



SCC0661 – MULTIMÍDIA E HIPERMÍDIA

AULA 06: VÍDEO DIGITAL

Prof. Dr. Marcelo G. Manzato
(mmanzato@icmc.usp.br)

Instituto de Ciências Matemáticas e de Computação – Sala 3-160

SUMÁRIO

- 1 – Princípios de Compressão de Vídeo
- 2 – Padrões para Compressão de Vídeo

1. PRINCÍPIOS DE COMPRESSÃO DE VÍDEO

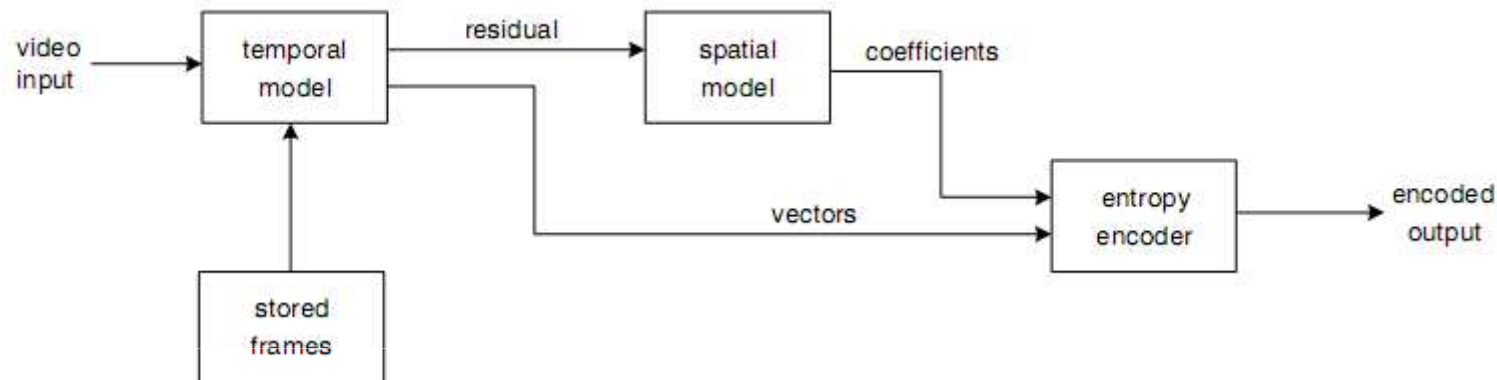
- 1.1 - Tipos de Quadros
- 1.2 - Estimativa e Compensação de Movimento

1. PRINCÍPIOS DE COMPRESSÃO DE VÍDEO

- Compressão : eliminação de redundâncias
 - Estatística (JPEG-LS)
 - Espacial (MJPEG)
 - Temporal (MPEG)



1. PRINCÍPIOS DE COMPRESSÃO DE VÍDEO



○ Modelo

- Representação dos dados codificados que pode ser usada para reconstruir o vídeo original
- Idealmente deveria utilizar poucos bits e recompor os dados com alta fidelidade

1. PRINCÍPIOS DE COMPRESSÃO DE VÍDEO

- Redundância entre quadros adjacentes
 - Redundância temporal.
- Técnicas para remoção de redundância temporal
 - Prever (“predizer”) o conteúdo de quadros sucessivos.
 - Apenas as diferenças são codificadas.
- Acuidade da predição
 - Quão bem o movimento é estimado.
 - Operação é chamada de **Estimativa de Movimento**.
 - Predição não é perfeita.
 - **Compensação de Movimento.**

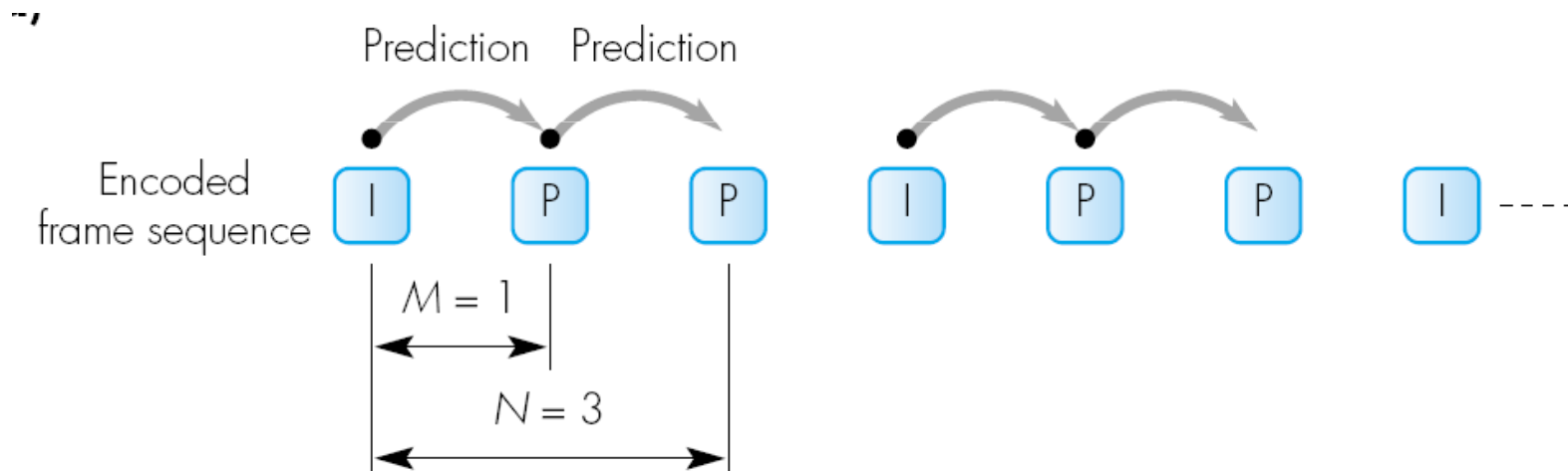
1.1 TIPOS DE QUADROS

- Dois tipos básicos de quadros:
 - Quadros codificados independentemente e quadros *predicted*.
 - *Intracoded frames* ou *I-frames* ou quadros I.
 - *Predicted frames*:
 - *Predictive* ou *P-frames* ou quadros P.
 - *Bidirectional* ou *B-frames* ou quadros B.
- Tipos de quadros especiais:
 - PB-frame
 - Agrupamento de quadros P e B num único fluxo
 - D-frame
 - Usado para operações de fast-forward e rewind

1.1 TIPOS DE QUADROS

- Quadros I.
 - São codificados sem nenhuma referência a outros quadros.
 - Cada quadro é tratado como uma imagem independente sendo Y, Cb e Cr codificados usando o algoritmo JPEG.
 - Aparecem no fluxo de saída em intervalos regulares.
 - $N = \text{GOP (group of pictures) span}$: número de quadros (3 a 12) entre dois quadros I sucessivos.

