

## Subsistema de Compressão e Codificação do Sinal de Vídeo dos Padrões HDTV (Parte II)

Luiz Rômulo Mendes, Yuzo Iano, Vicente Idalberto Becerra Sablón, Renê Togni Del Pietro

Universidade Estadual de Campinas - Faculdade de Engenharia Elétrica e de Computação  
Departamento de Comunicações

**Resumo:** Os dois principais padrões de HDTV já estabelecidos, um americano conhecido como Padrão da Grande Aliança (GA) [1][2][3][4], e outro europeu, Padrão DVB (“Digital Video Broadcasting”) [5][6][7], apresentam características específicas que os tornam diferentes. Por exemplo, o Padrão da GA e o DVB-T utilizam modulação 8-VSB (“Vestigial Side Band”) e modulação OFDM (“Orthogonal Frequency Division Multiplexing”) [8], respectivamente. Entretanto, o subsistema de compressão e codificação de vídeo desses padrões são muito semelhantes, pois ambos são baseados no Padrão MPEG-2 [9][10]. A utilização do Padrão MPEG-2 deve-se à sua boa eficiência no processo de compressão e ao fato de que este é um padrão já muito utilizado em diversas aplicações, por exemplo, transmissão de televisão digital via satélite. Dessa forma, as técnicas utilizadas pelos dois padrões para compressão de vídeo são basicamente as mesmas, com algumas diferenças nos parâmetros utilizados. Talvez a diferença mais significativa seja a possibilidade de codificação hierárquica de vídeo, proporcionada pelo Padrão DVB, que não é contemplada pelo Padrão da GA. Este artigo apresenta o Padrão MPEG-2 e descreve as técnicas de codificação de vídeo utilizadas por esse padrão.

### I. INTRODUÇÃO

Os sistemas de HDTV são totalmente digitais e estão projetados para a transmissão dos sinais de vídeo e áudio de alta qualidade dentro de faixas espectrais de mesmo tamanho que as utilizadas pelos sistemas analógicos atuais, ou seja, 6, 7 e 8 MHz [11][12].

Um sistema HDTV pode ser dividido em subsistemas, a saber: subsistema de compressão e codificação de áudio e vídeo, subsistema de multiplexação e transporte, e subsistema de transmissão que envolve codificação de fonte e modulação. O principal desafio do subsistema de vídeo é a compressão. Para verificar a importância da etapa de compressão de vídeo, considere quadros com 1080 linhas e 1920 pixels por linha, transmitidos a uma taxa de 30 quadros por segundo. Se cada pixel for representado por três componentes de cor de 8 bits cada, é necessário uma taxa de aproximadamente 1,5 Gbps para se transmitir esta seqüência. É praticamente impossível a transmissão a esta taxa dentro de uma faixa de 6 MHz utilizando técnicas de

modulação conhecidas. A solução para esse problema é a utilização de eficientes técnicas de compressão de vídeo que reduzem a taxa inicial de 1,5 Gbps para taxas de aproximadamente 20 Mbps, obtendo-se dessa forma fatores de compressão de até 75 vezes. O processo de compressão deve garantir não somente uma satisfatória redução da taxa de bits transmitidos, mas também uma qualidade de reprodução de vídeo que atenda aos requisitos exigidos pelos sistemas HDTV. Essas técnicas de compressão devem ser aliadas a técnicas de modulação e codificação de canal que sejam robustas em relação aos erros introduzidos devido às imperfeições do canal de transmissão.

Um padrão que atende aos pré-requisitos exigidos pelos sistemas HDTV é o MPEG-2. Esse padrão possibilita boas taxas de compressão de sinais de vídeo de alta resolução mantendo a qualidade de reprodução num patamar de qualidade satisfatório. O Padrão MPEG-2 foi escolhido para as etapas de compressão e codificação do sinal de vídeo dos sistemas HDTV americano e europeu. Além disso, o Padrão MPEG-2 também é utilizado nas etapas de multiplexação e transporte dos dois sistemas HDTV. Nas seções seguintes, são apresentadas técnicas utilizadas pelo Padrão MPEG-2 para compressão e codificação do sinal de vídeo digital.

