

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 Vídeo

- Taxas de transmissão de Vídeo Não Comprimido Superior às taxas de transmissão do meio físico utilizado na maioria dos casos:
 - SQCIF video alcança 10MBps
 - Televisão Digital ($YC_R C_B$ 4:2:0) chega a 162MBps
- Compressão é utilizada em todas as aplicações que envolvem transmissão (digital) de vídeo.
 - Existem vários padrões de compressão, indicados para as diferentes características de aplicações

Compressão de Vídeo

Introdução

- Vídeo é, por vezes, chamado “*moving picture*”, uma vez que o mesmo consiste da apresentação de imagens estáticas a uma frequência alta o suficiente para levar à impressão de continuidade do movimento pelo olho humano.
 - Cada imagem em um vídeo é chamado um quadro (*field*), ou imagem.
- Um mecanismo extremamente simplificado para codificação de vídeo consiste da codificação dos vários quadros, independentemente, através de um codificador JPEG. Esta metodologia é conhecida como MJPEG (*motion JPEG* ou *moving JPEG*).
 - A taxa de compressão neste caso é a mesma de uma imagem JPEG, na faixa de 1:10 a 1:20, o que não é o suficiente para diminuir significativamente as necessidades de banda passante de transmissão (televisão digital continuaria com cerca de 16Mbps).

Compressão de Vídeo

Redundância Espacial e Temporal

- Padrões de compressão de imagem como JPEG, operam sobre a redundância que é esperada entre os pixels vizinhos em um pequeno bloco de 8×8 pixels. Este tipo de característica de uma imagem é conhecida como Redundância Espacial, visto que a redundância ocorre no âmbito dos pixels.
- Ao se considerar um vídeo, porém, deve-se facilmente concordar que existe uma grande similaridade entre dois quadros vizinhos (ou próximos) de uma sequência de vídeo. Tal característica de quadros é conhecida como Redundância Temporal, uma vez que ela ocorre em relação a quadros que foram gerados em pequeno intervalo de tempo.
 - Em um filme usualmente as cenas tem pelo menos 3 segundos, que a 30 fps resultaria em um conjunto de pelo menos 90 quadros com grande redundância temporal.
 - Em uma transmissão de um telejornal, o que se move é, basicamente, o interlocutor.

Compressão de Vídeo

Estimação e Compensação de Movimento

- A técnica utilizada para aproveitar a redundância temporal é a predição do conteúdo de vários quadros a partir de uma combinação de um quadro anterior (e em alguns casos um posterior também).
 - Ao invés de transmitir os quadros como uma sequência de imagens não relacionadas entre si, somente algumas imagens são enviadas neste modo e para as demais somente a diferença entre o quadro e o quadro predito (*predicted*) são enviadas.
 - A precisão da predição depende do quão boa é a estimativa de movimento de um quadro para outro. Esta operação é conhecida como Estimativa de Movimento (*motion estimation*).
 - Como a estimativa de movimento não é precisa, também precisam ser enviados valores de diferença entre a posição estimada e a posição real dos elementos. Tal operação é conhecida como compensação do movimento (*motion compensation*).

Compressão de Vídeo

Tipos de Quadro

- Quadros codificados independentemente, ou quadros I (*intracode, I-frame*).
- Quadros codificados por predição:
 - Predição simples ou quadros P (*predictive, P-frames*).
 - Predição bidirecional ou quadros B *bidirectional, B-frames*
Também conhecidos como *intercoded frames* ou *interpolation frames*.

I

B

P

Compressão de Vídeo

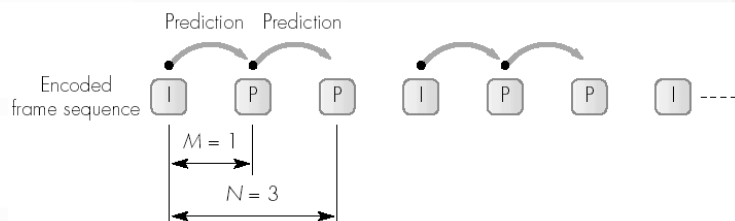
Quadros I

- São codificados sem referência a qualquer outro quadro, i.e. cada quadro é tratado como uma imagem (digitalizada) separada. as matrizes componentes Y , C_R e C_B são codificadas independentemente usando o algoritmo similar a JPEG (Transformada Discreta de Cosenos, Quantização, Codificação por Entropia), com a diferença que todos os coeficientes DCT são quantizados utilizando o mesmo valor constante. Quadros I são pouco comprimidos, portanto.
- Embora possa parecer vantajoso ter um único quadro I no início de cada cena do vídeo, algoritmo é independente do conteúdo do vídeo, i.e. ele não identifica cenas. Ademais, quadros I precisam estar presentes com certa frequência para permitir que o decodificador possa continuar decodificando o vídeo mesmo quando há erro de transmissão de um quadro I.
- O número de quadros entre quadros I sucessivos é conhecido como Grupo de Imagens (*group of pictures*), que é expresso pelo símbolo N e pode usualmente assumir valores entre 3 e 12.

Compressão de Vídeo

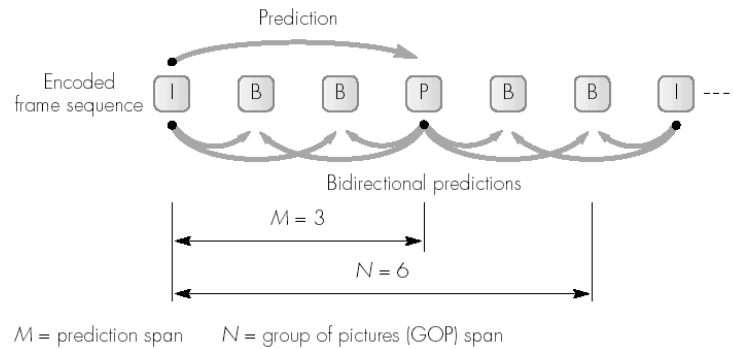
Quadros P

- A codificação de um quadro P é feito em relação ao quadro I ou P anterior. Tal codificação é feita através de uma combinação de estimação de movimento e compensação de movimento, o que leva a uma taxa de compressão elevada.
- Erros contidos em um quadro P são propagados para todos os quadros P posteriores, motivo pelo qual o número de quadros P antes de um novo quadro I ser limitado. O número de quadros P antes de um quadro I (*prediction span*) é denotado por M e varia de 1 a 3.
- Estimação de movimento limita-se a uma área reduzida, o que serve para a maioria dos vídeos onde o movimento é suave. Quando o movimento é menos suave a predição direta não é tão eficiente, motivo pelo qual existe a predição bidirecional.



Compressão de Vídeo

Quadros B



- Codificação Bidirecional utiliza buscas na estimação de movimento em ambos o primeiro quadro I ou P anterior e o primeiro quadro I ou P posterior ao quadro sendo codificado.
- Além de funcionar melhor para o caso no qual um objeto se move mais rapidamente, tal mecanismo também é mais adequado para sequências nas quais um objeto se move para frente (ou trás) de outro objeto.
- Erros em quadros B não se propagam.
- Quadros B são os mais “caros” no ponto de vista de processamento.

Compressão de Vídeo

Ordenação de Quadros I, P e B

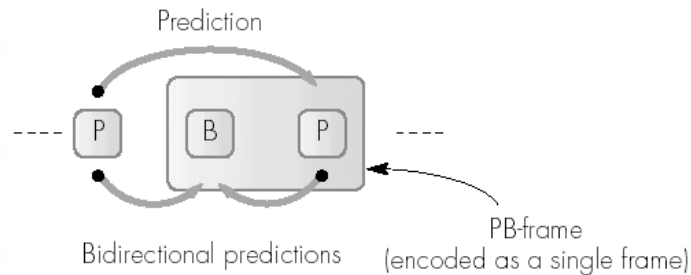
- Quadros I podem ser decodificados imediatamente na ordem de chegada para que o quadro original seja reconstruído.
- Quadros P, para serem reconstruídos, são decodificados e processados juntamente com a informação referente ao quadro I ou P anterior.
- Quadros B, por sua vez, por necessitarem de informação futura precisariam ser armazenados em buffer até que o próximo quadro I ou P seja decodificado para só então serem os vários quadros B serem finalmente reconstruídos.