1.1 TIPOS DE QUADROS

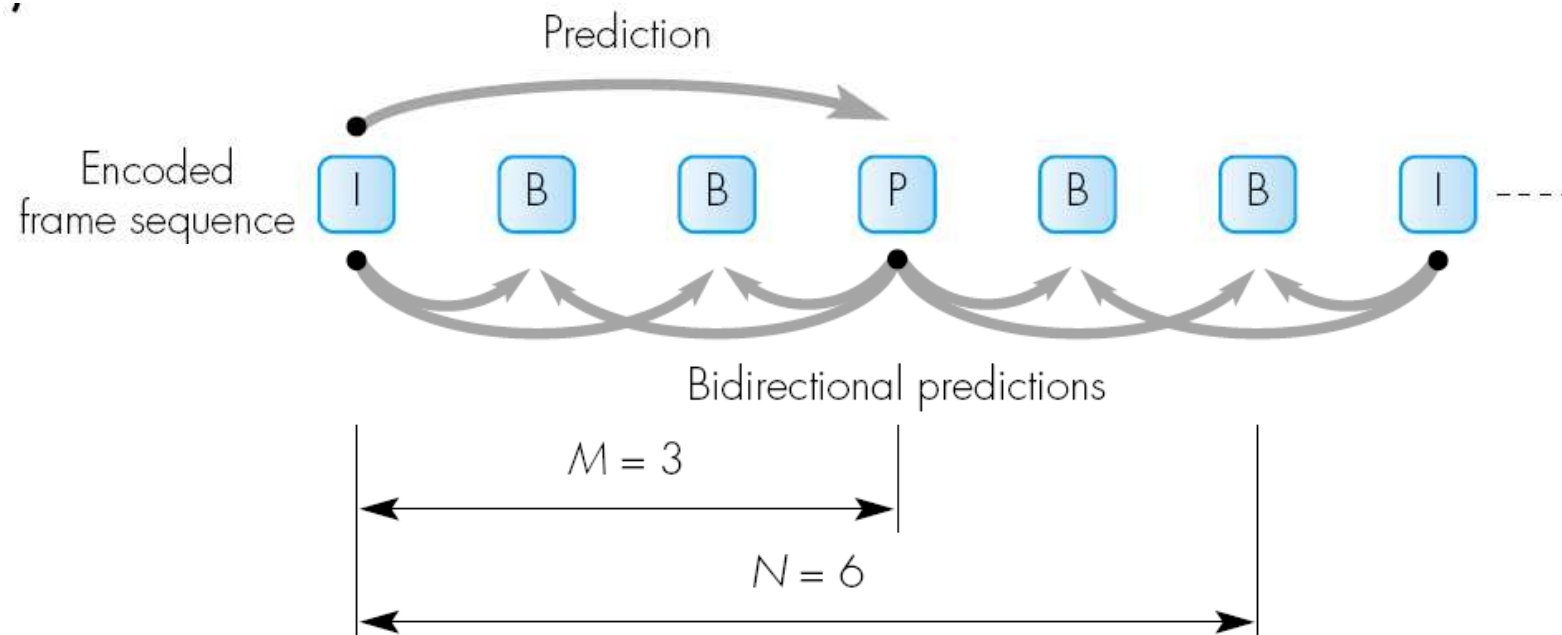


1.1 TIPOS DE QUADROS

○ Quadros P

- São codificados em relação ao conteúdo de um quadro I ou de um quadro P anterior.
- Usam combinação de estimativa e compensação de movimento - alcançam maiores taxas de compressão do que quadros I.
- **Propagam erros** - número de quadros P entre quadros I é limitado.
- $M = \textit{prediction span}$ - número de quadros entre um quadro P e o quadro I ou P imediatamente anterior.
- Desempenho: taxa de compressão entre 20:1 e 30:1.

1.1 TIPOS DE QUADROS



M = prediction span N = group of pictures (GOP) span

1.1 TIPOS DE QUADROS

○ Quadros B

- São codificados em relação ao conteúdo de um quadro I ou de um quadro P anterior e de um posterior.
- Envolve o processamento de 3 quadros: o quadro I ou P anterior, o quadro atual e o quadro I ou P posterior. Todos não codificados.
- Aumento no tempo (*delay*) para codificação e decodificação. É o tempo de esperar o próximo quadro I ou P.
- Provêem alta taxa de compressão: entre 30:1 e 50:1.
- Não propagam erros.

1.1 TIPOS DE QUADROS

○ Quadros B

- Decodificação:

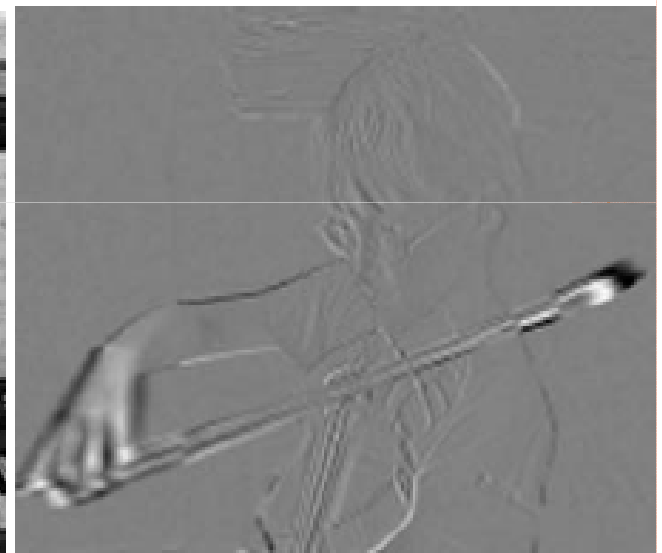
- $I_1 B_2 B_3 P_4 B_5 B_6 P_7 B_8 B_9 I_{10} \dots$

- Codificação:

- $I_1 P_4 B_2 B_3 P_7 B_5 B_6 I_{10} B_8 B_9 \dots$



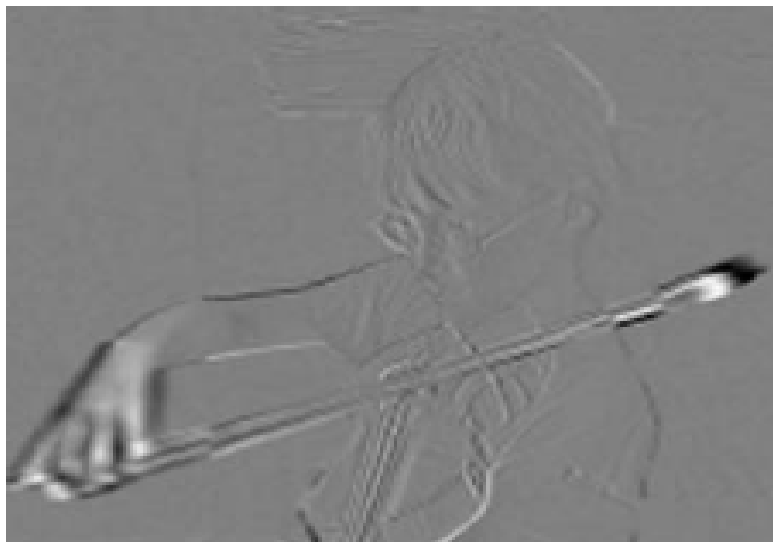
1.2 ESTIMATIVA E COMPENSAÇÃO DE MOVIMENTO



1.2 ESTIMATIVA E COMPENSAÇÃO DE MOVIMENTO

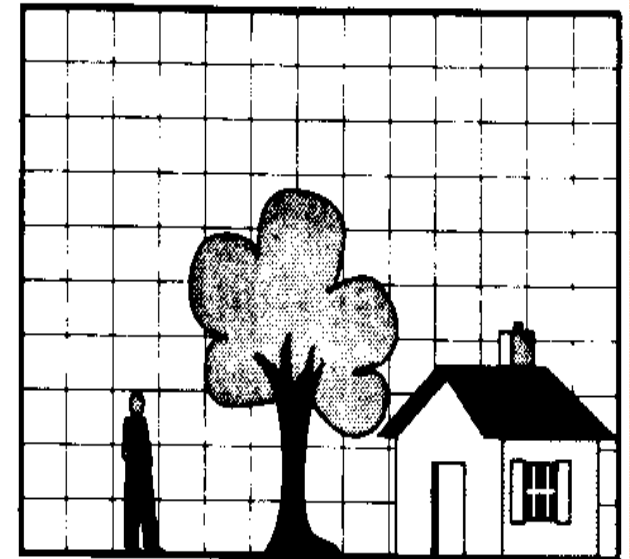
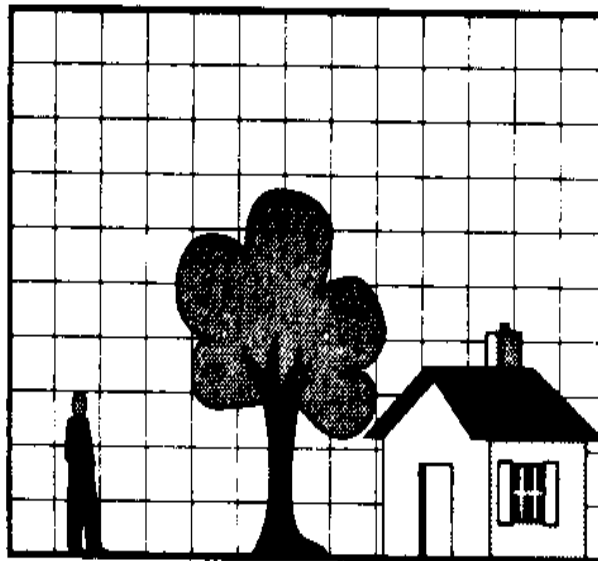
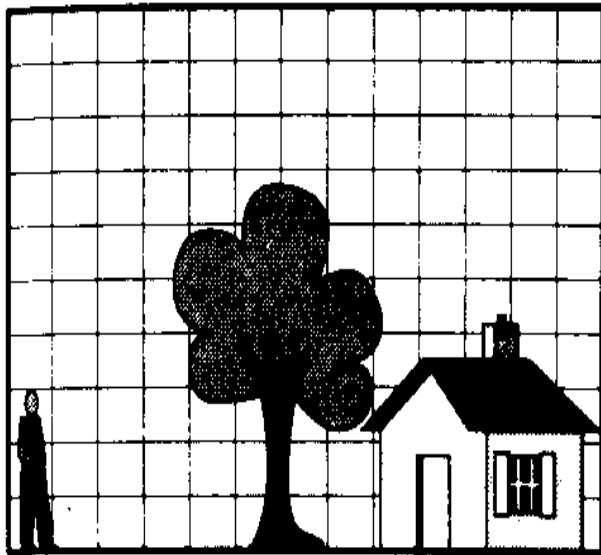
○ Fluxo óptico

- Previsão correta da maioria dos pixels do quadro atual por meio da movimentação dos pixels do quadro de referência
- Cálculo do fluxo óptico é computacionalmente caro
- Necessidade de transmitir todo vetor de fluxo óptico para o decodificador



1.2 ESTIMATIVA E COMPENSAÇÃO DE MOVIMENTO

- Estimativa e Compensação baseada em Bloco
 - Consiste em achar regiões na imagem que podem ser encontradas nas imagens seguintes.

