

Tópicos

- √ Introdução
- √ Representação de Informação Multimídia
- √ Digitalização
- Mídias Texto, Imagem, Áudio e Vídeo
- • Compressão, Padrões de Compressão (Texto, Imagem, Áudio e Vídeo)
- Comunicação Multimídia
- Protocolos de Rede, Redes Locais
- Redes Sem Fio, Bluetooth
- Sincronismo de Mídias
- Qualidade de Serviço
- Tópicos Avançados: Criptografia, Watermarking;
- Realidade Virtual
- Serviços Multimídia: Vídeo Sob Demanda, Videoconferência
- Sistemas Multimídia Avançados: Ambientes Virtuais Colaborativos

Compressão de Dados

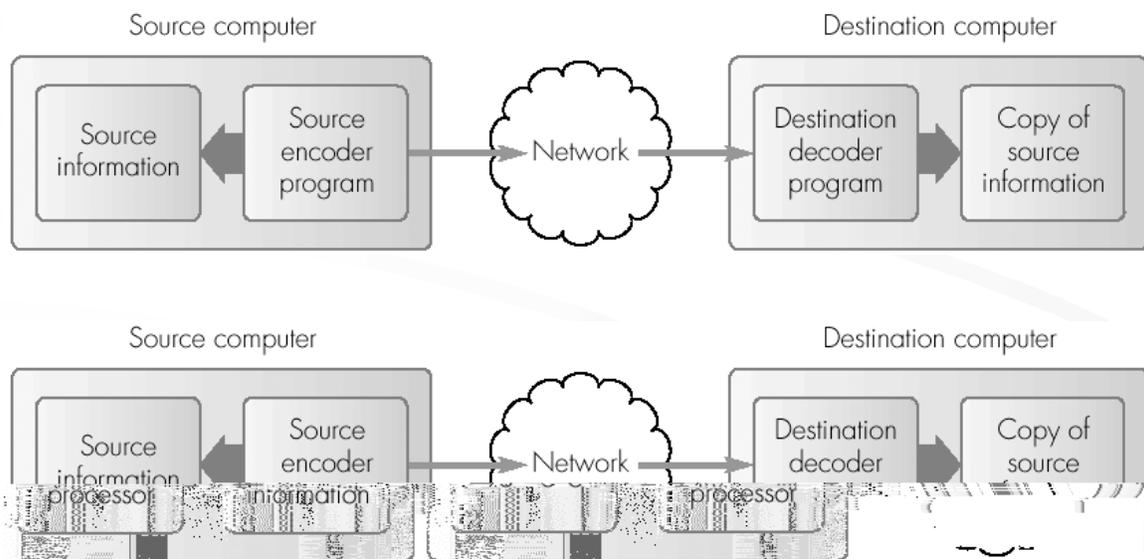
Motivação

- Como visto anteriormente, as diversas mídias podem requerer uma banda passante de rede não disponíveis.
 - Imagem: 18,9Mbits para imagem 1024×768 a 24bpp.
≈2,5Mbits para imagem VGA a 8bpp.
 - Áudio: 64 kbps (voz, mono, qualidade de telefone).
1,411 Mbps (música, stereo, qualidade de CD).
 - Vídeo: ≈ 332 Mbps for $YC_R C_B$ 4:2:2 video.
- Similarmente, usuários de redes com cobrança por quantidade de dados transmitidos (ou quantidade de tempo de conexão) se beneficiariam da diminuição de quantidade de dados enviados na rede.

Compressão de Dados

- Função realizada sobre dados antes da transmissão.
 - Codificador da Origem (*Source Coder*) ↘
 - Decodificador do Destino (*Destination Decoder*)
- Usada para reduzir o volume de informação a ser transmitida ou reduzir a banda passante necessária para transmissão dos dados.
- Compressão com perdas e sem perdas.

Compressão de Dados *Source Coder / Destination Decoder*



Compressão de Dados

Compressão Com Perdas e Sem Perdas

- Compressão sem perdas busca reduzir a quantidade de informação, mas no destino uma cópia exata dos dados originais são recuperados, a compressão é reversível.
 - Transferência de texto, arquivos binários, etc.
- Compressão com perdas busca permitir a recuperação de uma versão dos dados originais que são percebidos pelo usuário como sendo parecidos o suficiente com o original.
 - Transferência de imagens digitais, áudio, vídeo, pois o olho e ouvido humanos não são capazes de perceber pequena perda de qualidade no sinal.

Compressão de Dados

Codificação por Entropia

- É um modo de codificação sem perdas.
- É independente do tipo de informação sendo comprimido.
- Codificação em Carreira (*Run-Length*)
 - Usada quando o sinal a ser codificado contém uma longa sequência de bits repetidos.
 - Sequência de bits substituída pelo bit e indicação de quantidade do mesmo. 00000001111111110000011 ... → 0,7,1,10,0,5,1,2 ...
Assumindo que a sequência inicia com zeros → 7,10,5,2 ...

Compressão de Dados

Codificação Estatística

- Modelos de codificação utilizam o mesmo número de bits por valor (exemplo: ASCII). Alguns símbolos aparecem com maior frequência que outros. Símbolos que aparecem com maior frequência podem usar menos bits que aqueles que aparecem com menor frequência.
- Num texto a letra A aparece com maior frequência (probabilidade) que a consoante 'P', que aparece com maior frequência que 'Z'... Utiliza-se uma codificação com número de bits variável, de modo que na média se necessita menos bits para codificar o mesmo conteúdo.
- Propriedade do Prefixo.
 - Um símbolo não pode ser prefixo de um outro símbolo mais longo.
 - Codificação de Huffman
- A média mínima de bits necessários para transmitir uma determinada sequência de dados é chamada *Entropia*.

Compressão de Dados

Codificação Estatística

- A Entropia pode ser calculada pela fórmula de Shannon abaixo:

$$\text{Entropia}, H = -\sum_{i=1}^n P_i \log_2 P_i$$

onde n é o número de símbolos existentes, P_i é a probabilidade de um símbolo i aparecer.

- A eficiência de um codificador é dado pela razão da entropia do sinal e do número médio de bits por símbolo.
- O número médio de bits por símbolo é dado por:

$$\text{NúmeroMédioDeBits} = \sum_{i=1}^n N_i P_i$$

Compressão de Dados

Codificação Estatística

A statistical encoding algorithm is being considered for the transmission of a large number of long text files over a public network. Analysis of the file contents has shown that each file comprises only the six different characters M, F, Y, N, 0, and 1 each of which occurs with a relative frequency of occurrence of 0.25, 0.25, 0.125, 0.125, 0.125, and 0.125 respectively. If the encoding algorithm under consideration uses the following set of codewords:

M = 10, F = 11, Y = 010, N = 011, 0 = 000, 1 = 001

compute:

- (i) the average number of bits per codeword with the algorithm,
- (ii) the entropy of the source,
- (iii) the minimum number of bits required assuming fixed-length codewords.

Compressão de Dados

Codificação Estatística

- (i) Average number of bits per codeword

$$\begin{aligned} &= \sum_{i=1}^6 N_i P_i = (2(2 \times 0.25) + 4(3 \times 0.125)) \\ &= 2 \times 0.5 + 4 \times 0.375 = 2.5 \end{aligned}$$

- (ii) Entropy of source

$$\begin{aligned} &= \sum_{i=1}^6 P_i \log_2 P_i = - (2(0.25 \log_2 0.25) + 4(0.125 \log_2 0.125)) \\ &= 1 + 1.5 = 2.5 \end{aligned}$$

- (iii) Since there are 6 different characters, using fixed-length codewords would require a minimum of 3 bits (8 combinations).