Para minimizar o tempo necessário para decodificar um quadro B, os quadros de um vídeo mudam de ordem durante a codificação de modo que os quadros I futuros precedem os quadros B que dele dependem:

- IBBPBBPBB I...
- IPBBPBB IBBPBB...

Compressão de Vídeo

Quadros PB



- Consiste de um quadro P e um quadro B codificados como se fossem um quadro só.
- Permite que se aumente a taxa de quadros sem aumentar consideravelmente a taxa de bits

Compressão de Vídeo

Quadros D

- Somente utilizados em aplicações específicas: video-sob-demanda.
- Permite que o usuário execute operações de avanço, retrocesso, etc.
 - O vídeo precisa ser decodificado mais rapidamente.
 - Quadros D são incluídos entre os quadros I, P e B de modo a possibilitar a execução destas operações.
- Quadros D são quadros altamente comprimidos que são ignorados durante a decodificação I, B, P tradicional.
- Aproveita o fato de que o coeficiente DC contém a cor média de um bloco. Somente o coeficiente DC é utilizado para gerar uma imagem de baixa resolução (com altíssima taxa de compressão e velocidade de reconstrução), o que é necessário para operações FF, REW.

Compressão de Vídeo

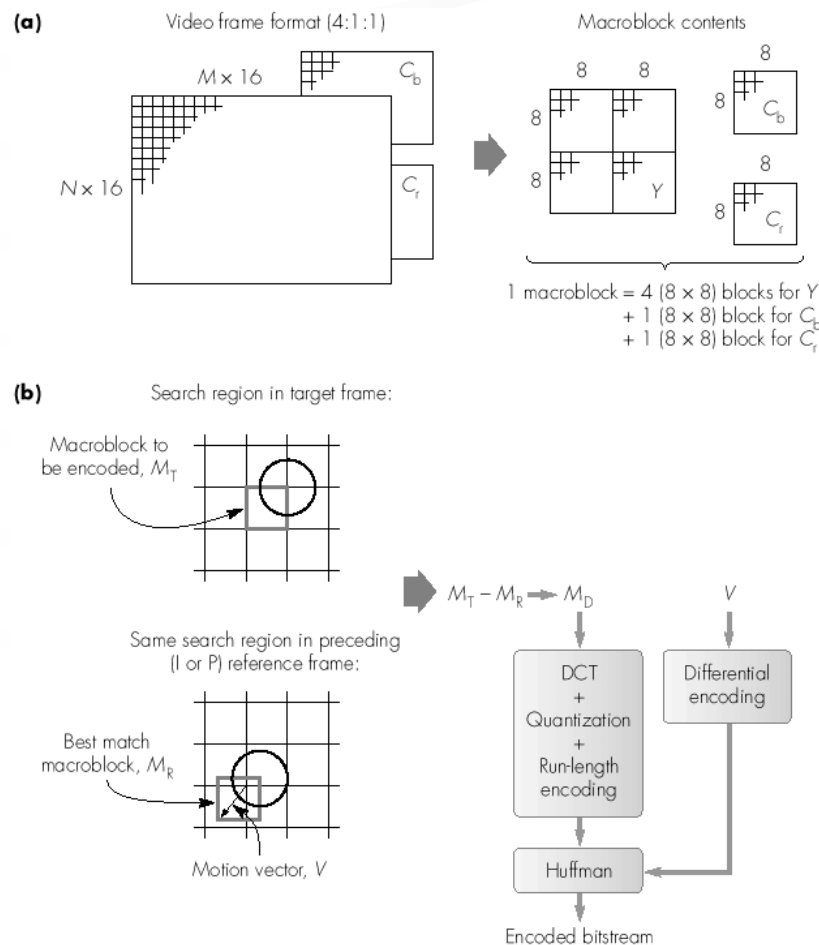
Estimação de Movimento - Quadro P

- O conteúdo dos quadros P e B passam pela estimação de movimento:
 - O conteúdo da matriz Y são divididos em matrizes de 16×16 pixels, denominados macroblocos. No caso da codificação 4:1:1 as matrizes dos componentes de crominância são matrizes 8×8 . O macrobloco, portanto, consiste de um conjunto de 4 blocos de luminância e 1 bloco de cada um dos componentes de crominância.
 - Na codificação de um quadro P, cada macrobloco é comparado pixel a pixel com o macrobloco correspondente no quadro de referência. Se for encontrado um bom grau de redundância, somente o endereço do macrobloco é codificado, caso contrário a busca segue na área ao redor do macrobloco de referência (incluindo outros macroblocos).
 - O modelo não define a área onde a busca deve ser efetuada, nem o mecanismo de busca a ser empregado. Somente o modo como o resultado da busca deve ser codificado é definido.
 - Normalmente só o componente Y é utilizado na estimação do movimento.

Compressão de Vídeo

Estimação de Movimento - Quadro P

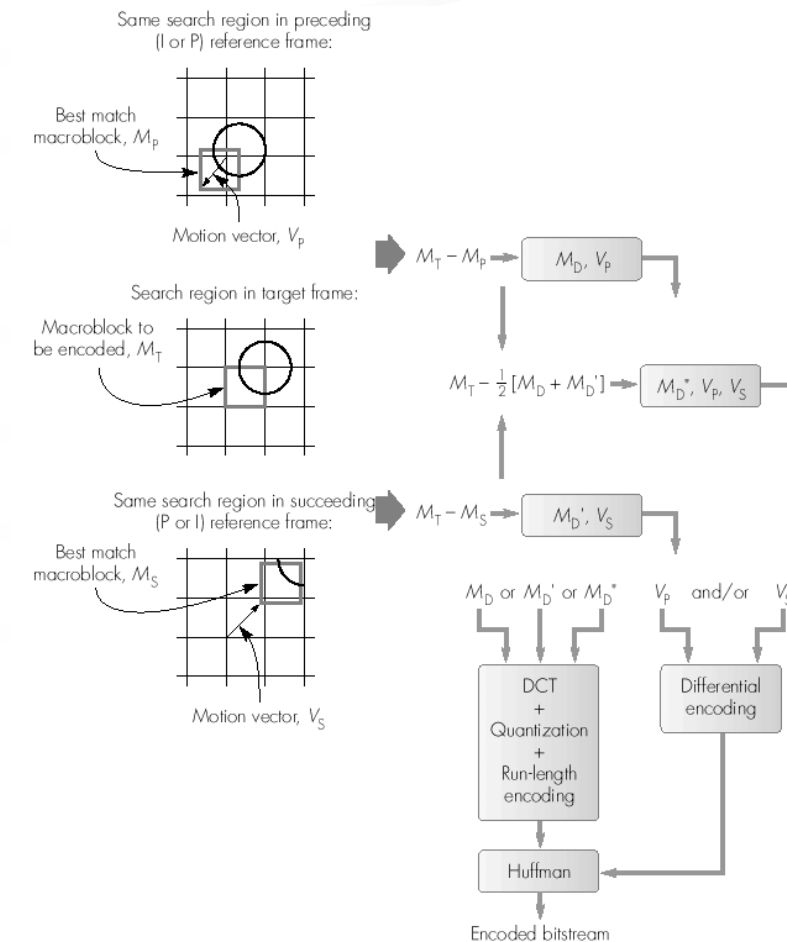
- Diz-se que um bloco é achado quando a média do erro absoluto entre o valor dos pixels do mesmo estiver abaixo de um determinado limite.
- Uma vez que um bloco é encontrado na estimação de movimento apenas dois parâmetros são codificados:
 - O Vetor de Movimento (*motion vector*), indica a posição (x,y) do macrobloco em relação à posição original, bem como a posição original.
 - A busca pode ser feita no domínio de blocos ou no domínio de pixels.
 - Erro de Predição (*prediction error*), que é representado por três matrizes (para componentes de luminância e crominância) que representam as diferenças entre aqueles no macrobloco destino e origem.
 - Como vetores de movimento pode conter números grandes e um objeto usualmente contém vários macroblocos (que terão vetores de movimento similares), tais vetores passam por codificação diferencial e Huffman.
 - As três matrizes de diferença são codificadas similarmente a quadros I, com DCT, quantização e codificação por entropia.
- Macroblocos não encontrados são codificados independentemente.