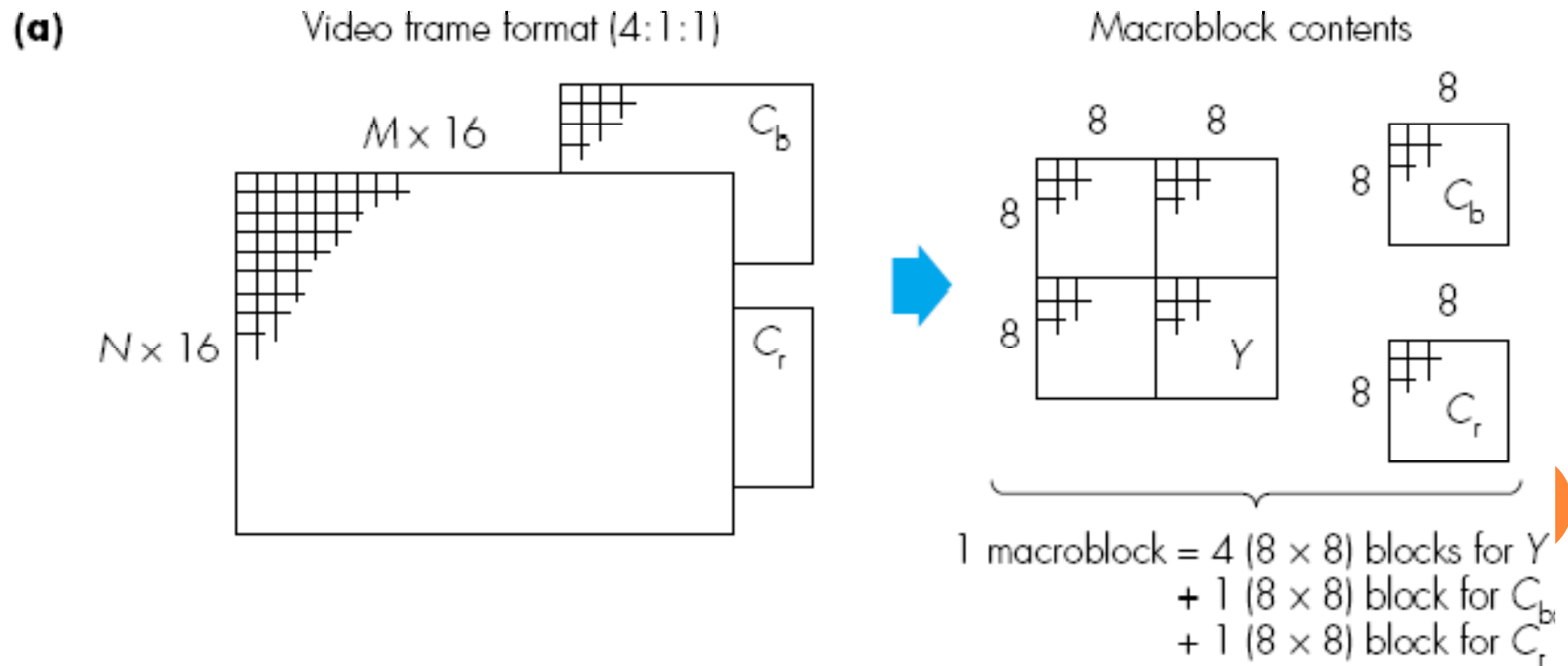


1.2 ESTIMATIVA E COMPENSAÇÃO DE MOVIMENTO

- Estimativa e Compensação baseada em Bloco
 - Buscar na imagem de referência uma região que melhor se assemelha à região do quadro atual (estimativa de movimento)
 - A região candidata torna-se o previsor do bloco do quadro atual, sendo subtraída do bloco atual para formar um resíduo (compensação de movimento)
 - O resíduo é codificado juntamente com o offset entre o bloco atual e a posição da região candidata (vetor de movimento)

1.2 ESTIMATIVA E COMPENSAÇÃO DE MOVIMENTO

- A imagem é dividida em macroblocos.
 - Y, Cr e Cb = matrizes de 16x16, 8x8 e 8x8 pixels (formato 4:1:1).

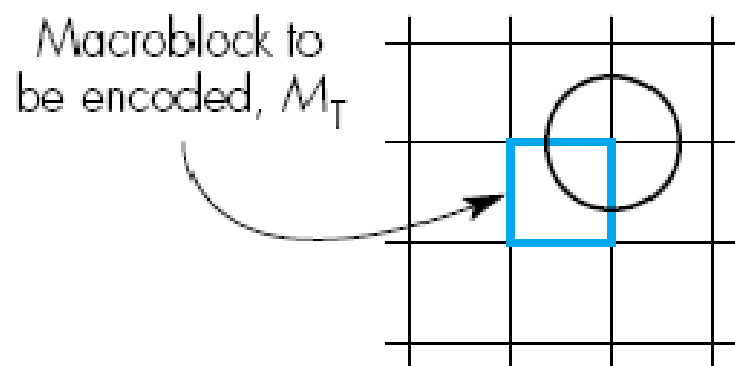


1.2 ESTIMATIVA E COMPENSAÇÃO DE MOVIMENTO

- Para codificar **quadros P**, cada macrobloco do quadro-alvo é comparado, pixel a pixel, com o macrobloco correspondente do quadro-referência (I ou P anterior).
 - Se o conteúdo combina (match), apenas o endereço do macrobloco é codificado. Senão, estende-se a busca para macroblocos vizinhos.
 - Normalmente utiliza-se apenas a componente Y

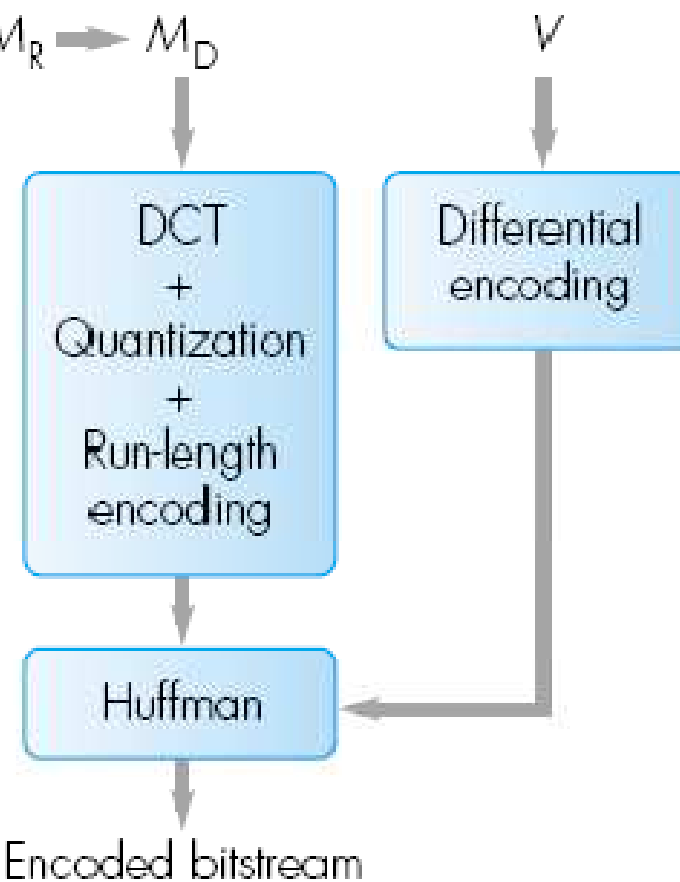
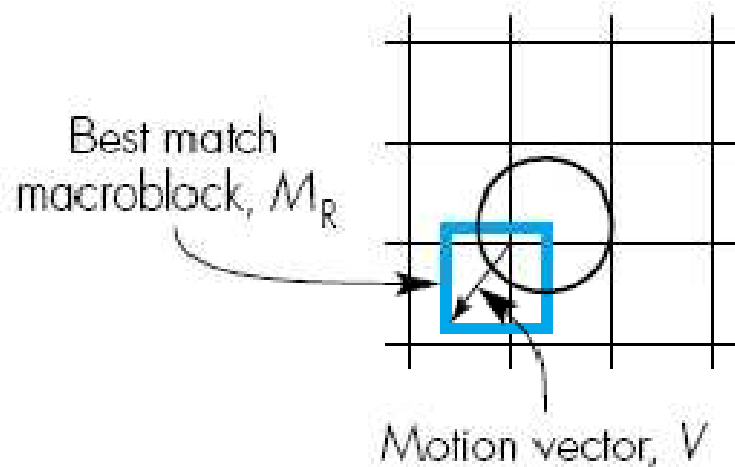
(b)

Search region in target frame:



$$M_T - M_R \rightarrow M_D$$

Same search region in preceding (I or P) reference frame:



1.2 ESTIMATIVA E COMPENSAÇÃO DE MOVIMENTO

- Quando a busca em macroblocos vizinhos encontra um macrobloco-par, dois parâmetros são codificados:
 - Um vetor de movimento (*motion vector*).
 - Indica o deslocamento (*offset*) do macrobloco.
 - Erro de predição: três matrizes – uma para cada componente (Y, Cr e Cb) – contendo as diferenças de valores entre os pixels do macrobloco-alvo e os pixels da área de busca.
 - É necessário pois a estimativa de movimento não é um método exato.