Compressão de Dados

Codificação por Diferença

- Usado quando a amplitude de um sinal pode assumir valores numa faixa larga, porém a diferença de amplitudes de valores sucessivos é relativamente pequena.
- Ao invés de usar um grande número de bits para representar a amplitude do sinal, apenas a diferença para o valor anterior é armazenado, com um número menor de bits.
- Pode acarretar compressão com perdas ou sem perdas, dependendo do número de bits escolhidos serem suficientes para armazenar a maior das diferenças entre amplitudes sucessivas.
- Um sinal que requer 12 bits para ser codificado pode ter diferenças com não mais de 3 bits, o que acarreta uma economia de 75% na quantidade de bits resultante da codificação das amostras.

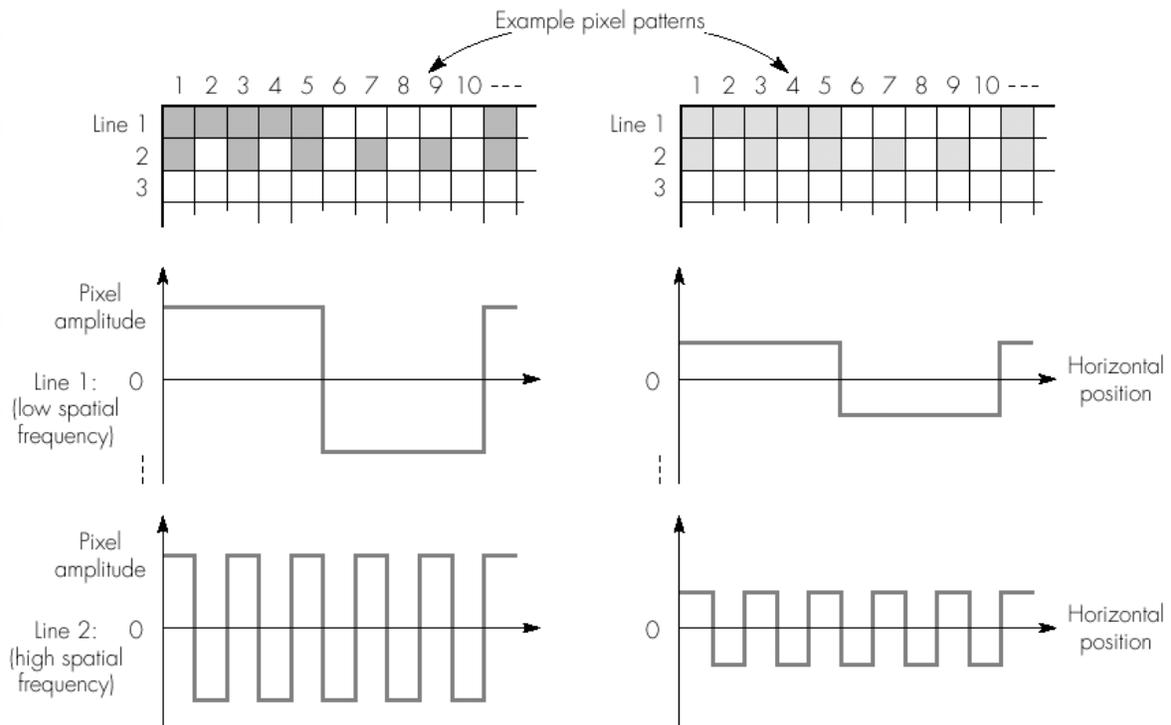
Compressão de Dados

Codificação por Transformadas

- Envolve a transformação da informação original de um formato para outro mais apropriado para compressão.
- Não existe perda pela transformação em si.
- Muito usado para compressão de imagens e vídeo.
- A taxa de mudança na magnitude dos valores dos pixels na matriz define o que denominamos *frequência espacial*, que pode ser horizontal e vertical.
- O olho humano é menos sensível às componentes espaciais de mais alta frequência. Adicionalmente, se a amplitude de uma componente de alta frequência estiver abaixo de um determinado valor, tal componente não será detectada pelo olho humano.
- Seria vantajoso converter os valores espaciais originais em valores baseados em componentes de frequência espacial, já que pode-se eliminar mais facilmente as componentes que pouco contribuem para a inteligibilidade da imagem, diminuindo a quantidade de informação a ser processada.

Compressão de Dados

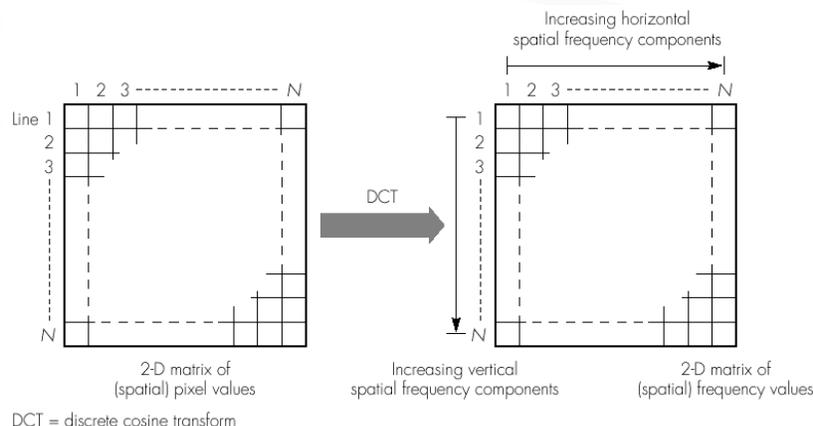
Componentes de Frequência Espacial



Compressão de Dados

Transformada Discreta de Cossenos - DCT

- Técnica matemática que transforma uma matriz de pixels em uma matriz equivalente de componentes de frequências espaciais. Tal matriz é chamada matriz de coeficientes.
- A DCT não introduz perda, mas o descarte de componentes de frequência mais altos introduzem perda irreversível de informação.

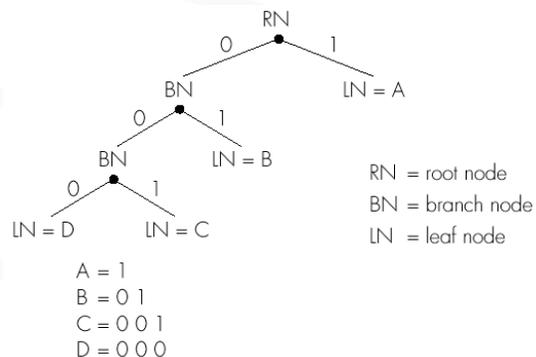


Compressão de Texto

- Necessário usar compressão sem perdas
- Utiliza-se codificação por entropia, mais precisamente codificação estatística. Dois métodos básicos são usados:
 - Utiliza-se caracteres individuais como base para a definição de símbolos para codificação dos dados.
 - Huffman
 - Codificação Aritmética
 - Utiliza-se uma string de caracteres de tamanho variável como base para a definição de símbolos para codificação dos dados.
 - Algoritmo Lempel-Ziv (LZ)
- A codificação pode ainda ser estática ou dinâmica (adaptativa)