Compressão de Vídeo

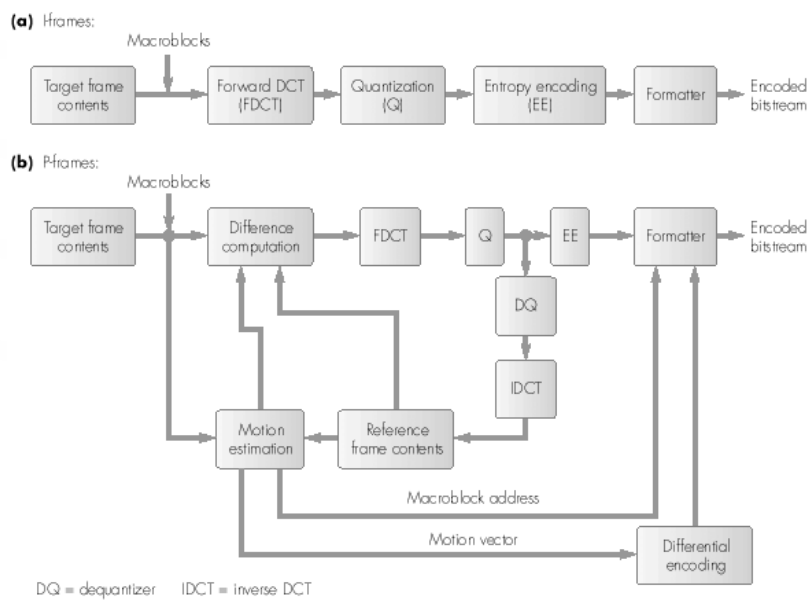
Estimação de Movimento - Quadro B

- Para quadros B tanto o quadro (I ou P) anterior quanto o posterior são considerados. Inicialmente a busca é feita utilizando o quadro anterior e em seguida usando-se o quadro posterior.
- Um terceiro vetor de movimento é então calculado utilizando a média dos dois vetores calculados no passo anterior.
- O conjunto com as menores matrizes de diferenças são selecionados.
- O restante da codificação é realizado de modo semelhante àquele utilizado na codificação de quadros P.
- Diz-se que o vetor de movimento tem uma precisão de uma fração de um pixel, e.g. resolução de meio-pixel (*half-pixel*).



Compressão de Vídeo

Unidades da Codificação de Quadros I e P



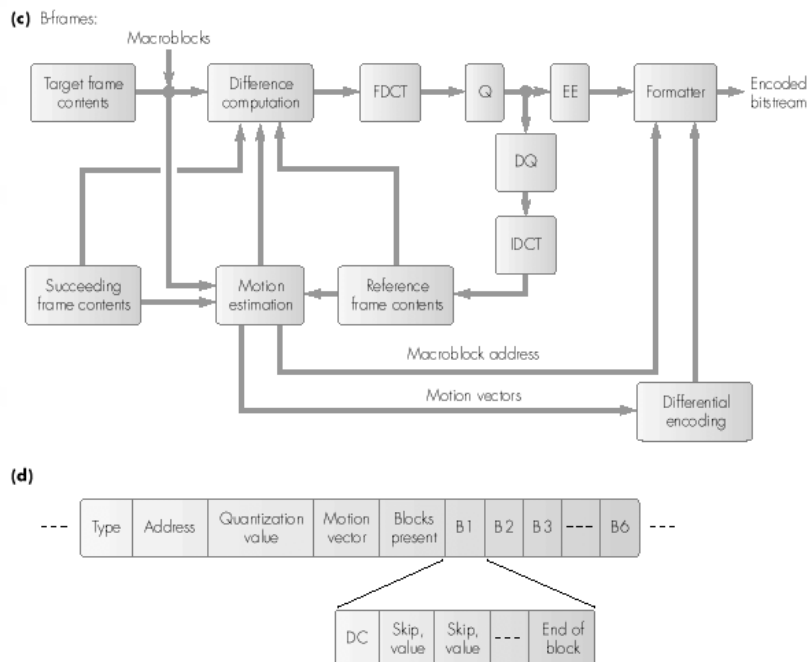
Compressão de Vídeo

Detalhes de Implementação - I e P

- A codificação de macroblocos que compõem quadros I é a mesma da codificação de blocos de uma imagem JPEG: DCT, Quantização e Codific. por Entropia. Cada macrobloco contém 6 blocos a codificar.
- A codificação de macroblocos (MB) que compõe quadros P depende do resultado da estimação de movimento, que por sua vez depende do conteúdo do macrobloco sendo codificado e do conteúdo do macrobloco de referência e da área de busca, podendo ser:
 - Se os conteúdos são idênticos, somente o endereço do MB é codificado.
 - Se os conteúdos são bastante próximos, tanto o vetor de movimento quanto as matrizes de diferença são codificadas.
 - Se não houver similaridade na área de busca o macrobloco é codificado como no caso de um macrobloco componente de um quadro I.
- Para decodificar MBs componentes de quadros P, é necessário utilizar uma cópia do quadro de referência. Tal operação envolve a decodificação do quadro de referência (dequantização e IDCT) ser aplicada aos dados oriundos da codificação do quadro de referência.

Compressão de Vídeo

Codificação de Quadros B e *Bitstream* Típica



Compressão de Vídeo

Detalhes de Implementação - B

- O formatador (*formatter*) é responsável por identificar como cada macrobloco foi codificado.
- O campo tipo (*type*) indica o tipo de quadro codificado (I, P ou B).
- O campo endereço (*address*) identifica a localização do macrobloco no quadro.
- O campo *quantization value* indica o limite utilizado na etapa de quantização.
- *Motion Vector* é o vetor de movimento codificado, caso exista um.
- *Blocs Present* indica quais dos seis blocos de 8×8 pixels que compõe o macrobloco estão presentes.
- Para aqueles blocos presentes os coeficientes DCT codificados de cada bloco seguem.
- Um codificador deste tipo gera um fluxo de dados VBR.
- A decodificação é BEM mais simples que a codificação em si.

