redundâncias psicovisuais
- 4) A matriz é vetorizada e passa por codificação de diferenças. Depois passa por codificação de carreira (redundância espacial) e finalmente por codificação estatística (redundância estatística)
- 5) Construção do frame.

#### Lista 4

1. Um sinal de vídeo nada mais é do que uma sequência de imagens (ou quadros) que, quando amostrada e reproduzida a certa frequência, provoca a sensação de movimento.

a. Qual a frequência (Hz ou fps – frames per second) utilizada pelo padrão de TV?

25 fps.

b. Qual a vantagem e desvantagem da amostragem progressiva e da entrelaçada?

A entrelaçada permite aumentar a taxa de refresh sem aumentar a quantidade de amostras

2. Em relação à compressão, qual a vantagem de se transformar o espaço de cores de um quadro de RGB para YCbCr?

O formato YCbCr consegue explorar o fato de que o olho humano é mais sensível à luminância do que da crominância, utilizando mais bits para transmitir a luminância do que as crominâncias assim consegue-se valores melhores do que usando o formato RGB.

3. Explique as diferenças e consequências de utilização dos formatos de amostragem YCbCr 4:4:4, 4:2:2 e 4:2:0.

O formato 4:4:4 apresenta componentes iguais para Y, Cb e Cr

O formato 4:2:2 apresenta a cada 4Y na horizontal 2 Cb e 2 Cr

O formato 4:2:0 para cada 4Y 1 Cb e 1 Cr;

4. Pelo fato de o vídeo se tratar de uma sequência de quadros, podemos aplicar a cada um as técnicas de compressão relacionadas a imagens, para remoção de redundâncias estatísticas, espaciais e psicovisuais. Além disso, o vídeo traz um novo tipo de redundância que pode ser explorado. Explique qual é este tipo e como ele pode ser explorado para aumentar a compressão.

O vídeo introduz o conceito de Redundância Temporal, ou seja, blocos de dados aparecem replicados ou muito parecidos, o que permite que sejam codificados menos vezes.

5. Explique o porquê de os métodos de previsão de movimento serem classificados como não perfeitos. O que se pode fazer para melhorar a estimativa?

6. Durante a codificação, cada quadro recebe um nome especial, dependendo de como será codificado.

a. Defina o que são quadros I, P e B.

Quadros I: são codificados sem nenhuma referência a outros quadros

Quadros P: são codificados em relação ao conteúdo de um quadro I ou de um quadro P anterior

Quadros B: são codificados em relação ao conteúdo de um quadro I ou de um quadro P anterior e de um posterior.

b. Defina GOP span e Prediction span.

GOP span: número de quadros (3 a 12) entre dois quadros I sucessivos

Prediction span: número de quadros entre um quadro P e o quadro I ou P imediatamente anterior

c. Qual a limitação encontrada nos quadros P?

Os quadros P propagam erros, por isso a quantidade de quadros P entre quadros I é limitada.

d. Qual a vantagem e desvantagem da utilização de quadros B?

A vantagem é que não propaga erros e a desvantagem é o alto custo computacional de codificação e decodificação.

7. A evolução das técnicas de codificação, bem como das tecnologias de transmissão e reprodução multimídia, resultou na criação dos padrões MPEG-4 e H.264.

a. Cite duas características do padrão MPEG-4 que o diferenciam de seus antecessores.

O conceito de objeto e alto nível de interação com o conteúdo

b. Cite vantagens da utilização do padrão H.264.

O padrão H.264 apresenta as maiores taxas de compressão e alta escalabilidade do codificador.

8. Muitas pessoas erroneamente entendem que a extensão de arquivo de vídeo (.AVI, por exemplo) representa a compressão que está sendo utilizada, sendo que na verdade, ela é somente um formato de representação. Explique a diferença entre formatos de representação e formatos de codificação.

Formatos de representação são como pacotes de dados e podem utilizar diversos formatos de codificação.