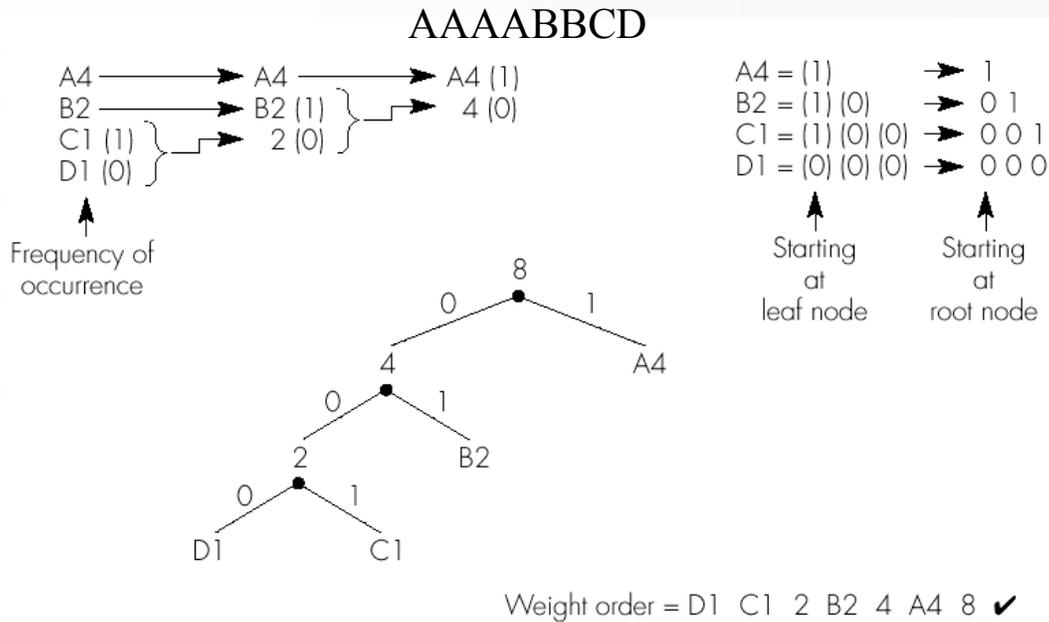
Compressão de Texto Codificação de Huffman (Estática)

- A informação a ser comprimida é analisada e uma árvore desbalanceada é construída baseada na frequência relativa de cada carácter (Árvore do Código de Huffman).
- A Árvore do Código de Huffman é uma árvore binária, com dígitos 0 e 1 atribuídos a cada galho da árvore. Os símbolos a serem codificados ficam nas folhas da árvore e a sequência de bits resultantes dos galhos percorridos da raiz até uma determinada folha indica a sequência de bits que representa o símbolo em questão.



Compressão de Texto

Codificação de Huffman (Estática)



Seriam necessários apenas 14 bits para transmitir a sequência original, que normalmente ocuparia 64 bits (8 bytes numa codificação ASCII)

Exercício

A series of messages is to be transferred between two computers over a PSTN. The messages comprise just the characters A through H. Analysis has shown that the probability (relative frequency of occurrence) of each character is as follows:

A and B = 0.25, C and D = 0.14, E, F, G, and H = 0.055

- Use Shannon's formula to derive the minimum average number of bits per character.
- Use Huffman coding to derive a codeword set and prove this is the minimum set by constructing the corresponding Huffman code tree.
- Derive the average number of bits per character for your codeword set and compare this with:
 - the entropy of the messages (Shannon's value),
 - fixed-length binary codewords,
 - 7-bit ASCII codewords.

(a) Shannon's formula states:

$$\text{Entropy, } H = - \sum_{i=1}^8 P_i \log_2 P_i \text{ bits per codeword}$$

Therefore:

$$H = -(2(0.25 \log_2 0.25) + 2(0.14 \log_2 0.14) + 4(0.055 \log_2 0.055)) \\ = 1 + 0.794 + 0.921 = 2.175 \text{ bits per codeword}$$

(b) The derivation of the codeword set using Huffman coding is shown in Figure 3.4(a). The characters are first listed in weight order and the two characters at the bottom of the list are assigned to the (1) and (0) branches. Note that in this case, however, when the two nodes are combined, the weight of the resulting branch node (0.11) is greater than the weight of the two characters E and F (0.055). Hence the branch node is inserted into the second list higher than both of these. The same procedure then repeats until there are only two entries in the list remaining.

The Huffman code tree corresponding to the derived set of codewords is given in Figure 3.4(b) and, as we can see, this is the optimum tree since all leaf and branch nodes increment in numerical order.

(c) Average number of bits per codeword using Huffman coding is:

$$2(2 \times 0.25) + 2(3 \times 0.14) + 4(4 \times 0.055) = 2.72 \text{ bits per codeword}$$

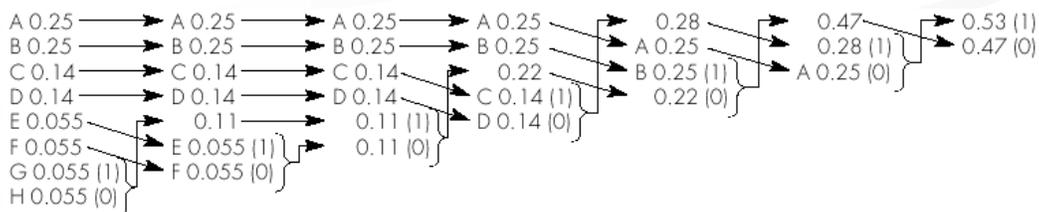
which is 99.8% of the Shannon value.

Using fixed-length binary codewords:

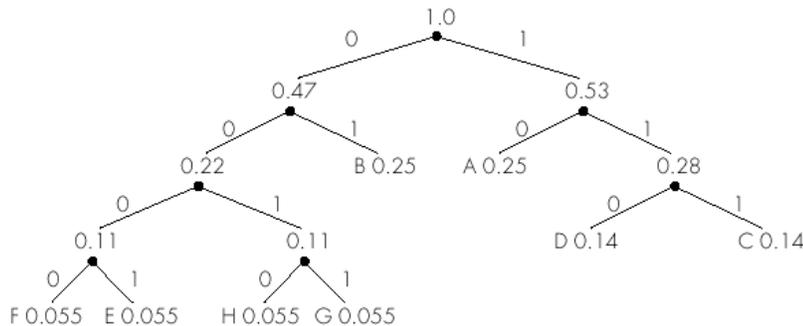
There are 8 characters – A through H – and hence 3 bits per codeword is sufficient which is 90.7% of the Huffman value.

Using 7-bit ASCII codewords:

7 bits per codeword
which is 38.86% of the Huffman value.



A = (0) (1) → 10
 B = (1) (0) → 01
 C = (1) (1) (1) → 111
 D = (0) (1) (1) → 110
 E = (1) (0) (0) (0) → 0001
 F = (0) (0) (0) (0) → 0000
 G = (1) (1) (0) (0) → 0011
 H = (0) (1) (0) (0) → 0010



Weight order = 0.055 0.055 0.055 0.055 0.11 0.11 0.14 0.14 0.22 0.25 0.25 0.28 0.47 0.53 ✓

Compressão de Texto

Codificação de Huffman (Estática)

- Para decodificar uma sequência de bits codificada através do código de Huffman é necessário realizar uma decodificação baseada em bits, ao invés de palavras de 7/8 bits para ASCII, 16 bits para Unicode, etc.
- Por causa da Propriedade do Prefixo, pode-se utilizar uma tabela de valores e algoritmo de busca para procurar sequências de bits válidas. Opcionalmente pode-se utilizar a própria Árvore do Código de Huffman para decodificação da sequência de bits.
- O decodificador precisa conhecer previamente, ou receber, a tabela de códigos válidos.
- Na codificação dinâmica o codificador/decodificador constroem a Árvore de Huffman dinamicamente, conforme os caracteres são transmitidos.