### II. O PADRÃO MPEG-2

Oficialmente, o Padrão MPEG-2 chama-se “Codificação Genérica de Quadros em Movimento e Informação de Áudio Associada: Vídeo” (“Generic Coding of Moving Pictures and Associated Audio Information: Video”), e foi publicado como padrão internacional ISO/IEC 13818-2 (ou, equivalentemente ITU-T 262). O termo “Generic” é utilizado, pois esse padrão não tem como objetivo uma aplicação específica, mas apresenta um conjunto de algoritmos e ferramentas que podem ser usados para variadas aplicações em diferentes condições operacionais, isto é, com diferentes taxas, diferentes meios de transmissão ou armazenamento, etc. A seguir é dada uma lista de aplicações nas quais pode-se utilizar o Padrão MPEG-2:

- TV - Transmissão broadcasting, via satélite, ou via cabo
- HDTV - Transmissão broadcasting, via satélite, ou via cabo
- Vídeo em meios de armazenamento digital - CD-ROM, CD's de alta densidade, DVD

- Vídeo em computadores - e-mail vídeo, sistemas de informação multimídia
- “Video on Demand”
- Vídeo em redes - Vídeo transmitido através de redes ATM, Ethernet e LAN’s
- HDTV em camadas - Transmissão compatível de TV e HDTV, HDTV progressivo (“higher progressive HDTV”)
- Vídeo profissional - Edição não linear, pós-produção de estúdio (“studio post production”)

Dentro desta vasta gama de aplicações, atualmente os sistemas de HDTV têm recebido atenção especial, pois vários países estão em processo de escolha do padrão HDTV a ser adotado. O Padrão MPEG-2, graças a eficientes técnicas de compressão utilizadas, foi o instrumento que viabilizou a transmissão de HDTV em faixas de mesmo tamanho que as utilizadas por TV analógicas, ou seja, 8, 7 e até 6 MHz para radiodifusão. Na seção seguinte, é apresentado um breve histórico do desenvolvimento do Padrão MPEG-2.

### III. HISTÓRICO DO PADRÃO MPEG-2

Em 1988, o comitê técnico da ISO/IEC (JTC 1 - Joint ISO/IEC Technical Committee) em tecnologia da informação criou o grupo de trabalho “Moving Pictures Experts Group” (MPEG), com a finalidade de desenvolver padrões para a representação codificada de vídeo e áudio associado. Originalmente, três padrões que almejavam taxas de até 1.5, 10 e 40 M bits/s estavam em estudos. Esses padrões eram conhecidos como MPEG-1, MPEG-2, MPEG-3, respectivamente.

Dessa forma, o Padrão MPEG-2 tinha como objetivo a compressão de vídeo digital a taxas de até 10 Mbits/s [14]. Para aplicações com taxas superiores a 10 Mbits/s, e.g. HDTV, outro padrão também em desenvolvimento, conhecido como MPEG-3, seria utilizado. Entretanto, em julho de 1992, o desenvolvimento do Padrão MPEG-3 foi interrompido, pois percebeu-se que o Padrão MPEG-2 poderia atender a aplicações que requerem taxas acima de 10 Mbits/s com alta qualidade de reprodução. Atualmente, o Padrão MPEG-2 é utilizado na etapa de compressão de vídeo dos sistemas HDTV europeu e americano.

Em 11 de novembro de 1994, foi aprovado o Padrão MPEG-2. Em 10 de julho de 1995, o ITU-T aprovou o Padrão H.262 que equivale às três primeiras partes do ISO 13818-2 (sistemas, áudio e vídeo). Desde sua aprovação, o Padrão MPEG-2 tem sido muito utilizado em diversas aplicações, como por exemplo em sistemas de transmissão de televisão digital (DTV - Digital Television) via satélite.

### IV. PERFIS E NÍVEIS DO PADRÃO MPEG-2

O Padrão MPEG-2 foi estabelecido para atender a uma grande variedade de aplicações e por isso apresenta uma vasta gama de funcionalidades. Por exemplo, codificação hierárquica, que pode ser

utilizada para transmitir TV digital e HDTV num mesmo canal, ou para permitir uma degradação de qualidade suave na presença de erros de transmissão.

