

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS MATEMÁTICAS E DE
COMPUTAÇÃO - SÃO CARLOS

SCC0661 - Multimídia e Hiperfídia

Prof. Marcelo Manzato

Listas de exercício - 1 a 4

Aluno:
Jonas DOURADO
6445442

São Carlos,
27 de setembro de 2011

Lista 1

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

Multimídia é a utilização de mais de um tipo de mídia com o intuito de transmitir uma mensagem.

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). Sim, o iPod suporta vídeos, fotos, músicas, texto. As tabelas seguintes classificam a mídia de acordo com a finalidade

Percepção	Mídia
Visual	vídeo, foto, texto
Auditiva	vídeo, música

Representação	Mídia
ASCII	texto
MPEG	vídeo
MP3	música
JPEG	foto

Apresentação	Mídia
LCD	vídeo, foto, texto
Headphone	vídeo, música

Armazenamento	Mídia
Memória flash interna	vídeo, música, foto, texto

Transmissão	Mídia
USB	vídeo, música, foto, texto

Granularidade	Mídia
Contínua	vídeo, música
Discreta	foto, texto

3. Aponte qual a motivação em comprimir dados digitais.

O principal motivo de se comprimir dados é contornar a capacidade de armazenamento e transmissão, ou seja, trabalhar com um conteúdo de maior qualidade utilizando a mesma capacidade de armazenamento e transmissão (Não esquecer que há necessidade de maior processamento).

4. É aconselhável realizar compressão lossy em textos? Por quê?
- Não, pois texto é uma mídia onde há a necessidade de ser apresentada em sua forma original, caso contrário, pode não ser entendida.
5. Seja um determinado tipo de informação que é usualmente codificado por longas strings que alternam zeros e uns, como, por exemplo, 0101101010100101010101110101. 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)
- Uma boa forma de se realizar a compressão é ao ter dois pares de números binários iguais seguidos, o próximo par identificaria o número de cópias seguidas. Utilizando essa codificação ficaria 0101001010100101110111010100.
6. Dê exemplo de uma situação em que a codificação por diferenças pode ocasionar perdas.
- A perda pode ocorrer quando a variação entre amplitudes sucessivas é maior do que a maior delta possível de acordo com o número de bits escolhidos para o delta. Um bom exemplo é quando uma som gravado em WAV com muitos instrumentos de frequências e alturas bem distintas, ao ser codificado por diferenças, ocorre distorções.
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 na árvore de Huffman significa que nenhum código é prefixo de outro. A vantagem é que ao decodificar, como nenhum código é prefixo de outro código, o código que se encontra no início do arquivo comprimido não apresenta ambiguidade. Pode-se simplesmente identificar este código inicial, traduzi-lo de volta ao caracter original e repetir o processo no restante do arquivo comprimido.
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, intrinsecamente relacionada com o hardware em que ela é executada.
- A desvantagem intrinsecamente relacionada ao hardware é que o tamanho do código é determinado pela precisão de ponto flutuante da máquina.
9. Decodifique a string representada pelo código 0,44220. Dado o alfabeto formado por A, C, O e S cujas probabilidades são: $P(A) = 0,4$ $P(C) = 0,3$ $P(O) = 0,1$ $P(S) = 0,1$ $P(.) = 0,1$
- 0,44220 está entre 0,4 e 0,7 C 0,44220 está entre 0,4 e 0,52 A 0,44220 está entre 0,4 e 0,448 A 0,44220 está entre 0,4384 e 0,4432 S 0,44220 está entre 0,44176 e 0,44224 O 0,44220 está entre 0,442192 e 0,44224 .
10. Explique como uma palavra é decodificada utilizando o algoritmo LZW.
- A palavra é decodificada utilizando o dicionário criado dinamicamente.

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.

A amostragem realiza aferições em tempo predeterminado, convertendo a onda mecânica do som em sinal elétrico. A quantização atribui um nível (dentro os pré determinados) equivalente a tensão do sinal elétrico.

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?

É o teorema de Nyquist que diz: “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”. Ou seja, a taxa mínima é pelo menos o dobro da maior frequência amostrada.

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

Ocorre o aliasing, que é quando um sinal de maior frequência é interpretado como um sinal de menor frequência devido a baixa amostragem do sinal.

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?

Tem que ser de tamanho suficiente para que seja semelhante com o sinal amostrado e pequeno o suficiente de forma que não fique muito grande.

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

Acontece distorções, pois sinais de níveis visivelmente diferentes acabam sendo quantizados na mesma faixa de valor.

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.

O algoritmo faz lookahead durante a compressão para adaptar a escala de diferenças de acordo com o tamanho da mudança. O lookahead é necessário para verificar se há a necessidade de mudar a escala de diferenças.

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 de audição → A resposta psicológica a cada frequência de som.

Mascaramento de frequência → Uma frequência em conjunto com outra próxima pode não ser percebida

Mascaramento temporal → Após ouvir um som alto, por um pequeno instante, o ouvido humano não consegue perceber sons mais baixos.