Compressão de Texto

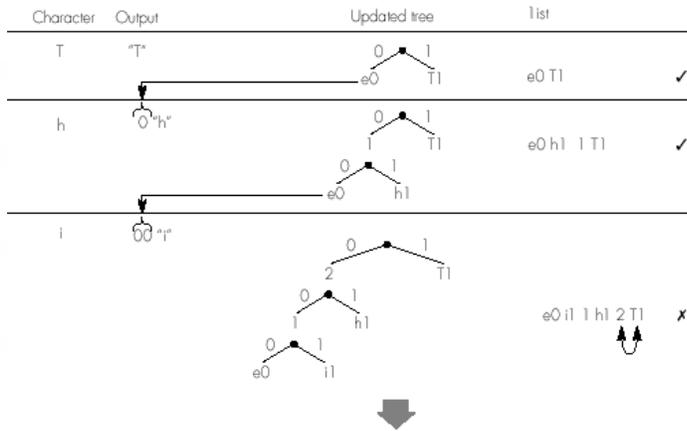
Codificação de Huffman (Dinâmica)

- Quando um caractere a ser codificado já existe na árvore ele é codificado de acordo com a sequência de bits especificada. Quando um caractere é novo, ele é enviado sem compressão para o destino. O codificador então atualiza sua árvore seja incrementando o contador de ocorrências do caractere ou criando nova entrada na árvore.
- O decodificador realiza as mesmas operações, derivando a mesma árvore.
- Utilizado em aplicações que envolvem transmissão de texto.

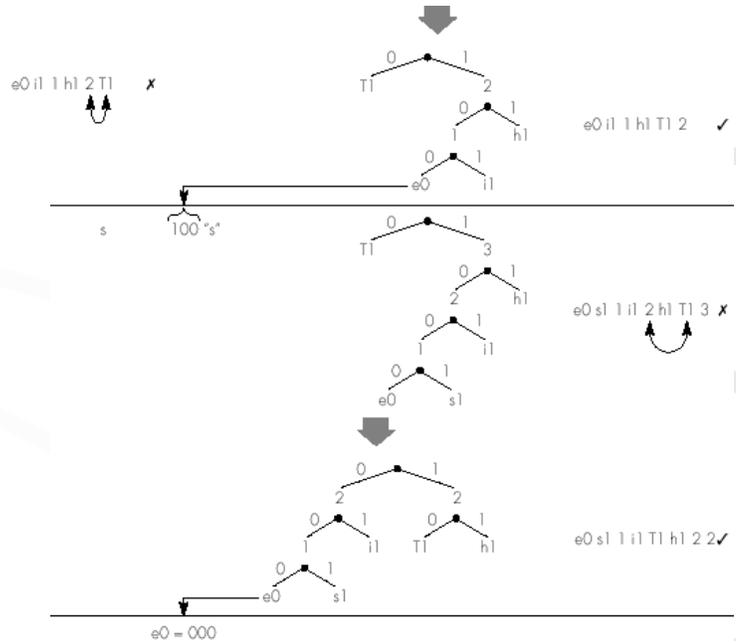
Input string = This is simple
 Initialized tree:

␣ = Space character

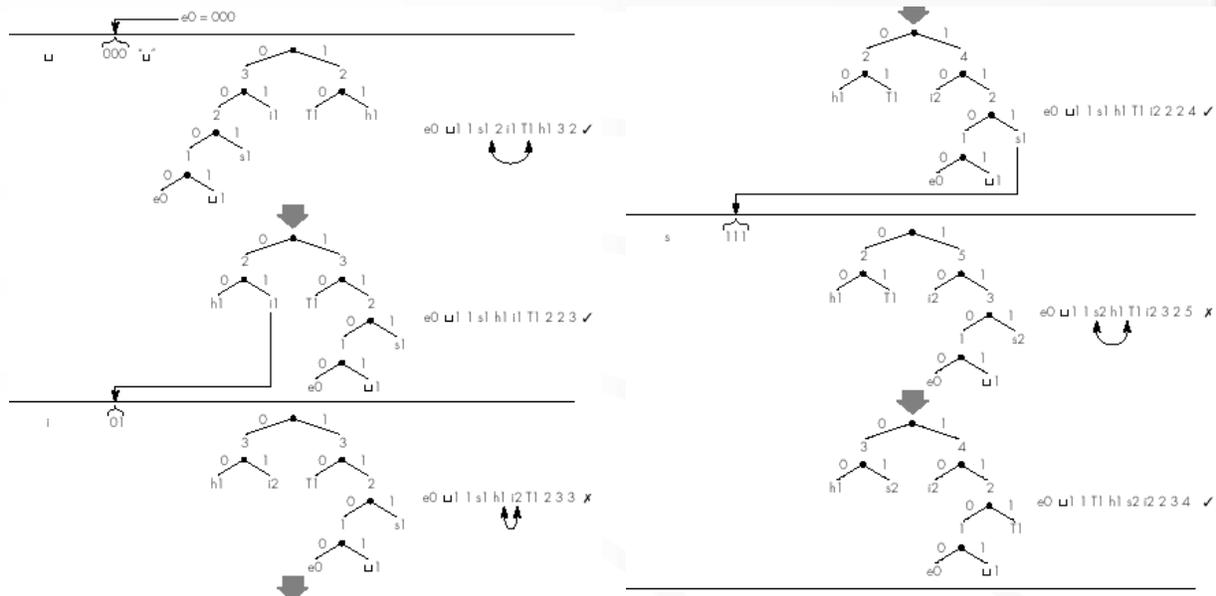
e0 = empty leaf list



Compressão do texto: *This is simple*



Compressão do texto: *This is simple*



Compressão de Texto

Codificação Aritmética

- Codificação de Huffman só alcança o desempenho indicado pela equação de Shannon se as probabilidades de ocorrência de caracteres for dado por números inteiros potências de $1/2$, o que leva a poucos casos onde as sequências de bits são ótimas.
- As sequências de bits produzidos pela Codificação Aritmética, por outro lado, sempre alcançam o desempenho indicado pela equação de Shannon.
- É bem mais complexo que codificação de Huffman.
- Codificação Aritmética cria uma sequência de bits para cada combinação de caracteres, incluindo o caracter terminador.
- Defini-se faixas baseados nas probabilidades de frequência.
- Cada combinação subsequente é definida como um subintervalo do intervalo anterior.

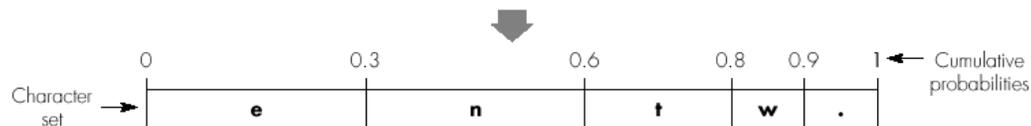
Compressão de Texto

Codificação Aritmética

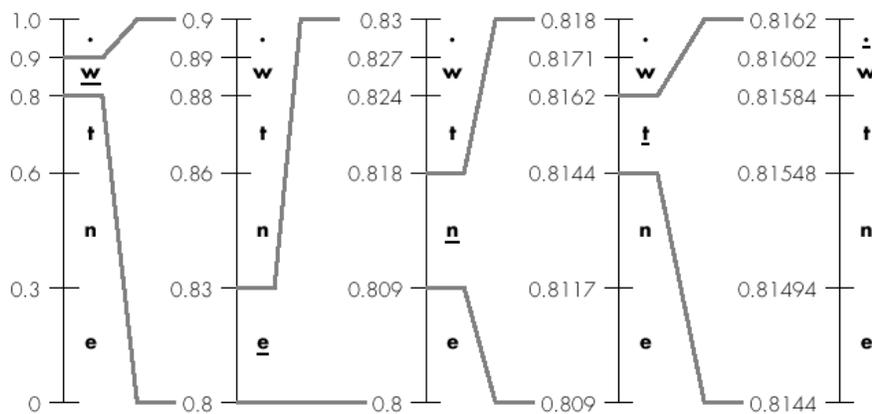
(a)