Entretanto, a utilização da sintaxe completa, incluindo todas as funcionalidades, pode inviabilizar a implementação prática na maioria das aplicações, pois torna o sistema demasiadamente complexo. Por isso, o padrão foi dividido em subconjuntos que restringem as funcionalidades e limitam os parâmetros utilizados pelo sistema. Esta divisão foi realizada através da introdução dos conceitos de perfis “Profiles” e níveis “Levels” para estipular um padrão de conformidade entre equipamentos que não suportam o padrão completo [13][15].

Os perfis definem restrições de sintaxe que correspondem à utilização de um subconjunto de algoritmos dentro do conjunto completo de algoritmos estabelecidos no padrão. Os níveis estabelecem limitantes superiores para os parâmetros utilizados por esses algoritmos.

A Tabela 1 apresenta os possíveis perfis com suas funcionalidades. Os perfis superiores englobam todas as funcionalidades dos perfis inferiores, por exemplo, o perfil “Scalable SNR” contempla todas as funcionalidades do perfil “Main” e ainda possibilita codificação hierárquica em dois níveis através de requantização. A Tabela 2 apresenta os níveis suportados pelo padrão MPEG-2, onde os valores apresentados são limitantes superiores dos parâmetros.

Profile	Funcionalidades
<i>High</i>	<ul style="list-style-type: none"> <li>• Escalonamentos espacial e através de requantização.(3 camadas, no máximo)</li> <li>• Formato 4:2:2 para as componentes de YCrCb</li> </ul>
<i>Spatial</i>	<ul style="list-style-type: none"> <li>• Escalonamento Espacial (2 camadas)</li> <li>• Representação 4:2:0 para as componentes YUV</li> </ul>
<i>Scalable SNR</i>	<ul style="list-style-type: none"> <li>• Escalonamento por requantização (2 camadas)</li> <li>• Representação 4:2:0 para as componentes YUV</li> </ul>
<i>Main</i>	<ul style="list-style-type: none"> <li>• Não suporta escalonamentos de vídeo</li> <li>• Codificação de vídeo entrelaçado</li> <li>• Acesso aleatório aos quadros da sequência.</li> <li>• Interpolação utilizando quadros tipo B. Formato 4:2:0 para as componentes YCrCb</li> </ul>
<i>Simple</i>	<ul style="list-style-type: none"> <li>• Inclui as funcionalidades do profile Main, entretanto não suporta quadros interpolados (B)</li> <li>• Formato 4:2:0 para as componentes de YCrCb</li> </ul>

Tabela 1. Perfis (“Profiles”) do padrão MPEG-2

O Padrão HDTV da Grande Aliança utiliza o perfil “Main” e o nível “High”. Esta combinação é representada pela notação MP@HL (“Main Profile” e “High Level”).

Na Europa, as combinações MP@H1440L (“Main Profile” e “High 1440 Level”) e SSP@H1440L (“Scalable SNR Profile” e “High Level”) têm sido

as preferidas, apesar de outras combinações também serem admitidas no padrão europeu.

Observa-se que um dos perfis utilizados pelo padrão europeu, o “Scalable SNR” possibilita codificação hierárquica de vídeo em duas camadas através de requantização dos coeficientes[16][17], enquanto que no padrão americano a codificação hierárquica não é possível. O padrão europeu tem esta funcionalidade adicional, pois o subsistema de transmissão do DVB-T (“DVB-Terrestrial”) permite transmissão modulação hierárquica em duas camadas de diferentes prioridades. Nesse caso, a codificação de canal é realizada separadamente para cada uma das duas camadas, e o mapeamento na constelação QAM (*Quadrature Amplitude Modulation*) é tal que oferece maior robustez na transmissão da camada de alta prioridade.