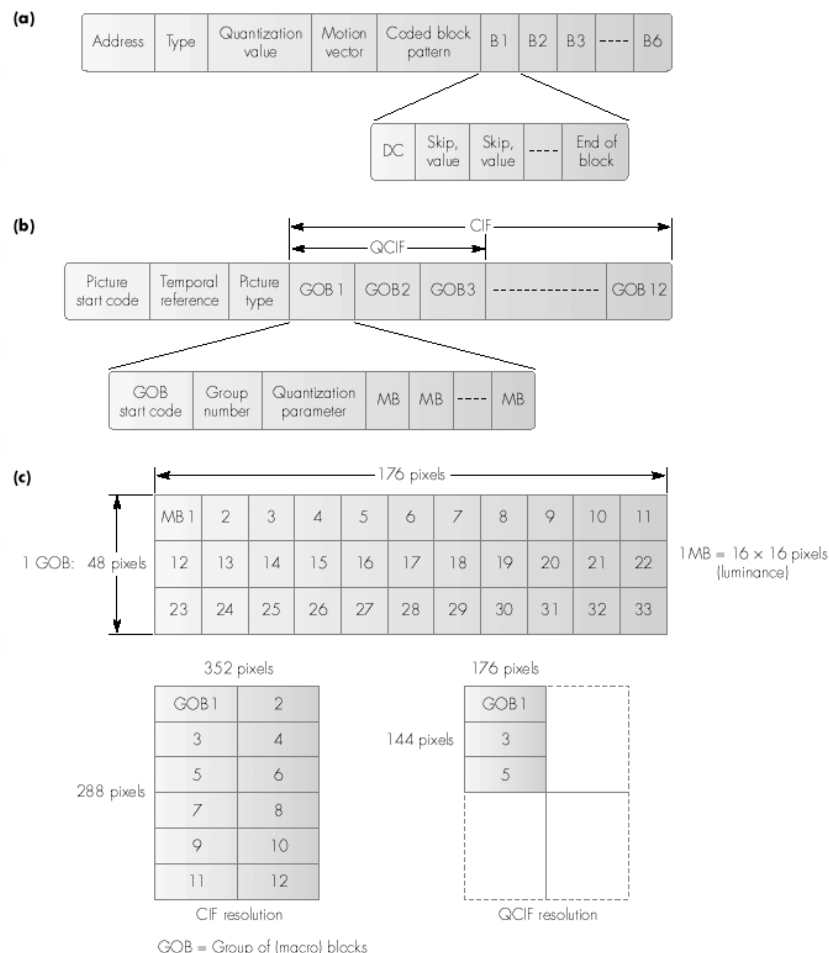
Compressão de Vídeo

Desempenho

- Quadros I
 - A codificação de quadros I alcança taxa de compressão similar àquela de imagens JPEG, da ordem de 10:1 a 20:1.
 - O tempo de processamento para comprimir um quadro I é relativamente baixo.
- Quadros P
 - A codificação de Quadros P depende muito do método de busca de blocos na estimação de movimento, usualmente chegando a compressão da ordem de 20:1 a 30:1.
 - O tempo de processamento para comprimir quadros P é intermediário.
- Quadros B
 - A codificação de Quadros B depende muito do método de busca de blocos na estimação de movimento, usualmente chegando a compressão da ordem de 30:1 a 50:1.
 - O tempo de processamento para comprimir quadros B é relativamente alto.

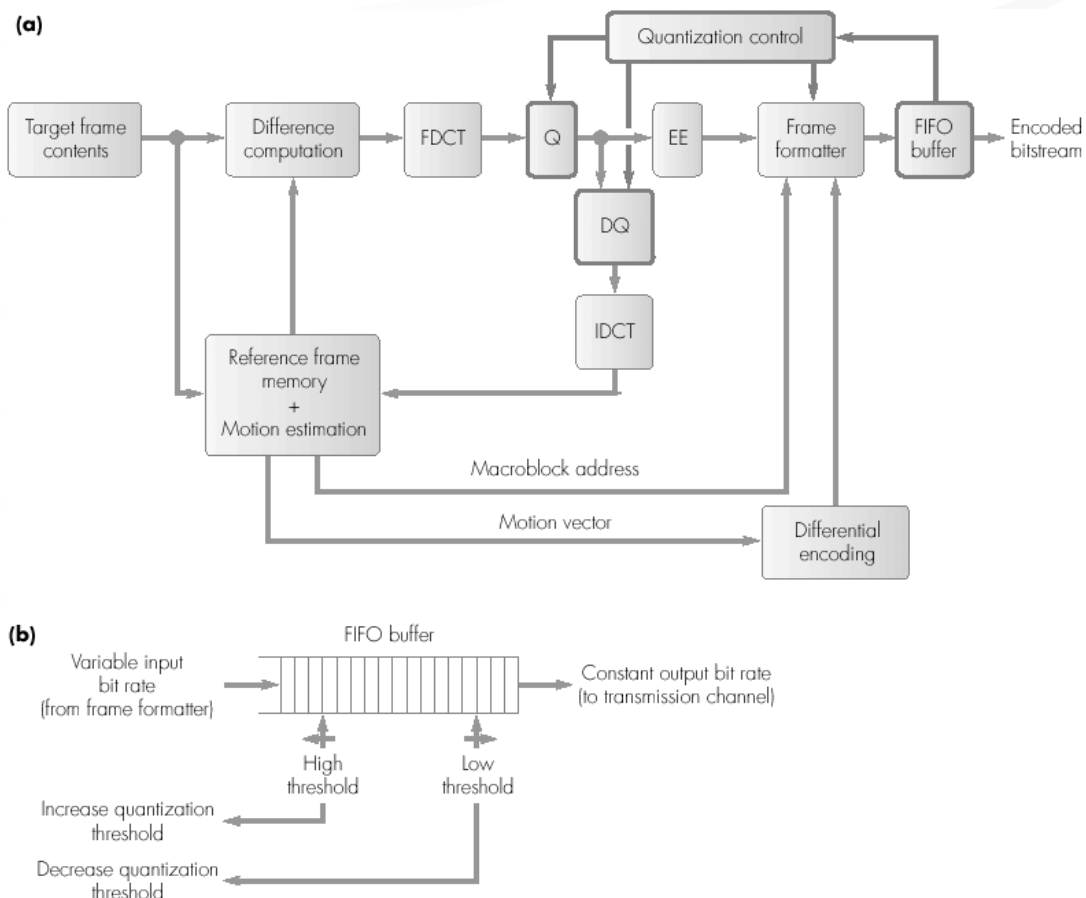
H.261 ou $p \times 64$

- Padrão de compressão de vídeo definido pela ITU-T objetivando a implementação de serviços de videotelefonia e videoconferência sobre Redes Digitais de Serviços Integrados (RDSI ou *ISDN*) - canal com taxa de transmissão múltipla de 64kbps, o padrão também é conhecido como $p \times 64$, onde p varia de 1 a 30 (64 a 1920kbps). (*Video codec for audiovisual services at $p \times 64$ kbit/s*)
- Define dois tamanhos de imagem:
 - CIF: $Y=352 \times 288$, $C_b = C_r = 176 \times 144$
 - QCIF: $Y=176 \times 144$, $C_b = C_r = 88 \times 72$
- Os quadros são divididos em macroblocos de 16×16 pixels.
- Somente quadros I e P são utilizados pelo padrão, objetivando a realizar a compressão em menor tempo, visto que as aplicações em questão são aplicações em tempo real. Três quadros P são utilizados entre quadros I.



H.261 ou $p \times 64$

- O início de cada quadro inicia com um PSC (*Picture Start Code*).
- Cada quadro tem uma referência de tempo associada ao mesmo, de modo que pode-se calcular o exato momento em que o quadro deve ser apresentado.
- Macroblocos são agrupados no que se chama Grupo de Blocos (GOB). Um grupo de blocos consiste de uma matrix de 3×11 macroblocos (CIF contem 6 GOBs, 2×3 , QCIF contém 3 GOBs, 1×3).
- Cada GOB também tem um marcador de início.
- Em caso de erro de transmissão o decodificador pode aguardar o próximo *GOB start code* ou PSC, por este motivo os *Start Codes* também são conhecidos como marcadores de resincronização.
- CBR é obtida variando-se o limite de quantização. Tal operação é realizada no *FIFO buffer*. Existem dois limites, que são utilizados para controlar o aumento ou diminuição do nível de quantização, objetivando manter a taxa de transmissão constante (CBR).