1.2 ESTIMATIVA E COMPENSAÇÃO DE MOVIMENTO

- Codificação dos parâmetros:
 - Vetores de movimento
 - Codificados usando codificação por diferenças
 - Resultado é acoplado na codificação Huffman
 - Erro de predição
 - Codificado como um quadro I
 - Mas as matrizes contém apenas as diferenças entre os macroblocos do quadro alvo e do quadro referência.
- Se um “casamento” não é encontrado:
 - Macrobloco é codificado independentemente.
 - Codificação segue os passos de um quadro I: DCT, quantização e codificação por entropia.

1.2 ESTIMATIVA E COMPENSAÇÃO DE MOVIMENTO

○ Quadros B:

- Estima-se, primeiro, o vetor de movimento e as matrizes de diferenças usando-se o quadro P ou I anterior.
- Depois, estima-se os mesmos parâmetros usando-se o quadro P ou I posterior.
- Calcula-se um terceiro conjunto de parâmetros usando o macrobloco-alvo e a média dos valores previstos nos dois passos anteriores.
- O conjunto com os menores valores é escolhido para ser codificado – como em um quadro P.

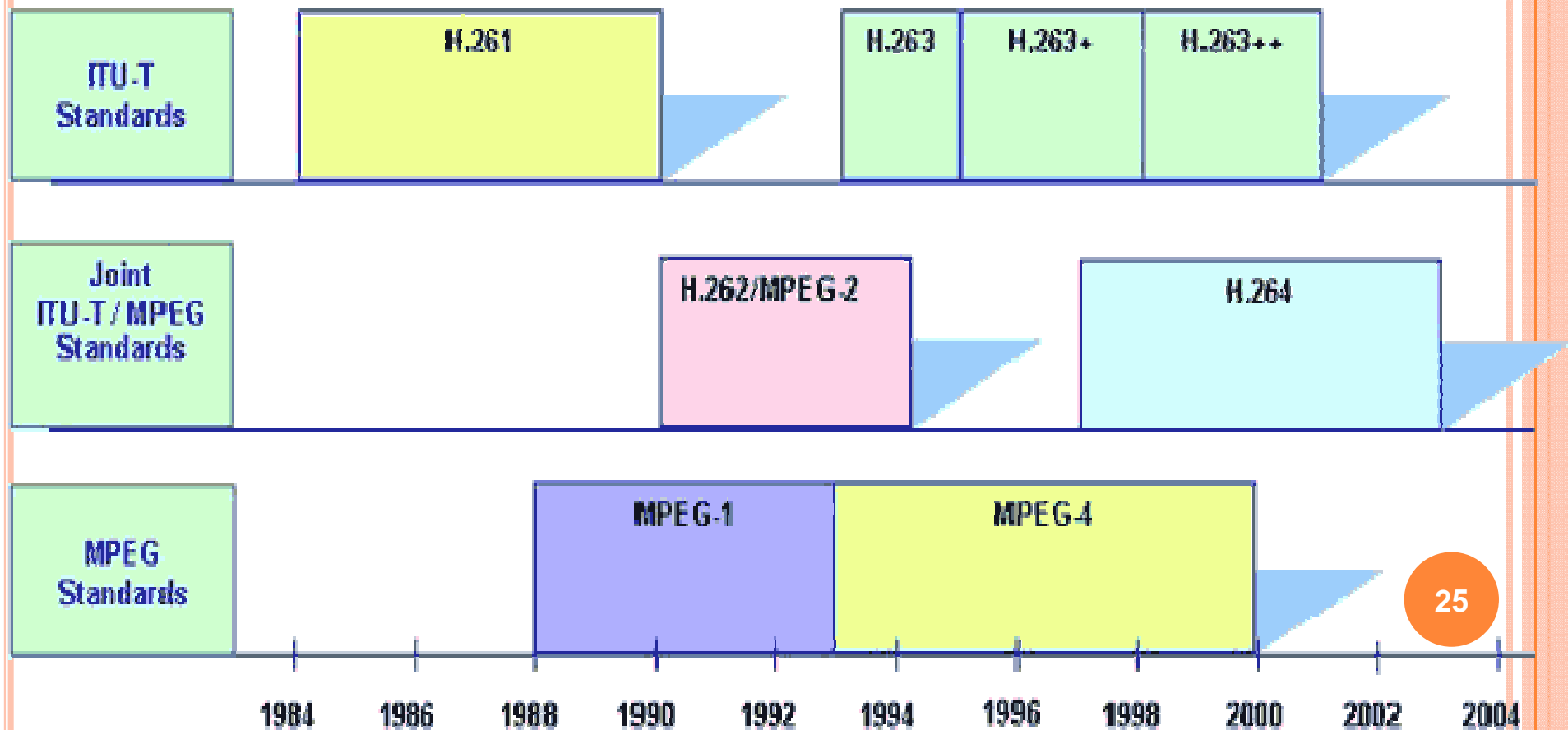
2. PADRÕES PARA COMPRESSÃO DE VÍDEO

- 2.1 – H.261
- 2.2 – H.263
- 2.3 – MPEG-1
- 2.4 – MPEG-2
- 2.5 – MPEG-4
- 2.6 – H.264



2. PADRÕES PARA COMPRESSÃO DE VÍDEO

- Evolução dos padrões:



2.1 H.261

- Padrão de compressão de vídeo definido pela ITU-T em 1990.
- Projetado para:
 - Videoconferência.
 - Aplicações de vídeo-telefone em linhas ISDN (*Integrated Services Digital Network*).
- Codificador de vídeo para transmissão a px64 Kbit/s.
 - p varia de 1 a 30 (64 a 1920Kbps).
- Padrão também é conhecido como px64.