Level	Parâmetros
High	1920 amostras/linha 1152 linhas/quadro 60 quadros/segundo 80 Mbits/segundo
High 1440	1440 amostras/linha 1152 linhas/quadro 60 quadros/segundo 60 Mbits/segundo
Main	720 amostras/linha 576 linhas/quadro 30 quadros/segundo 15 Mbits/segundo
Low	352 amostras/linha 288 linhas/quadro 30 quadros/segundo 4 Mbits/segundo

Tabela 2. Níveis (“Levels”) do padrão MPEG-2

O padrão americano estabelece formatos de quadro de acordo com as limitações impostas pela combinação perfil/nível adotada. Esses formatos são apresentados na Tabela 3.

Número de Linhas Verticais	Número de Pixels por linha	Razão de Aspecto	Taxa de quadros por segundo
1080	1920	16:9	60I,30P,24P
720	1280	16:9	60P,30P,24P
480	1280	16:9	60I, 60P,30P,24P
480	640	4:3	60P,60I,30P,30I

Tabela 3. Formatos de quadro do sistema HDTV americano

Nesta tabela, P refere-se à seqüência com varredura de quadros progressiva (varredura progressiva), e I refere-se à seqüência com varredura intercalada, em que são apresentados os campos pares e ímpares de cada quadro, alternadamente. Razão de aspecto é a relação entre as dimensões horizontal e vertical de um quadro da seqüência.

V. DESCRIÇÃO DOS PROCESSOS DE CODIFICAÇÃO E DECODIFICAÇÃO

Uma seqüência de vídeo digital é um conjunto de quadros ordenados temporalmente. Cada quadro é

uma matriz cujos elementos contêm informações de cor e representam uma imagem. Os pontos desta matriz são chamados de pixels. No Padrão MPEG-2 as cores são representadas através de uma combinação de três componentes, um de luminância, Y, e dois de crominância, Cr e Cb. A luminância representa a intensidade, ou seja, a imagem em preto-e-branco (brilho da cena). As componentes de crominância contêm a informação de cor (associada a matiz e saturação em sistemas convencionais - neste caso, fase e módulo do sinal de crominância modulado em quadratura). Em HDTV, tem-se os dois sinais diferença de cor separados Cr e Cb que podem ser associados ao sinal de crominância.

Devido às características psico-visuais humanas, é possível subamostrar os sinais de crominância sem perda de qualidade perceptível. Isso ocorre porque o sistema visual humano é menos sensível às altas freqüências dos sinais de crominância do que às altas freqüências do sinal de luminância. Por isso, os padrões HDTV utilizam o formato 4:2:0, em que os sinais de crominância são subamostrados por um fator de 2 nas duas direções, horizontal e vertical. A subamostragem deve ser precedida de uma pré-filtragem, pois caso contrário ocorrerá uma distorção do sinal devido ao “alising”, que é uma sobreposição de componentes espectrais. Assim, da mesma forma que nos sistemas convencionais, os detalhes da imagem são reproduzidos pelo sinal de luminância.

A Figura 1 apresenta a disposição espacial das amostras de luminância e crominância para o formato 4:2:0. Observa-se que as amostras de crominância estão posicionadas no sentido vertical, entre as amostras de luminância, e por isso é necessário que as amostras de crominância sejam obtidas através de interpolação.

Tipicamente, as seqüências de vídeo apresentam uma alta correlação, tanto temporal quanto espacial. Dessa forma, o Padrão MPEG-2 comprime seqüências de vídeo digitalizadas basicamente explorando os três tipos de redundância presentes nestas seqüências:

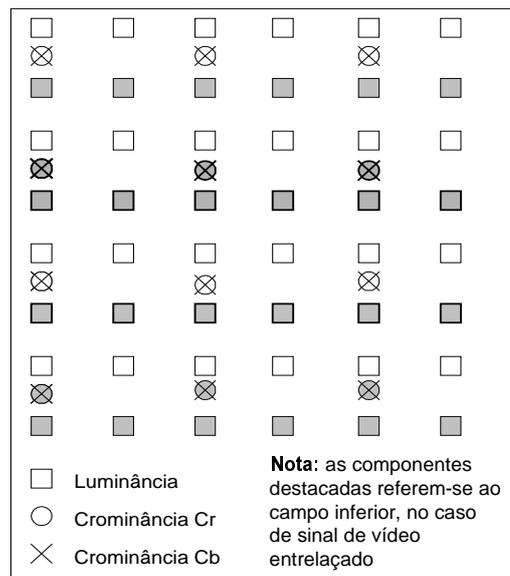


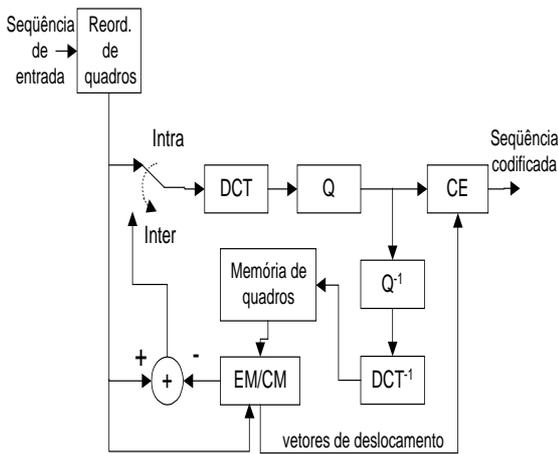
Figura 1. Localização espacial das amostras de luminância e crominância para o formato 4:2:0

- (1) redundância espacial: através da transformada DCT (Discrete Cosine Transform) combinada com codificação de entropia;
- (2) redundância temporal: através de codificação preditiva entre quadros;
- (3) redundância psicovisual: através de subamostragem da crominância.

O diagrama geral de um codificador MPEG-2 é apresentado na Figura 2. Cada etapa desse processo será brevemente discutida; entretanto, primeiro é apresentada a estrutura adotada para a seqüência de vídeo, ou seja, a forma como os dados são divididos, a fim de serem codificados.

**VI. ESTRUTURA DA SEQÜÊNCIA DE VÍDEO**

Cada quadro da seqüência é dividido em unidades utilizadas no processo de codificação. A menor unidade de codificação no algoritmo MPEG é o bloco. Ele é utilizado na aplicação da transformada DCT, e tem tamanho 8x8 pixels. O bloco pode ser de luminância (Y), ou de crominância (Cr ou Cb).



**Figura 2.** Diagrama geral de um codificador MPEG-2

**Legenda:**

- CE - Codificação de entropia (Huffman e corrida de zeros).
- DCT - Transformação DCT de blocos de 8x8 pixels
- DCT<sup>-1</sup> - Transformação DCT inversa
- Q - Quantizador
- Q<sup>-1</sup> - Quantizador inverso
- EM/CM - Estimativa e compensação de movimento

O macrobloco é a unidade básica de codificação no algoritmo MPEG. Ele consiste de segmentos de 16 x 16 pixels, e é a unidade utilizada nas etapas de estimativa e compensação de movimento. Na Figura 3 é apresentado um exemplo de macrobloco de formato 4:2:0, composto de quatro blocos de luminância e dois de crominância. O formato 4:2:0 refere-se à forma como subamostra-se a crominância do sinal, nesse caso, a crominância é subamostrada por um fator de dois nas duas direções, horizontal e vertical. Outros dois formatos são também possíveis: 4:2:2 em que a crominância é subamostrada por um fator de dois somente na horizontal, e o 4:4:4, em que a crominância não é subamostrada.

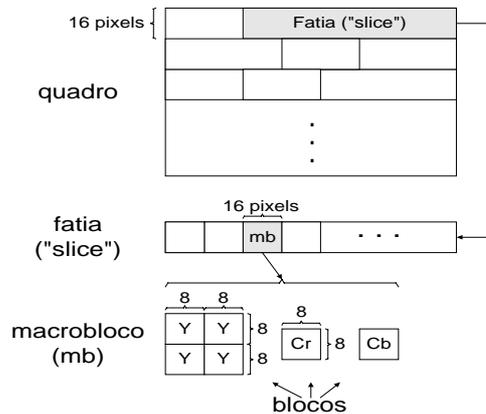
O “slice” é uma fatia horizontal de macroblocos e serve como unidade de resincronismo. Uma fatia horizontal de macroblocos de um quadro pode conter mais de um slice. Além disso, podem existir áreas do quadro onde não há slices e que, portanto, não são codificadas. Entretanto, todos os perfis (“profiles”) do padrão utilizam uma estrutura de slice restrita (“restricted slice structure”), em que todo o quadro é segmentado em slices, como no exemplo da Figura 3.