H.263

- Padrão de compressão de vídeo definido pela ITU-T objetivando a implementação de aplicações de vídeo sobre redes sem fio e PSTN. Tais aplicações incluem videotelefonia, videoconferência, segurança ... (*Video coding for low bit rate communication*).
- PSTN requer modems para envio de dados digitais, o que limita a taxa de transmissão para algo em torno de 28kbps-56kbps.
- Baseado no padrão H.261, mas com qualidade de imagem superior (na prática H263 tem substituído, com vantagens, o H.261 na maioria das aplicações. Com menos de 64kbps H261 produz efeito de blocagem, que é explicado por uma quantização muito agressiva no *FIFO buffer*).
- H263 inclui um conjunto de opções avançadas de codificação que permitem uma taxa de compressão extremamente alta.
- O padrão não limita o número de quadros entre quadros I sucessivos.

H263

Formatos e Tipos de Quadro

- CIF: $Y=352 \times 288$, $C_b = C_r = 176 \times 144$
- QCIF: $Y=176 \times 144$, $C_b = C_r = 88 \times 72$
- SQCIF: $Y=128 \times 96$, $C_b = C_r = 64 \times 68$
- 4CIF: $Y=704 \times 576$, $C_b = C_r = 356 \times 288$
- 16CIF: $Y=1408 \times 1152$, $C_b = C_r = 704 \times 576$
- H263 possui quadros dos tipos I, P e PB.

H.263

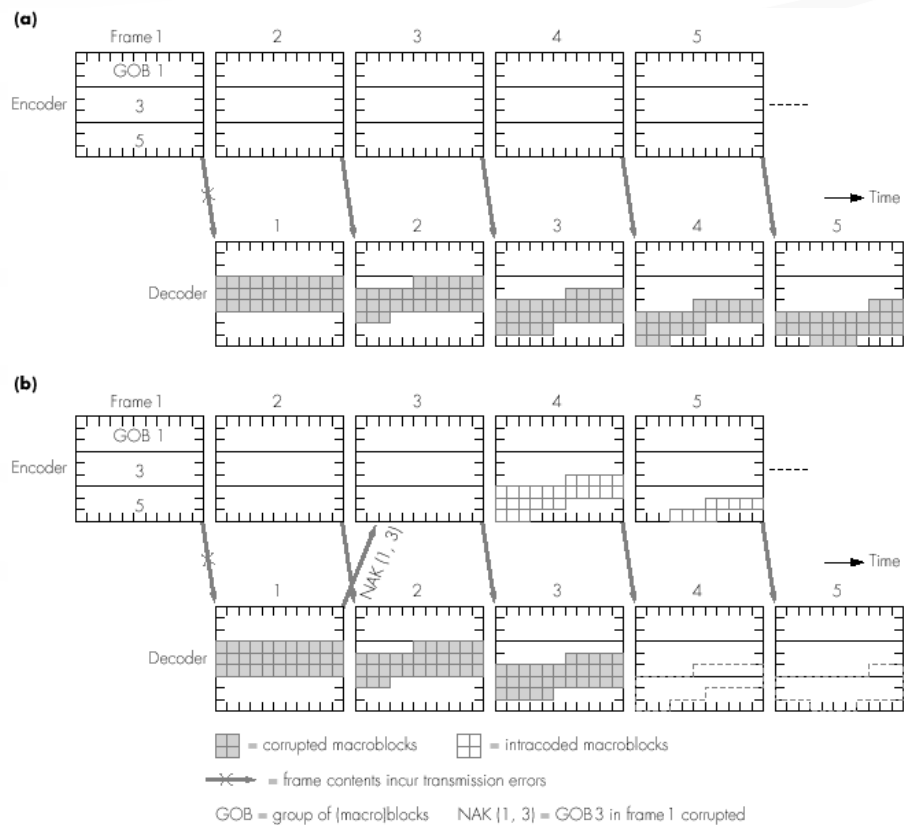
Opções Avançadas

- *Unrestricted Motion Vector*
 - Vetores de movimento associados com um macrobloco usualmente são limitados a uma área bem definida em volta do macrobloco sendo codificado. Ademais, a área é usualmente restrita à aresta da imagem. H.263, por sua vez, introduz o modo **opcional** de vetor de movimento irrestrito, que permite que o vetor aponte inclusive para fora da imagem.
- *Error Resilience*
 - Permite que o decodificador consiga identificar um erro e corrigir tal erro, por exemplo, repetindo partes da imagem anterior. Tal característica é importante considerando o fato de que o número de quadros entre quadros I consecutivos pode ser bastante grande, pois o padrão não especifica um número máximo.

H.263

Opções Avançadas

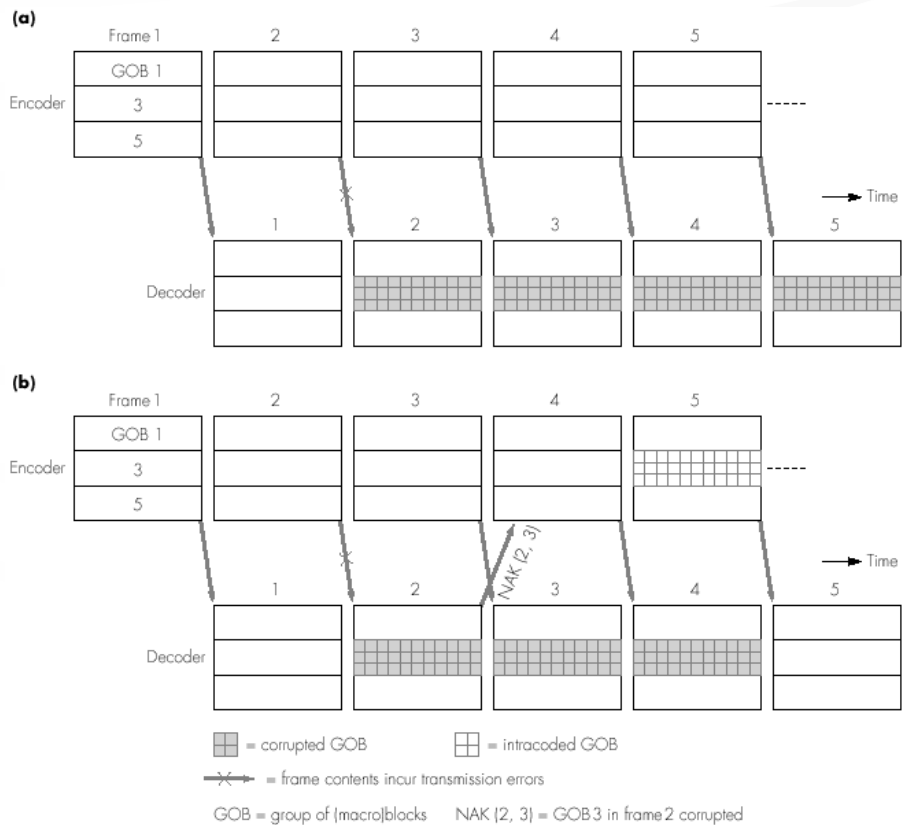
- *Error Tracking*
 - Aproveitando a comunicação entre transmissor e receptor, o receptor indica ao transmissor que um erro em um GOB foi detectado (através de NAK).
 - Posteriormente, partes de um GOB são transmitidos no modo *intra-mode*, após análise de possível propagação do erro.
 - Erros em um GOB são identificados quando, entre outros, acontece:
 - Um ou mais vetores de movimento fora da faixa.
 - Um ou mais palavras de tamanho variável (Huffman) inválidas.
 - Um ou mais coeficientes DCT fora de faixa.
 - Um número excessivo de coeficientes em um macrobloco.