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

O algoritmo do MP3 explora sensibilidade de audição, mascaramento de frequência e mascaramento temporal obter uma boa compressão de áudio.

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.

Os cones são responsáveis pela captação de cores e os bastonetes são responsáveis pela captação de intensidade luminosa. O olho humano é mais sensível a intensidade luminosa.

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

- a. Explique seu funcionamento.

O sensor CCD é um dispositivo que tem uma malha de células e cada célula captura a intensidade da luz.

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

O sensor é o mesmo, a diferença é que ao capturar cor, um filtro separa cada cor(RGB) e direciona para um CCD diferente(há a necessidade de ter 3 sensores CCD).

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)

$X \times Y \times \text{Bits/pixel}$

$1024 \times 768 \times 24 = 18.874.368 \text{ bits} = 18 \text{ Mbps}$

$18 / 1.5 = 12 \text{ segundos}$

São necessários 12 segundos para transmitir essa imagem sem compressão.

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.

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

Redundância estatística é a transmissão da mesma informação utilizando símbolos mais frequentes, representações menores. Podemos aplicar codificação Huffman e aritmética.

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.

A redundância espacial leva em conta a interdependência entre pixels vizinhos, enquanto a estatística, leva em conta a probabilidade de cada símbolo. Um técnica lossless é a codificação delta(por diferenças) e uma com perdas é a codificação por transformadas.

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. A transformada DCT é importante pois ela diminui a redundância psicovisual dos símbolos.
7. Explique o processo de compressão JPEG, destacando em que momentos ocorrem as remoções de redundâncias espacial, estatística e psicovisual.

A compressão JPEG começa dividindo a imagem em blocos de 8x8 pixels. Em seguida ocorre a transformada DCT, que auxilia no processo de remoção de redundância psicovisual. Após a transformada DCT, ocorre a quantização, que de fato remove a redundância psicovisual. Logo depois, ocorre a codificação da imagem, que realiza uma codificação por diferença e uma codificação runlength que eliminam a redundância espacial e então realiza uma codificação por Huffman que elimina a redundância estatística. Por fim, ocorre a construção do quadro, onde grava os headers e os blocos.

(Exercício extra - opcional)
8. Faça uma pesquisa sobre o funcionamento do padrão GIF. Após entender como ele funciona, responda: Para a compressão de uma foto de paisagem qual método irá obter a melhor taxa de compressão, JPEG ou GIF? Por quê?

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?

As frequências dos sinais de cores mais difundidos são aproximadamente 30Hz para o padrão NTSC e 25Hz para o padrão PAL.

- b. Qual a vantagem e desvantagem da amostragem progressiva e da entrelaçada?] A vantagem da amostra progressiva é a menor largura de banda necessária, e a desvantagem é a

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

A vantagem é que o espaço de cores YCbCr separa a luminância e crominância, o que facilita a manipulações de vídeos (ex. redução da redundância psicovisual), já que o olho humano tem maior sensibilidade a luminância

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

4:4:4 - Um componente Y, Cb e Cr para cada pixels - Mantém a qualidade de crominância da imagem, entretanto ocupa espaço maior. 4:2:2 - A cada 4 Y na horizontal, existem 2 Cb e Cr - Balanço entre qualidade de crominância com economia de espaço. 4:2:0 - A cada 4 Y, existem um Cb e um Cr - Economiza espaço, entretanto diminui a qualidade de crominância da imagem.

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 além das compressões aplicáveis a imagens podem também ser comprimido reduzindo-se a redundância temporal. Essa redundância temporal pode ser explorada usando predição entre as cenas.

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?

O vídeo pode apresentar movimentos que são maiores que a predição pode prever, para melhorar é necessário incluir mecanismos de compensação de movimento.

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 quadros codificados independentemente, usando JPEG.

Quadros P - São quadros que armazenam estimativa e compensação de movimentos a partir do quadro anterior. Propagam erros.

Quadros BI - São quadros que armazenam estimativa e compensação de movimentos a partir do quadro anterior e posterior. Não propagam erros.

b. Defina GOP span e Prediction span.

GOP span: É o número de quadros entre dois quadros I.

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

c. Qual a limitação encontrada nos quadros P?

A limitação dos quadros P é que eles propagam erros.

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

A vantagem de utilização de quadros B é o aumento da compressão, e a desvantagem é que aumenta o tempo de codificação e decodificação, pois há a necessidade de esperar por um próximo quadro I ou P.

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.

As duas características que diferencia o padrão MPEG-4 dos seus antecessores é a inclusão do conceito de objetos e a inclusão da possibilidade de interagir com a cena.

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

As vantagens são observadas no armazenamento e na transmissão dos dados, pois o H.264 permite uma maior compressão de vídeo.

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.

Um formato de representação pode conter diversos formatos de codificação desde que suportados pelo formato de representação.

Referência

1. <http://www.ic.unicamp.br/~rezende/ensino/mo417/2010s2/Slides/Aula17.pdf>