2.1 H.261

- Define dois formatos de imagem:
 - CIF: $Y = 352 \times 288$, $Cb = Cr = 176 \times 144$.
 - Varredura progressiva, 30 fps.
 - Videoconferência.
 - QCIF: $Y = 176 \times 144$, $Cb = Cr = 88 \times 72$.
 - Varredura progressiva, 15 ou 7.5 fps.
 - Vídeotelefonia.
- Macroblocos de 16×16 pixels.
- Dois tipos de quadros: I e P.
- GOP *Span*: $N = 4$.
 - Três quadros P entre dois quadros I.

2.1 H.261

- Algoritmo de compressão baseado em DCT.
- Otimiza a utilização de largura de banda estabelecendo um compromisso entre qualidade contra movimento.
 - Imagens com rápidas mudanças têm pior qualidade que imagens quase estáticas.

2.2 H.263

- *Video Coding for Low Bit Rate Communication*
- Padrão de compressão de vídeo definido pela ITU-T em 1995.
- Projetado para:
 - Videoconferência.
 - Vídeo sobre redes sem fio e PSTN (*Public Switched Telephone Networks*).
- PSTN (analógico) requer modems para envio de dados digitais, o que limita a taxa de transmissão para algo em torno de 28Kbps-56Kbps.
 - Requisito do codificador: vídeo a taxas de bits muito baixas

2.2 H.263

- Baseado no padrão H.261:
 - Algoritmo de codificação semelhante ao do H.261
- Qualidade de imagem superior.
 - H.261 usa só quadros I e P, e para manter baixa a taxa de bits, precisa usar limiares de quantização elevados → efeitos de blocagem
- H.263 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.

2.2 H.263

- Formatos e Tipos de Quadro
 - Mandatórios:
 - SQCIF: $Y=128 \times 96$, $Cb = Cr = 64 \times 68$.
 - QCIF: $Y=176 \times 144$, $Cb = Cr = 88 \times 72$.
 - Opcionais:
 - CIF: $Y=352 \times 288$, $Cb = Cr = 176 \times 144$.
 - 4CIF: $Y=704 \times 576$, $Cb = Cr = 356 \times 288$.
 - 16CIF: $Y=1408 \times 1152$, $Cb = Cr = 704 \times 576$.
- H.263 possui quadros dos tipos I, P, B e PB.

2.3 MPEG-1

- Recomendação ISO 11172
- Características gerais:
 - Domínio de aplicação:
 - Armazenamento de áudio e vídeo, com qualidade VHS, a taxas de 1,5 Mbps.
 - Resolução de vídeo baseada no formato SIF (*Source Intermediate Format*).
 - NTSC: $Y=352 \times 240$, Cr e Cb = 176×120 .
 - PAL: $Y=352 \times 288$, Cr e Cb = 176×144 .
 - Varredura progressiva.
 - 30Hz p/ NTSC e 25 Hz p/ PAL.

2.3 MPEG-1

- Utiliza compressão:
 - Intra-quadros: redundância espacial
 - Mesma técnica do JPEG.
 - Inter-quadros: redundância temporal
 - Estimativa e compensação de movimento.
- Tipos de quadros:
 - MPEG-1 usa três tipos de quadros I, P e B.
- Não utiliza quadros D.
 - Acesso aleatório (*VCR-like*) – usa-se quadros I.
 - Acesso aleatório no máximo em 0.5 segundo.
 - Separação máxima entre quadros I
 - IBBPBBPBBI... (PAL). IBBPBBPBBI... (NTSC).

EXERCÍCIO

- Um vídeo digitalizado precisa ser comprimido usando o padrão MPEG-1. Assumindo uma sequência de quadros de:

IBBPBBPBBPBBI...

e razões de compressão de 10:1 para quadros I, 20:1 para quadros P e 50:1 para quadros B, deriva a taxa média de bits que será gerada pelo codificador para ambos os formatos NTSC e PAL.

2.4 MPEG-2

- Recomendação ISO 13818
- Definido em uma série de documentos, subconjuntos da recomendação 13818.
- Domínio de aplicação:
 - Armazenamento e transmissão de áudio e vídeo com qualidade de estúdio.
- Utiliza compressão intra-quadros e inter-quadros.
- Possui 4 **Níveis** de resolução de vídeo:
 - Low, Main, High 1440 e High.
- Cada nível possui cinco **Perfis**: simple, main, spatial resolution, quantization accuracy e high.
 - Tradeoff entre necessidades da aplicação e características do perfil@nível.
- Níveis superiores são compatíveis com níveis inferiores.

2.4 MPEG-2

○ Nível Low:

- Resolução: 352 x 288 (SIF).
- Taxa de quadros: 23.976, 24, 25, 29.97 ou 30
- Compatível com MPEG-1.
- Taxa de bits: 4Mbps.

○ Nível Main:

- Resolução: 720 x 576.
- Taxa de quadros: 23.976, 24, 25, 29.97 ou 30
- Áudio e vídeo com qualidade de estúdio.
- Múltiplos canais de áudio.
- Taxa de bits: 15Mbps.

2.4 MPEG-2

○ Nível High 1440:

- Resolução: 1440 x 1152.
- Taxa de quadros: 23.976, 24, 25, 29.97, 30, 50, 59.94 ou 60
- HDTV.
- Taxa de bits: 60Mbps.

○ Nível High:

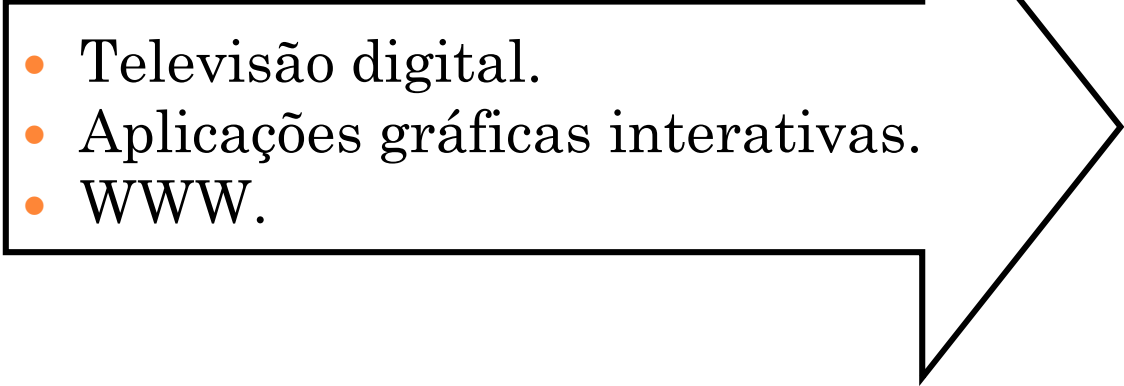
- Resolução: 1920 x 1152.
- Taxa de quadros: 23.976, 24, 25, 29.97, 30, 50, 59.94 ou 60
- Wide-screen HDTV.
- Taxa de bits: 80Mbps.