H.263

Opções Avançadas

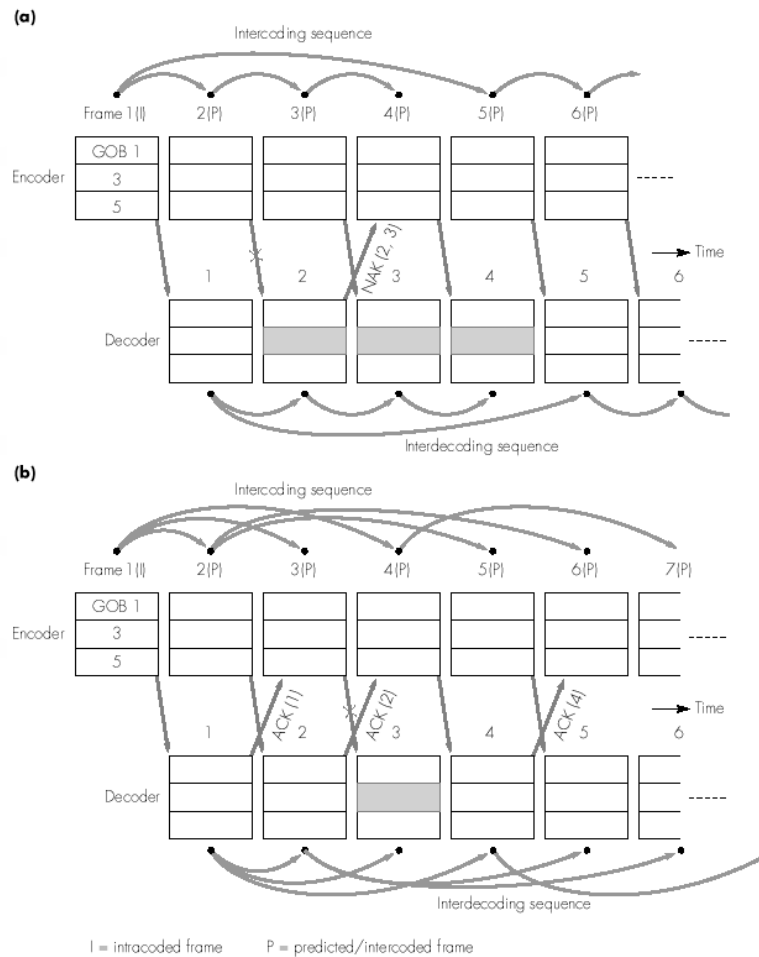
- *Independent Segment Decoding*
 - O objetivo desta técnica é evitar que erros em um GOB se propagem para GOBs vizinhos nos quadros seguintes. Isto é implementado através da partição de um vídeo em subvídeos definidos pelos GOBs.
 - Esta técnica limita os vetores de movimento aos limites de um GOB.
 - Por conta da limitação na estimação de movimento esta técnica não costuma ser utilizada sozinha. Ela é mais empregada em conjunção com *error tracking* ou, mais frequentemente, com *reference picture selection*.



H.263

Opções Avançadas

- *Reference Picture Selection*
 - Similar a *error tracking*, uma vez que o mesmo também objetiva evitar a propagação de erros.
 - Possui dois modos de operação
 - Modo NAK: O decodificador pode selecionar qualquer um dos quadros decodificados anteriormente
 - Modo ACK: Todos os quadros recebidos sem erro são confirmados através de envio de ACK, somente os quadros que foram confirmados são utilizados como quadros de referência.



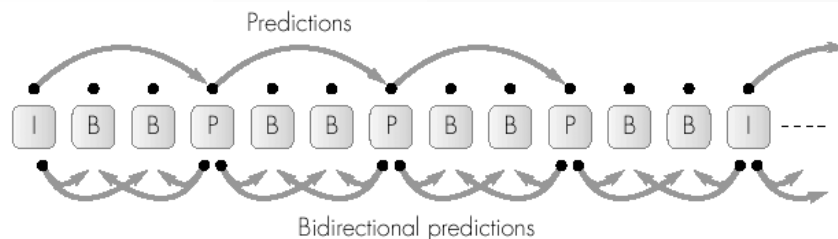
MPEG

- O grupo *Motion Pictures Expert Group* (MPEG) foi formado pela ISO para desenvolver um conjunto de padrões para aplicações multimídia que utilizam vídeo e som.
 - MPEG-1
 - Recomendação ISO 11172. Resolução de vídeo baseada no formato SIF (Source Intermediate Format) com 352×288 pixels. Objetiva armazenar vídeo com qualidade VHS em CDROM com taxas de 1.5Mbps.
 - MPEG-2
 - Recomendação ISO 13818 (H.262). Objetiva o armazenamento e transmissão de áudio e vídeo com qualidade de studio. 4 Resoluções de vídeo:
 - Baixa, baseado no formato SIF de 352×288 pixels. Compatível com MPEG-1. Áudio com qualidade de CD
 - Principal: Digitalização 4:2:0 com resolução de até 720×576 pixels. Vídeo com qualidade de studio e áudio com multiplas faixas com qualidade de CD.
 - Alta 1440: Digitalização 4:2:0 com resolução de 1440×1152 pixels. HDTV com até 60Mbps (ou 80Mbps com formato 4:2:2).
 - Alta: Digitalização 4:2:0 com resolução de 1920×1152 pixels. *Wide-screen* HDTV com taxa de bits de 80Mbos (ou 100Mbps com formato 4:2:2).
 - MPEG-4
 - Aplicações Multimídia Interativas. Permite que um usuário acesse componentes individuais da sequência de vídeo.
 - MPEG-7 (indexação), MPEG-21, ...

MPEG-1

- Utiliza um mecanismo de compressão similar àquele de H.261.
- Utiliza o formato SIF como base para os formatos de compressão, porém, como utiliza-se macroblocos de 16×16 pixels, tal resolução tem de ser diminuída de 360 to 352 pixels para produzir um número inteiro de macroblocos (22):
 - NTSC: $Y=352 \times 240$ pixels, $C_r=C_b=176 \times 120$ pixels
 - PAL: $Y=352 \times 288$ pixels, $C_r=C_b=176 \times 144$ pixels
- Varredura progressiva com 30fps (NTSC) ou 25fps (PAL).
- O padrão permite o uso de:
 - Somente quadros I
 - Somente quadros I e P.
 - Quadros I, P e B.
- Não suporta quadros D. Quadros I são utilizados para FF e REW.
- Não deve haver mais do que 0,5 segundos entre quadros I.

MPEG-1



- Exemplos de sequências são:
 - IBBPBBPBBI - PAL (25fps)
 - IBBPBBPBBPBBI - NTSC (30fps)
- Quadros são reordenados para permitir decodificação de quadros de referência antes dos quadros que daqueles dependem.

Exercício 1

An MPEG-1 system uses the frame sequence shown in Figure 4.20.

- (i) Define the terms M and N and hence determine their values for the sequence shown in the figure.
- (ii) Derive a suitable reordered sequence that ensures firstly, only two frames must be stored in the decoder, and secondly, the required I- and/or P-frames are available to decode each P- and B-frame as they are received.

Answer:

- (i) As we described earlier in Section 4.3.1 under the subheading of “Frame types”, M is the distance (in frames) between a P-frame and the immediately preceding I- or P- frame, and N is the number of frames between two successive I-frames. The latter is known as a group of pictures or GOP. Hence for the frame sequence shown in Figure 4.20, $M = 3$ and $N = 12$.
- (ii) A suitable reordered frame sequence that meets the defined requirements is:
IPBBPBBPBBIBBPBB ...

Exercício 2

A digitized video is to be compressed using the MPEG-1 standard. Assuming a frame sequence of:

IBBPBBPBBPBL...

and average compression ratios of 10:1 (I), 20:1 (P) and 50:1 (B), derive the average bit rate that is generated by the encoder for both the NTSC and PAL digitization formats.

Answer:

Frame sequence = IBBPBBPBBPBL...