Example character set and their probabilities:

$$e = 0.3, n = 0.3, t = 0.2, w = 0.1, . = 0.1$$



(b)



Encoded version of the character string **went.** is a single codeword in the range $0.81602 \leq \text{codeword} < 0.8162$

Compressão de Texto

Codificação Aritmética

- Qualquer valor no intervalo final serve para representar a string em questão.
- Cada novo caracter representa mais um dígito no número de ponto flutuante sendo gerado.
- A precisão de números de ponto flutuante limita o número de caracteres que podem ser comprimidos.
- Uma string pode, portanto, ser representada por uma série de números de ponto flutuante.
- Modo de compressão dinâmico é bem mais complexo.
- Modo de compressão binário é bem mais complexo.

Compressão de Texto

Codificação Lempel-Ziv

- Usa sequências de caracteres como base para a codificação.
- Sequências de caracteres são catalogados em uma tabela e apenas a posição de uma determinada sequência é transmitida/armazenada.
- É um modelo de compressão baseado em dicionário.

The LZ algorithm is to be used to compress a text file prior to its transmission. If the average number of characters per word is 6, and the dictionary used contains 4096 words, derive the average compression ratio that is achieved relative to using 7-bit ASCII codewords.

Answer:

In general, a dictionary with an index of n bits can contain up to 2^n entries. Now $4096 = 2^{12}$ and hence an index of 12 bits is required.

Using 7-bit ASCII codewords and an average of 6 characters per word requires 42 bits.

Hence compression ratio = $42/12 = 3.5:1$.

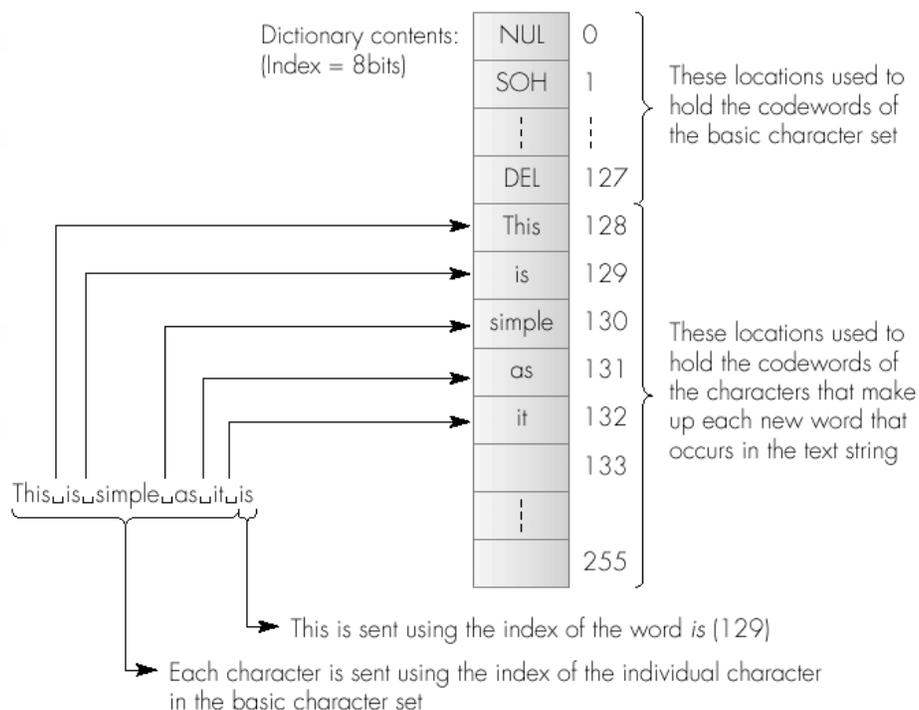
Compressão de Texto

Codificação Lempel-Ziv-Welsh

- O algoritmo LZW é similar ao LZ, mas o dicionário é construído dinamicamente, o que implica que somente as palavras em uso de fato existem no dicionário.
- Inicialmente os dicionários (codificador e decodificador) contém somente os caracteres básicos (ASCII por exemplo). O dicionário é criado de modo que espera-se ter um número máximo de palavras, sendo que as primeiras são os caracteres básicos e as demais palavras adicionadas dinamicamente ao dicionário.
- Somente palavras contendo letras são armazenadas no dicionário. Os demais caracteres são usados como delimitadores.
- Dicionários podem iniciar com um tamanho reduzido e ir aumentando o número de verbetes dinamicamente.

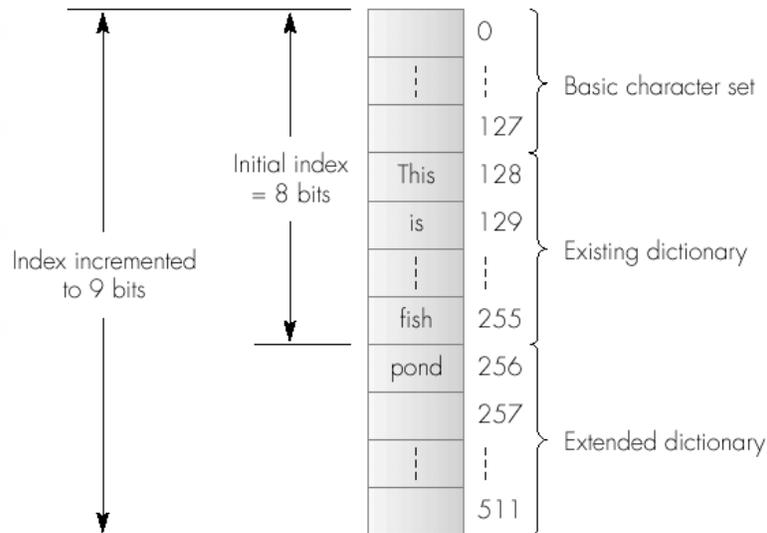
Compressão de Texto

Codificação Lempel-Ziv-Welsh



Compressão de Texto

Codificação Lempel-Ziv-Welsh



Compressão de Áudio

Pulse Code Modulation

- PCM (*Pulse Code Modulation*) é o mecanismo utilizado para digitalização de sebaus de áudio.
 - Consiste da amostragem do sinal à taxa mínima do dobro da componente de frequência máxima do sinal.
 - No caso de transmissão do sinal por um canal de comunicação com limitação de banda, a amostragem pode ser feita baseada na banda limitada do canal.
 - Uma vez quantizados, sinais de voz são usualmente codificados com 12 bits por amostra e sinais de som genérico com 16 bits por amostra.
 - Codificação A-law e μ -law levam o sinal amostrado a ser armazenado em 8 bits por amostra (voz) com qualidade similar à codificação com 12 bits.