2.4 MPEG-2

- Main Profile at the Main Level (MP@ML)
 - Broadcasting de TV Digital.
 - Formato de digitalização 4:2:0 a 30Hz para NTSC ou 25Hz para PAL.
 - Resolução de 720x480 @ 30Hz ou 720x576 @ 25Hz
 - Taxa de bits entre 4 e 15 Mbps.
 - Esquema de codificação similar ao MPEG-1.
 - Principal diferença: uso de varredura entrelaçada.
 - Codificação poder ser progressiva ou entrelaçada.

2.5 MPEG-4

- Visa atender a três áreas:

- 
- Televisão digital.
 - Aplicações gráficas interativas.
 - WWW.

Convergência

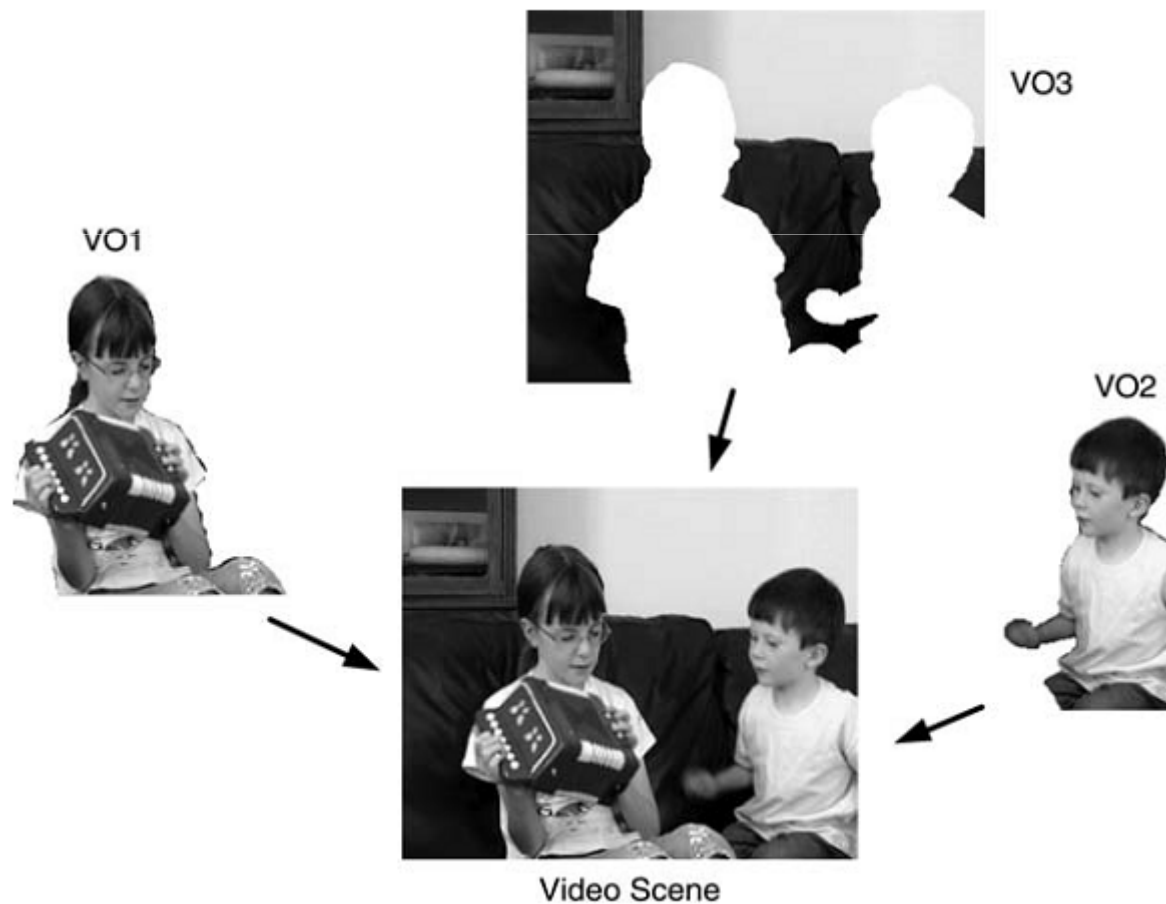
- Visa fornecer padrões para integrar a produção, distribuição e acesso ao conteúdo audiovisual.
- Principal diferença em relação aos outros padrões MPEG: conceito de Objetos.

2.5 MPEG-4

- Funcionalidades baseadas em conteúdo
 - Antes da compressão, a cena é definida em termos de um plano de fundo e um ou mais objetos audiovisuais (AVOs)
 - Cada AVO é definido na forma de um ou mais objetos de vídeo ou de áudio
 - Descrição e manipulação de cada objeto por meio de descritores de objetos (object descriptor)
 - Binary format for scenes (BIFS)
 - Descrição da cena (scene descriptor) em termos de diferentes AVOs
 - Multiplexação dos dados para armazenamento ou transmissão