Hence: 1/12 of frames are I-frames, 3/12 are P-frames, and 8/12 are B-frames.

and Average compression ratio = $(1 \times 0.1 + 3 \times 0.05 + 8 \times 0.02) / 12$
= 0.0342 or 29.24:1

NTSC frame size:

Without compression = $352 \times 240 \times 8 + 2 (176 \times 120 \times 8)$
= 1.013760 Mbits per frame

With compression = $1.01376 \times 1/29.24$
= 34.670 kbits per frame

Hence bit rate generated at 30fps = 1.040 Mbps

PAL frame size:

Without compression = $352 \times 288 \times 8 + 2 (176 \times 144 \times 8)$
= 1.216512 Mbits per frame

With compression = $1.216512 \times 1/29.24$
= 41.604 kbits per frame

Hence bit rate generated at 25fps = 1.040 Mbps

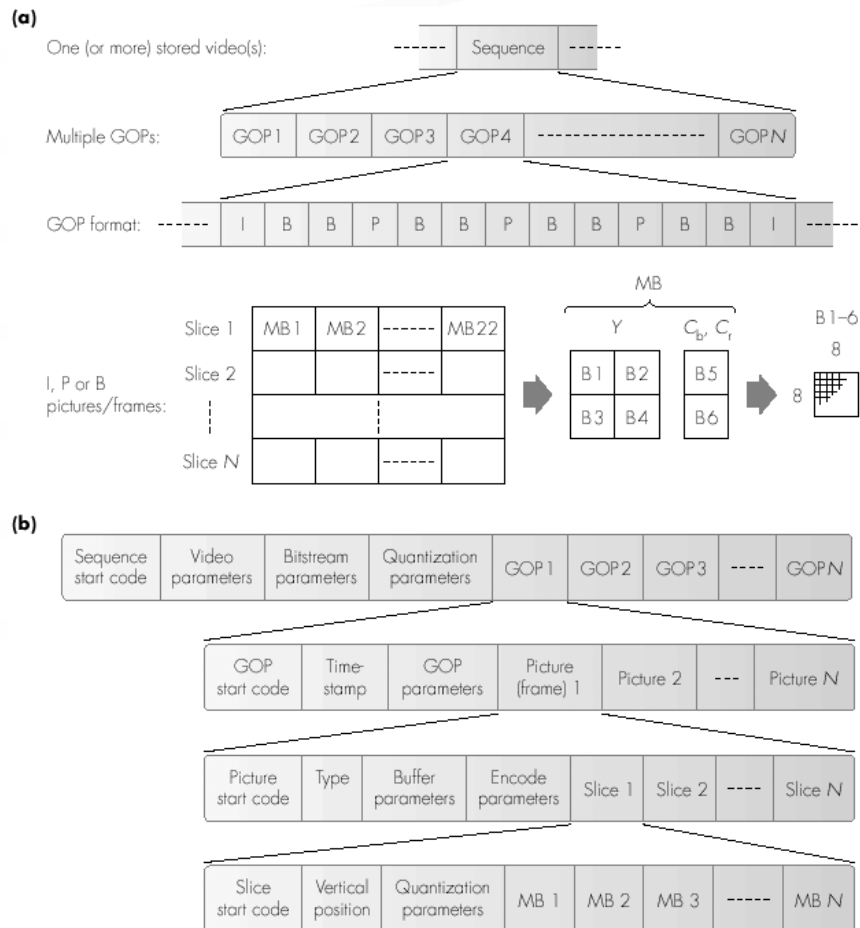
Normally, allowing for packetization and multiplexing overheads, a bandwidth of 1.2 Mbps is allocated for the video. Hence, assuming a maximum bit rate of 1.5 Mbps, this leaves 300 kbps for the compressed audio stream.

MPEG-1

- AS principais diferenças para H.261 são:
 - *Time-Stamp*s (referências de tempo) podem ser incluídas dentro de um quadro, o que permite o decodificador se re-sincronizar mais rapidamente no caso de erros em macroblocos. O número de macroblocos entre referências de tempo é chamada fatia (*slice*). Uma fatia pode ter de 1 ao máximo número de macroblocos no quadro. O valor mais comumente utilizado, entretanto, é 22, que é o número de macroblocos em uma linha
 - A introdução de quadros B, o que aumenta o intervalo entre quadros I e P (o que é compensado pelo aumento no tamanho da janela de busca).
- Taxas de compressão típicas são da ordem de:
 - 10:1 para quadros I
 - 20:1 para quadros P
 - 50:1 para quadros B

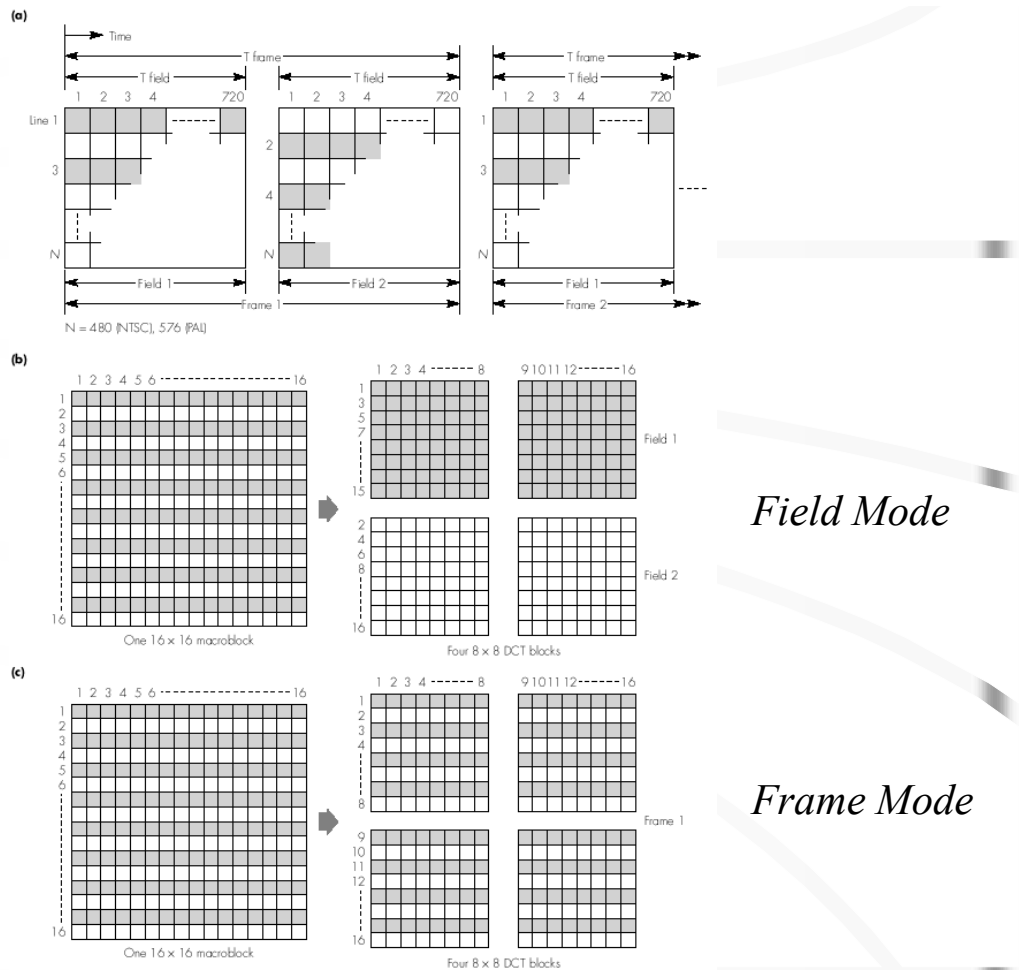
MPEG-1

- Uma vez comprimido os dados são organizados num esquema hierárquico. O vídeo comprimido completo é chamado “sequência” e consiste de:
 - Cada sequência tem um marcador de início (*sequence start code*) seguido de um número de parâmetros que se aplicam à fatia:
 - Parâmetros de vídeo (dimensões, formato)
 - Parâmetros de *stream* (taxa de bits, quantidade de memória necessária para decodificar a sequência).
 - Parâmetros de quantização (tabelas de quantização)
 - Composta de GOPs.
 - Grupos de Imagens (GOP)
 - Conjunto de imagens I, P e B.
 - Cada Figura é composta de fatias, que são compostas de macroblocos, que são compostos por blocos que são compostos por pixels.
 - Possui um *GOP Start Code*, *time-stamp*, etc.