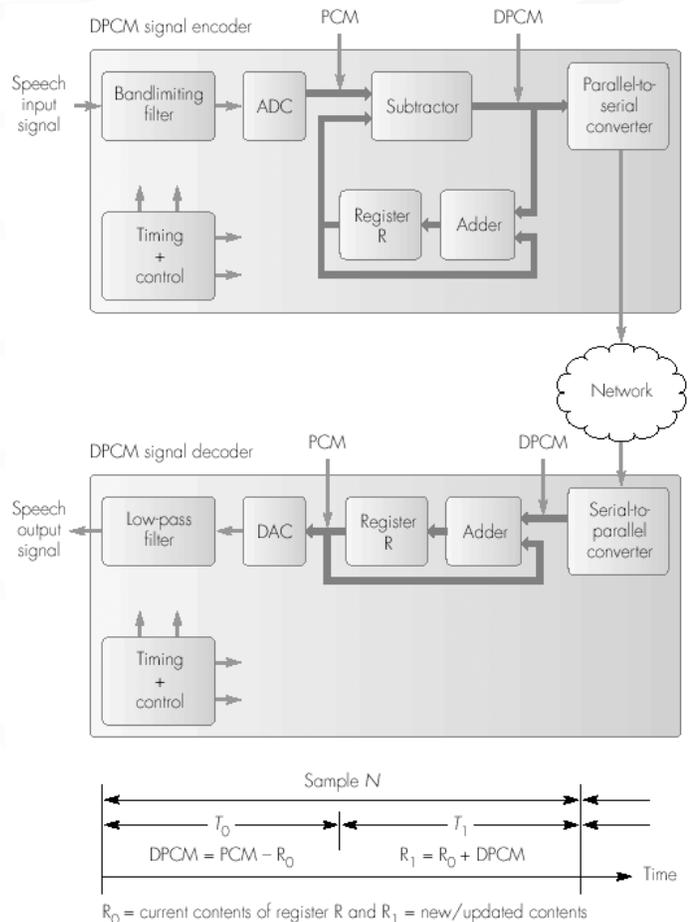
Compressão de Áudio

DPCM (*Differential PCM*)

- É derivado da codificação PCM
- Explora o fato de que a diferença de amplitude entre duas amostras PCM é menor que o valor a faixa de frequência do sinal.
 - Compressão é alcançada através do uso de um número menor de bits para armazenar a diferença entre amostras, ao invés do próprio valor (amplitude) das amostras.

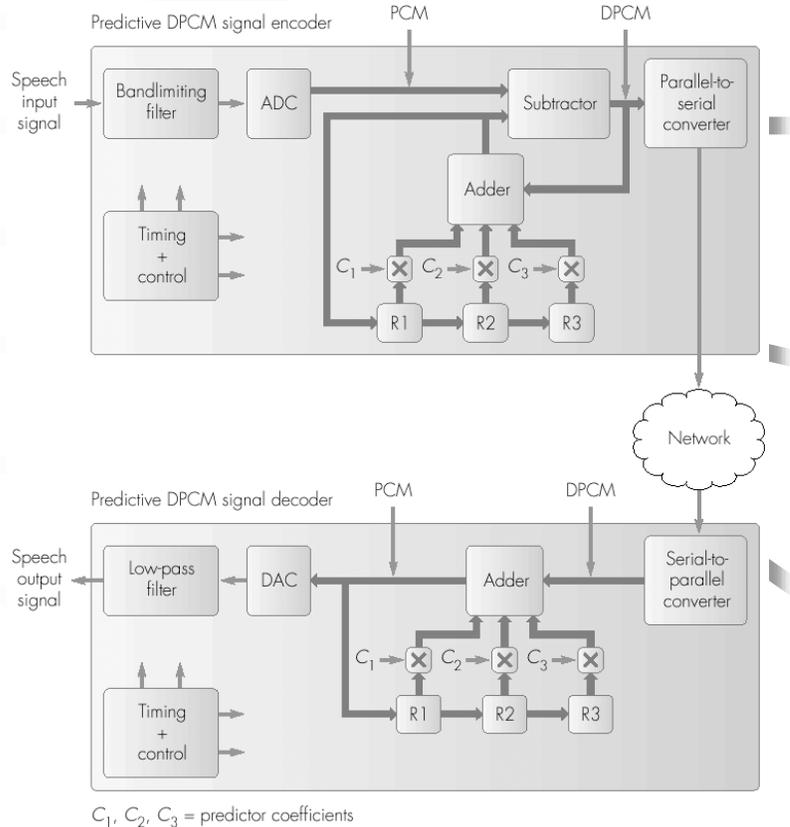
Differential PCM

- O valor das amostras DPCM dependem da precisão do valor armazenado no registrador R.
- Erros de quantização podem, potencialmente, se acumular, ocasionando um efeito cascata na distorção do sinal armazenado no registrador R.
- O valor armazenado no registrador é, portanto, apenas uma aproximação do valor real.
- Outras técnicas para o armazenamento do valor original foram propostas com o intuito de minimizar o problema.



Differential Pulse Code Modulation

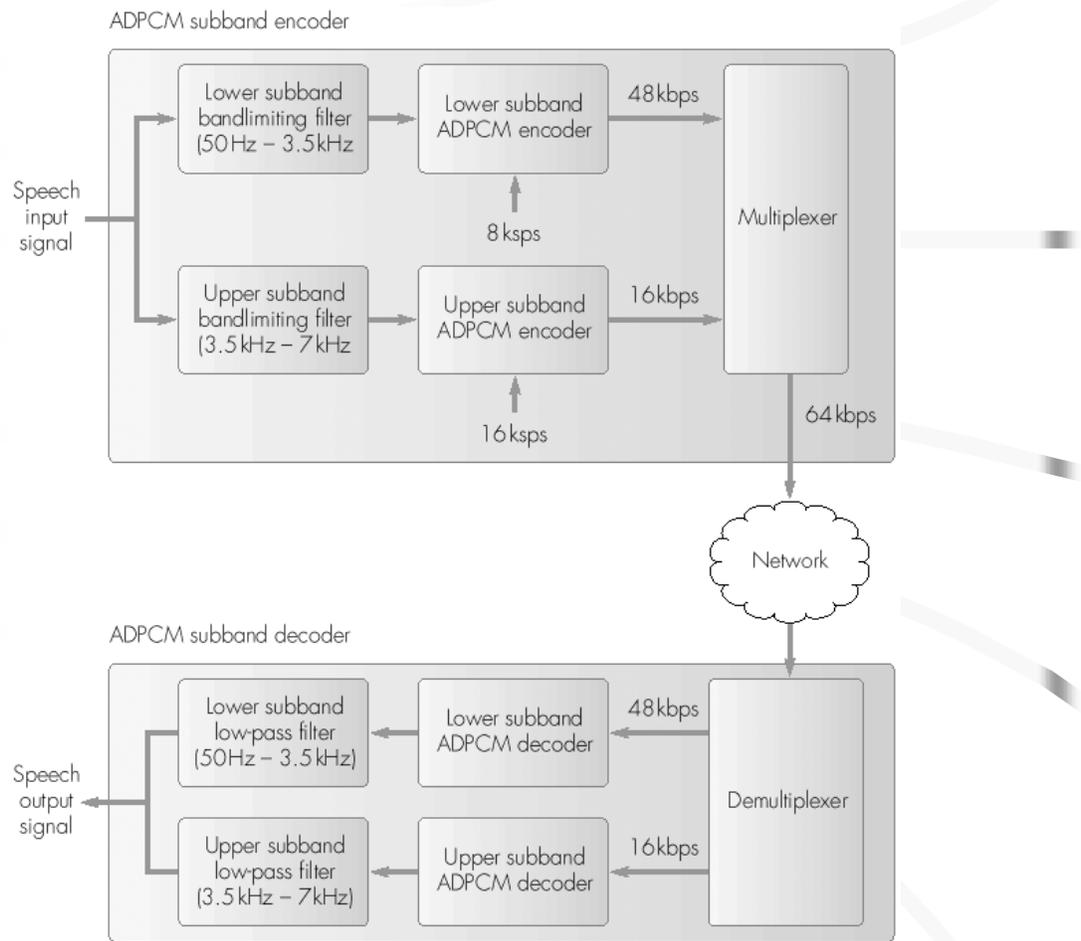
- Utiliza técnicas de estimativa (predição) para possibilitar que um valor que melhor representa a amplitude do sinal original seja utilizado pelo codificador.
- O valor é estimado com base em vários valores no histórico do sinal, que são armazenados através de uso de pesos conhecidos como coeficientes de predição.
- Taxa de bits de 32Kbps (6 bits por amostra) com qualidade similar a PCM.