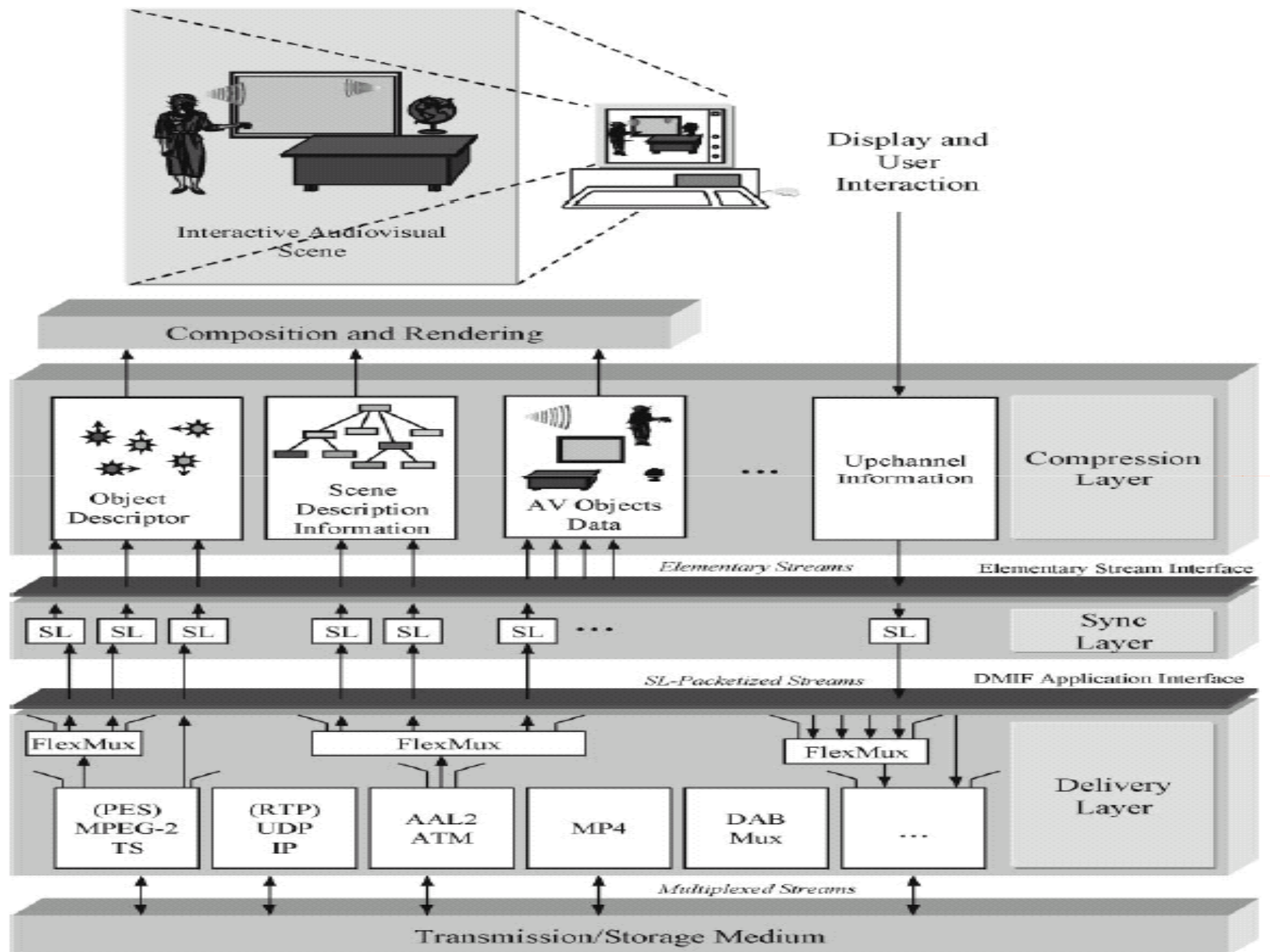
MPEG-4

- Video Object Plane (VOP)



2.5 MPEG-4

- Prevê um conjunto de tecnologias para atender a:
 - **Autores**: possibilita a criação de conteúdo re-usável e flexível (HDTV, animações, WWW), e proteção de direitos autorais.
 - **Provedores de serviços de rede**: descritores genéricos de QoS para cada tipo de mídia MPEG-4.
 - **Usuários finais**: altos níveis de interação com o conteúdo.
- Objetivos são alcançados através da padronização de:
 - Composição de objetos, criação de objetos complexos.
 - Codificação de objetos de mídia (naturais ou sintéticos).
 - Multiplexação e sincronização.
 - Interação com a cena audiovisual.
 - Compressão.



2.5 MPEG-4

- Base: compressão intra e inter-quadros. Utilização de quadros I, P e B.
- Altas taxas de compressão.
 - DivX, AVI, .mp4...
- Novas ferramentas em compensação de movimento.
 - Mais vetores de movimento por macrobloco
 - Vetores de movimento irrestritos
- Codificação de formas arbitrárias.
- Escalabilidade (vídeo em múltiplas camadas):
 - Espacial.
 - Temporal.
- Amplo *range* de taxas de dados: 5Kbps-1Gbps
- Formatos: SQCIF – HDTV.
 - Progressivo e entrelaçado.

2.5 MPEG-4

○ Escalabilidade Espacial



Figure 5.58 Original video frame



Figure 5.59 Sub-sampled frame to be encoded as base layer



Figure 5.60 Base layer frame (decoded and upsampled)



Figure 5.61 Residual to be encoded as enhancement layer

2.5 MPEG-4

- Parte 1: Systems (ISO/IEC 14496-1)
- Parte 2: Visual (ISO/IEC 14496-2)
- Parte 3: Audio (ISO/IEC 14496-3)
- Parte 4: Conformance Testing (ISO/IEC 14496-4)
- Parte 5: Reference Software (ISO/IEC 14496-5)
- Parte 6: Delivery Multimedia Integration Framework (DMIF) (ISO/IEC 14496-6)
- Parte 7: Optimized Visual Reference Software (ISO/IEC 14496-7-TR)
- Parte 8: MPEG-4 over IP (ISO/IEC 14496-8:2002)
- Parte 9: Reference Hardware Description (ISO/IEC 14496-9:2003-TR)
- Parte 10: Advanced Video Coding (ISO/IEC 14496-10:2003)

2.6 H.264

- Advanced video coding (AVC) for generic audiovisual services
 - Padrão de compressão de vídeo definido pela JVT em 2002.
 - H.264/AVC/H26L
 - Mais avançado padrão de compressão disponível atualmente.
 - Usa técnicas de compressão não disponíveis no MPEG-2, MPEG-4 e H.263

2.6 H.264

- Codificador extremamente escalonável.
- Aplicação:
 - Broadcast;
 - DVD;
 - Videoconferência;
 - Vídeo em demanda;
 - Transmissão e mensagens multimídia.

2.6 H.264

Cenário de Uso	Resolução e Taxa de Frame	Exemplo de Taxa de Dados
Conteúdo Móvel	176x144, 10-24 fps	50-60 Kbps
Internet/Definição Padrão	640x480, 24 fps	1-2 Mbps
Alta Definição	1280x720, 24p	5-6 Mbps
Alta Definição em Tela Cheia	1920x1080, 24p	7-8 Mbps

OUTRAS MÍDIAS DE REPRESENTAÇÃO

- Mais comuns:
 - AVI (microsoft)
 - MOV (apple)
 - MKV (padrão aberto)
- Servem como “pacotes” de dados.
- Usam codificadores diversos para comprimir vídeo.
 - Indeo, Cinepack, DivX, ADPCM, H.264...

PARA SABER MAIS

- Luther, A. C. Using Digital Video. AP Professional, 1995. (capítulo 2 e apêndice A).
- Richardson, L. E. G. H.264 and MPEG-4 Video Compression, Wiley, 2003. (capítulos 2 e 3).
- Halsall, F. Multimedia Communications: Applications, Networks, Protocols, and Standards, Addison-Wesley Publishing, 2001. ISBN: 0201398184. Capítulo 3, seção 4.3.
- H.261 e H.263:
 - http://www.compression-links.info/H.261_H.263
- Padrões MPEG:
 - <http://www.chiariglione.org/mpeg/>