MPEG-2

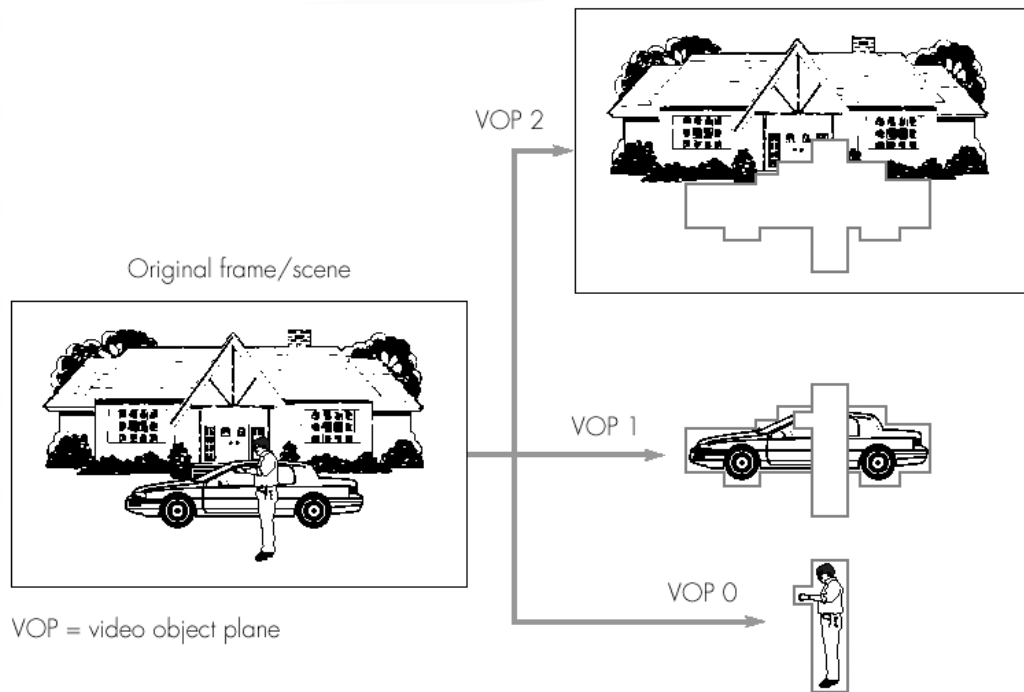
- MPEG-2 suporta 4 níveis (baixo, principal, alto 1440 e alto), cada um direcionado a um tipo de aplicação. Cada nível possui, ainda, 5 perfis (simples, principal, resolução espacial, precisão de quantização e alto), que montam uma tabela, juntamente com os níveis, que é usada como referência para toda e qualquer aplicação que venha a utilizar o MPEG-2.
- MP@ML
 - É uma das combinações da tabela. É um padrão indicado para broadcasting de televisão digital, com ou 30fps (NTSC) ou 25fps (PAL), digitalização 4:2:0 com resolução de 720×480 (NTSC) ou 720×576 (PAL).
 - Taxa de transmissão entre 4 e 15Mbps
 - Codificação similar ao MPEG-1, exceto pelo uso de varredura interlacada ao invés de progressiva.



MPEG-4

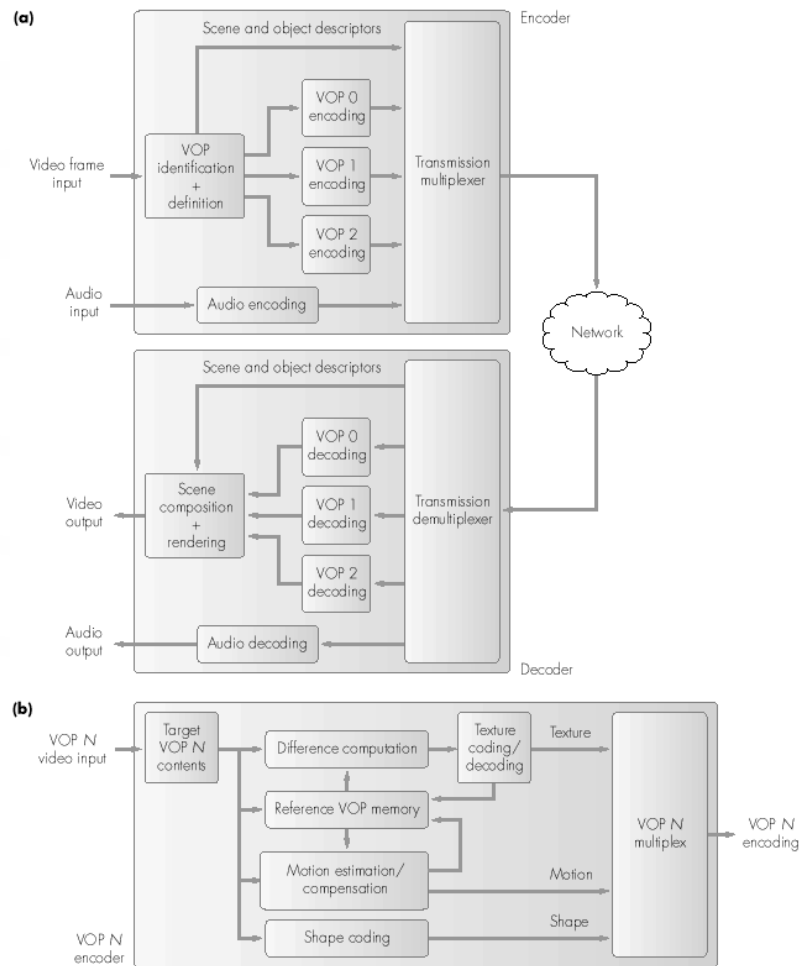
Composição de Cena

- A principal diferença entre MPEG-4 e os outros padrões de compressão de vídeo é o fato de que MPEG-4 contém recursos chamados *content-based functionalities*.
 - Antes de ser codificada cada cena é decomposta em um background e um número de objetos audiovisuais (AVOs). Cada AVO é então definida como um conjunto de um ou mais objetos de vídeo e/ou objetos de áudio.
 - Cada objeto de áudio ou vídeo tem um descritor de objetos (*object descriptor*) associado a ele, o que permite a um usuário manipular o objeto antes do mesmo ser decodificado e representado.
 - A linguagem usada para descrever e modificar objetos é chamada BIFS (*binary format for scenes*), o que inclui comandos para apagar objetos, modificar o formato, aparência e animar o objeto em tempo real (se apropriado).



MPEG-4

- Uma cena é composta, num nível superior, por um descritor de cenas (*scene descriptor*), o qual define como os vários AVOs se relacionam uns com os outros.
- Um quadro é particionado em planos de objetos de vídeo (*video object planes* - VOPs). Cada um destes planos corresponde a um AVO de interesse.
- Cada VOP é encapsulada em um retângulo, que é composto do número mínimo de macroblocos que completamente envolvem o objeto.
- AS coordenadas do VOP são relativas às coordenadas do macrobloco do topo superior esquerdo do frame.
- Cada VOP é codificado independentemente usando um procedimento similar àqueles anteriormente descritos.
- Os vários VOPs codificados são então multiplexados num *bitstream* que também contém o descritor da cena.



Decodificador MPEG-4