Compressão de Áudio

ADPCM (*Adaptive Differential PCM*)

- ADPCM é baseado na idéia de utilizar um número menor de bits dependendo na amplitude da diferença das amplitudes das amostras resultantes da digitalização de um sinal de áudio.
- Padronizado como Recomendação ITU-T G.721.
 - Utiliza predição de oitava ordem ao invés de terceira ordem como DPCM.
 - Utiliza-se 6 bits por amostra para resultar em uma qualidade melhor que DPCM ou 5 bits por amostra quando maior compressão é necessária (16kbps)
- Recomendação ITU-T G.722, um segundo padrão ADPCM derivado do G.721, resulta em melhor qualidade de som às custas de maior complexidade no algoritmo.
 - Utiliza codificação por subbandas (*subband coding*).
 - A faixa de frequência do sinal é estendida para 50Hz a 7kHz



Compressão de Áudio

ADPCM (*Adaptive Differential PCM*)

- ITU-T G.722 pode operar também a taxas de transmissão de 56kbps e 48kbps, além dos 64kbps já indicados.
- Uma outra recomendação baseada na codificação ADPCM é a Recomendação ITU-T G.726
 - Também utiliza codificação por sub-bandas, mas no domínio de frequências de voz de 3.4kHz.
 - As taxas de transmissão nominais previstas são de 40, 32, 24 ou 16kbps.

Compressão de Áudio

APC (Adaptive Predictive Coding)

- Níveis de compressão ainda maiores podem ser alcançados através da utilização de algoritmos mais complexos, por exemplo através da adaptação dos coeficientes de predição que, portanto, mudam constantemente.
- O sinal de áudio é dividido em pequenos segmentos de tempo fixos e cada um é analisado separadamente. Para cada segmento um conjunto de coeficientes otimizado é selecionado.
- A codificação APC chega a comprimir o sinal de áudio a uma taxa de transmissão de 8kbps, ainda mantendo uma qualidade aceitável.

Compressão de Áudio

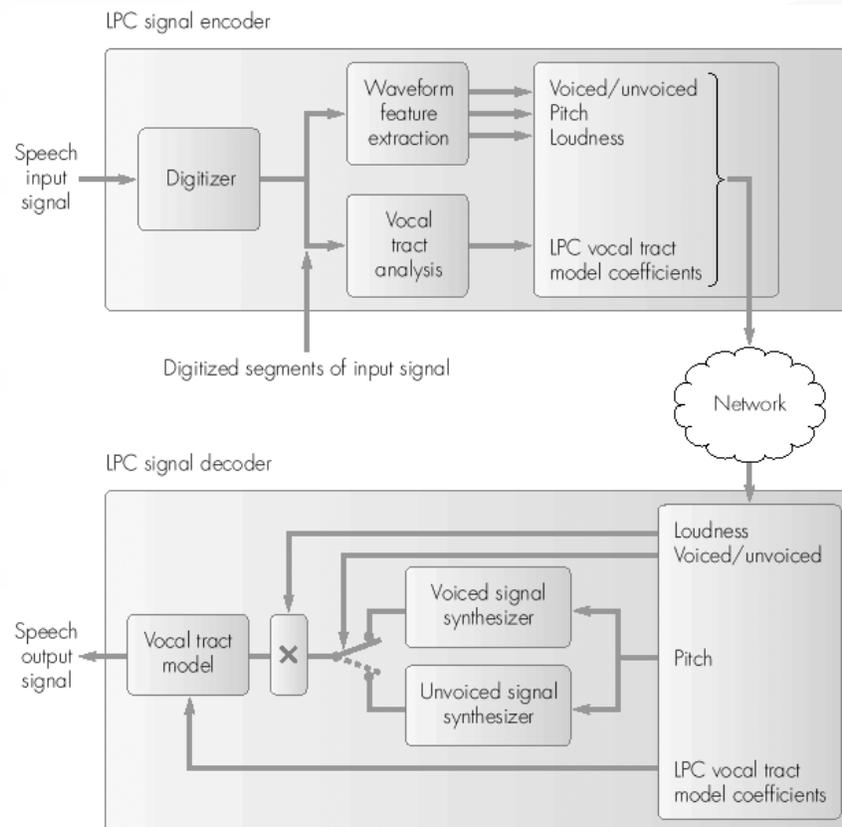
LPC (Linear Predictive Coding)

- Com a disponibilidade de circuitos de processamento digital de sinais de baixo custo, novas técnicas de compressão puderam ser empregadas.
- Uma destas técnicas consiste da análise da onda do sinal de áudio com o objetivo de identificar um conjunto de características de percepção do sinal. Tais características são então quantizados e utilizadas para gerar um sinal sonoro comparável ao sinal original. Tal som pode parecer sintético, mas continua inteligível.
- O conjunto de características do sinal escolhido tem importância fundamental na qualidade do som gerado por um decodificador LPC.
- Taxa de compressão muito alta, com cerca de 2.4kbps (ou mesmo 1.2kbps). Voz a tais taxas de transmissão tão baixas é bastante sintético, sendo mais usado em aplicações militares ou naquelas onde taxa de transmissão é um fator crucial a ser minimizado.

Compressão de Áudio

LPC (*Linear Predictive Coding*)

- As características principais que determinam a percepção de um sinal pelo ser humano são:
 - Tom de voz (*pitch*): Relacionado com a frequência do sinal. É importante porque o ser humano é mais sensível a frequências na faixa de 2 a 5kHz que em faixas de frequência mais altas ou mais baixas que tal faixa.
 - Período: Duração do sinal.
 - Volume: Quantidade de energia no sinal.
- Além dos parâmetros acima, as origens do som, denominados Parâmetros de Excitação do Trato Vocal, também são de importância. Tais características são classificadas como:
 - Sons vocais: Gerados pelas cordas vocais. Exemplos incluem os sons produzidos por m, v.
 - Sons não vocais: Quando gerados as cordas vocais estão abertas. Exemplos incluem os sons produzidos por f e s.



Compressão de Áudio

CELP (*Code Excited LPC*)

- Os sintetizadores utilizados na codificação LPC padrão são baseados num modelo simplificado do trato vocal. Um modelo mais sofisticado, conhecido como modelo de predição linear (CELP). Tal modelo é um exemplo de uma família de modelos conhecidos como *Enhanced Excitation (LPC) models*.
- É utilizado quando baixos requisitos de transmissão são cruciais, mas uma qualidade aceitável para um sistema multimídia é necessária.
- No modelo LPC cada segmento é tratado independentemente, diferentemente da codificação CELP, onde um conjunto limitado de segmentos é utilizado. Cada segmento válido é denominado um “*waveform template*”. Um conjunto pre-computado de *templates* são compartilhados pelo codificador e decodificador no que é conhecido como *template codebook*.

Compressão de Áudio

CELP (*Code Excited LPC*)

- Cada amostra digitalizada que compõe uma *template* são codificados por diferença. A *template* selecionada para codificar um determinado segmento é escolhida de modo que a diferença tenha um erro de quantização mínimo, o que resulta em uma melhor qualidade do sinal codificado.
- As recomendações da ITU-T G.728, G.729, G.729(A) e G.723.1, que cria uma boa qualidade de voz com grande taxa de compressão.
- Codificadores CELP introduzem uma série de retardos durante a codificação do sinal. Tais retardos devem ser considerados quando da seleção do algoritmo adequado a uma determinada aplicação.

Compressão de Áudio

CELP (*Code Excited LPC*)

- O retardo introduzido pela codificação PCM é proporcional ao intervalo de tempo entre amostras. A 8kHz o delay é de 0.125 ms - telefonia.
- G.728 tem um retardo total de 0.625ms, com transmissão a 16kbps - telefonia com baixa taxa de bits.
- G.729 tem um retardo total de 25ms, com transmissão a 8kbps - telefonia celular.
- G.729(A) idem ibidem - transmissão de voz e dados (DSDV).
- G.723.1 tem um retardo total de 67.5ms, com transmissão de 5.3/6.3kbps - vídeo e áudio na Internet.

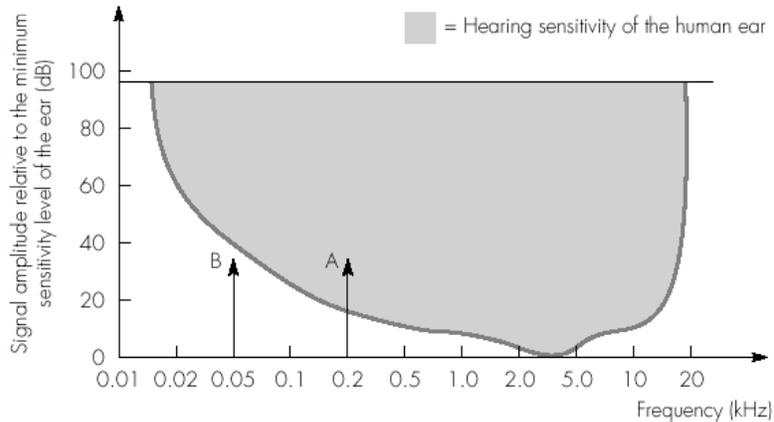
Compressão de Áudio

Perceptual Coding

- Tanto LPC como CELP são usados principalmente para aplicações de telefonia, motivo pelo qual os mesmos são orientados para compressão de voz. *Perceptual Coding* (PC), por sua vez, foram desenvolvidos para codificação de sinais de áudio em geral (TV digital, etc.). PC também utiliza um modelo para a codificação dos sinais de áudio: O modelo psicoacústico, que explora as limitações do aparelho auditivo humano. O processo de codificação do sinal é semelhante àquele realizado em codificações CELP, porém somente as características do sinal que são relevantes para a percepção humana são transmitidas, tais características de interesse do sinal incluem:
 - Sensibilidade da Audição
 - Mascaramento de Frequência
 - Mascaramento Espacial

Perceptual Coding

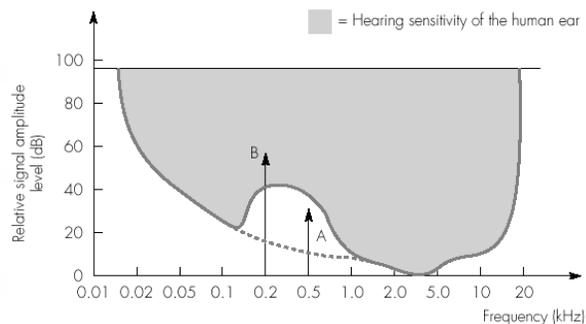
Modelo Psicoacústico - Sensibilidade da Audição



- Sensibilidade do Ouvido Humano: é a razão entre a amplitude máxima e mínima do sinal, medida em dB.

Perceptual Coding

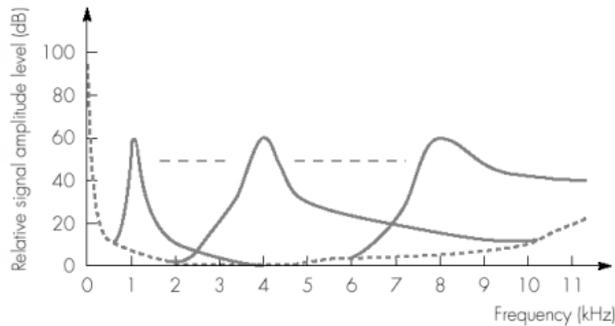
Modelo Psicoacústico - Mascaramento de Frequência



- Mascaramento de Frequência: A sensibilidade da audição humana muda conforme a amplitude relativa do sinal. A figura acima mostra como a sensibilidade muda na vizinhança de um som de amplitude alta, caso no qual o som A não será escutado, embora ele o fosse se ocorresse sozinho.

Perceptual Coding

Modelo Psicoacústico - Mascaramento de Frequência

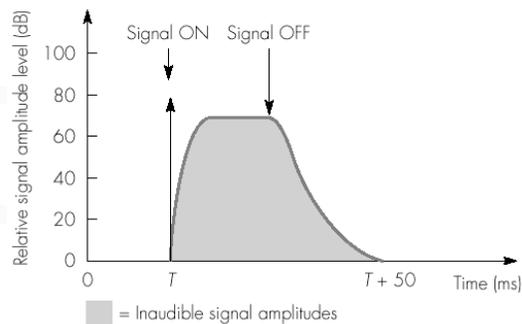


- O efeito de mascaramento de frequência varia de acordo com a frequência em questão.
- A largura (eixo x) das curvas de frequência é conhecido como largura de banda crítica, e define a faixa de frequências afetadas por mascaramento de frequência resultante de um dado sinal.
 - Para frequências abaixo de 500Hz a largura de banda crítica é de 100Hz.
 - Para frequências acima de 500Hz a largura de banda crítica aumenta aproximadamente linearmente em múltiplos de 100Hz. Um sinal de 1kHz (2×500) teria uma LBC de 200Hz (2×100) enquanto que um sinal de 5kHz (10×500) teria uma LBC de 1kHz (10×100).
 - Pode-se, portanto, calcular as frequências afetadas, que não precisam ser enviadas.

Perceptual Coding

Modelo Psicoacústico - Mascaramento Temporal

- Após escutar um som alto, o ouvido humano fica surdo a sons mais baixos por um intervalo de tempo. Tal fato é denominado mascaramento temporal.
- Depois que um som alto termina são necessários alguns décimos de segundo para o ouvido humano “se recuperar” e a amplitude do sinal ouvido reduza. Neste intervalo de tempo o ser humano é incapaz de escutar sons com amplitude abaixo do limite indicado na figura acima. Tais sons não precisam, portanto, ser transmitidos.



Codificadores de Áudio MPEG

- O *Motion Pictures Expert Group* (MPEG) foi formado pela ISO para criar uma série de padrões relacionados a aplicações multimídia que transmissão de áudio e vídeo. Os codificadores relacionados com compressão de áudio são conhecidos como codificadores de áudio MPEG.
- Alguns codificadores de áudio MPEG utilizam *Perceptual Coding*. As operações de processamento de sinal em um codificador de áudio MPEG são realizadas digitalmente (um esquema aparece a seguir).
- Padrão ISO 11172-3, que explicita a definição de 3 níveis de processamento do sinal, cada um com mais informações sobre o sinal de áudio (camada 1, por exemplo não apoveita máscara temporal.