Os quadros também são divididos em grupos chamados GOP (“Group of pictures”). O GOP é um conjunto de quadros que possibilita acesso randômico. No Padrão MPEG-2, o conceito de GOP é um pouco diferente do Padrão MPEG-1. No primeiro caso sua utilização é opcional, enquanto no segundo é obrigatória. Tipicamente, 6 a 15 quadros são agrupados em um GOP e é estabelecida uma ordem de tipos de quadros (I, P, ou B) que se repete ao longo da seqüência. Os tipos de quadros são apresentados na seção seguinte.

**VII. TIPOS DE QUADROS**

Os quadros podem ser divididos em três tipos, de acordo com a codificação preditiva utilizada:

- quadro I (intracodificado): este quadro é codificado de maneira independente dos outros quadros.
- quadros P (predito): este quadro é codificado utilizando-se o quadro de referência anterior como preditor.
- quadros B (interpolado): o preditor utilizado na codificação deste quadro é baseado em dois quadros de referência, um anterior e outro posterior.



**Figura 3.** Estrutura de um quadro não entrelaçado de formato 4:2:0

Cada quadro P ou B é codificado tomando-se a diferença entre os quadros originais e uma estimativa desse quadro. Para o cálculo desta estimativa, no caso de quadros P, o quadro de referência anterior é utilizado em um processo que envolve estimativa e compensação de movimento (EM / CM). Para os quadros B, o processo é semelhante, só que, nesse caso, são utilizados dois quadros de referência (um anterior e outro posterior) que, após o processo de EM/CM, são interpolados para formarem o quadro

estimado. Os quadros de referência são os quadros P ou I temporalmente mais próximos do quadro atual, um anterior e outro posterior. Esses quadros são recuperados no codificador e armazenados na

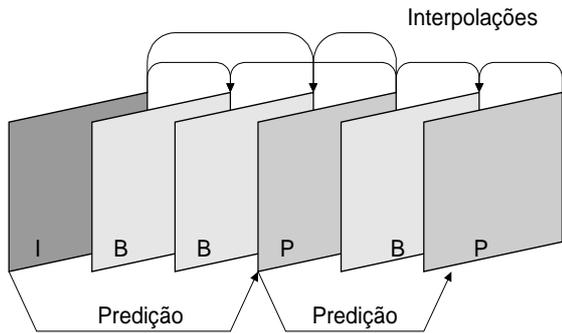


Figura 4. Exemplo de estrutura temporal de seqüência

memória de quadros para serem utilizados no processo de EM/CM. Os quadros do tipo B nunca são usados como referência. A Figura 4 apresenta um exemplo de uma seqüência de quadros, ordenados temporalmente. Nesse exemplo, para cada quadro, é dado o tipo de predição e os quadros de referência relacionados.

Os quadros a serem codificados geralmente estão dispostos em ordem temporal, ou seja, de acordo com a ordem de apresentação na tela quando a seqüência é reproduzida. Quando os quadros são codificados, eles são incluídos no "bitstream" de acordo com a ordem em que são decodificados, sendo esse ordenamento chamado ordem de codificação. No decodificador é feito um novo reordenamento e os quadros são novamente dispostos em ordem temporal.

Quando a seqüência não contém quadros do tipo B, a ordem de codificação é igual à ordem temporal. Entretanto, se existem quadros B na seqüência, a ordem de codificação é diferente da ordem temporal, pois deve-se codificar as duas referências de um quadro B antes de codificá-lo. O seguinte exemplo ilustra esta situação:

Na entrada do codificador tem-se:

I<sub>1</sub> B<sub>2</sub> B<sub>3</sub> P<sub>4</sub> B<sub>5</sub> B<sub>6</sub> P<sub>7</sub> P<sub>8</sub> B<sub>9</sub> I<sub>10</sub> P<sub>11</sub>

Na saída do codificador e na entrada do decodificador tem-se:

I<sub>1</sub> P<sub>4</sub> B<sub>2</sub> B<sub>3</sub> P<sub>7</sub> B<sub>5</sub> B<sub>6</sub> P<sub>8</sub> I<sub>10</sub> B<sub>9</sub> P<sub>11</sub>

Na saída do decodificador tem-se:

I<sub>1</sub> B<sub>2</sub> B<sub>3</sub> P<sub>4</sub> B<sub>5</sub> B<sub>6</sub> P<sub>7</sub> P<sub>8</sub> B<sub>9</sub> I<sub>10</sub> P<sub>11</sub>

**VIII. ESTIMAÇÃO E COMPENSAÇÃO DE MOVIMENTO**

Para se explorar a correlação temporal da seqüência, utilizam-se técnicas DPCM ("Differential Pulse Code Modulation"). Basicamente, esta técnica consiste em codificar a diferença entre o sinal e sua predição ao invés de codificar o próprio sinal. No caso do Padrão MPEG-2, esta diferença é calculada entre o macrobloco atual e sua predição, que pode ser uma região de um quadro anterior ou uma média entre regiões de um quadro anterior e outro posterior.

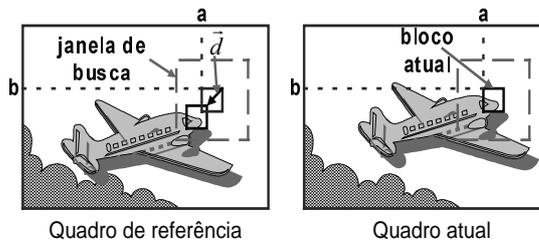


Figura 5. Estimação de movimento

O modo mais simples de se calcular uma predição para um macrobloco é utilizar macroblocos de um ou dois quadros de mesma região espacial. Esta técnica é eficiente em regiões da cena em que não há movimento. Entretanto, em áreas onde há movimento, a correlação entre pixels de mesma posição espacial geralmente é pequena. Por isso, são utilizadas técnicas de estimação e compensação de movimento para tornar mais eficiente esse processo de predição.

As técnicas de estimação de movimento utilizadas pelo Padrão MPEG-2 são baseadas em casamento de blocos. Para cada bloco é realizada uma busca do vetor de deslocamento dentro de uma região do quadro de referência chamada janela de busca. Essa busca consiste em minimizar uma medida de disparidade, por exemplo, o erro quadrático médio (MSE - "Mean Square Error"), definido pela Equação (1).

$$MSE(i, j) = \frac{1}{M \cdot N} \sum_{m=a}^{M+i} \sum_{n=b}^{N+j} [Q_a(m, n) - Q_{ref}(m + i, n + j)]^2 \quad (1)$$

onde,  $Q_a(x, y)$  é o quadro atual,  
 $Q_{ref}(x, y)$  é o quadro de referência,  
 $(i, j)$  é o deslocamento entre os dois blocos,  
 $(a, b)$  é a posição do bloco atual.

A Figura 5 apresenta um exemplo de estimação de movimento em que o deslocamento do bloco atual em relação à mesma posição espacial no quadro de referência é representado pelo vetor  $\vec{d}$  (vetor de deslocamento).

**IX. CODIFICAÇÃO DE UM MACROBLOCO**

Quanto à forma de predição utilizada, basicamente, um macrobloco pode ser codificado de quatro modos:

- intracodificado,
- predito utilizando-se o quadro de referência anterior,
- predito utilizando-se o quadro de referência posterior,
- ou interpolado (utiliza os dois quadros de referência para formar a predição).

O Padrão MPEG-2 contempla algumas variações desses quatro modos básicos, por exemplo, o "dual-prime" e o 16x8 MC[18][19]. Neste último o macrobloco é dividido em dois e o processo de EM/CM é realizado para cada uma das partes de tamanho 16x8 pixels. Esse modo só é utilizado quando os campos do quadro são codificados individualmente.

Os processos de predição e interpolação envolvem estimação e compensação de movimento, técnicas que foram apresentadas na seção anterior. Se o macrobloco atual é predito, seja em relação ao quadro de referência anterior ao posterior, o macrobloco do quadro de referência que apresentou melhor casamento é utilizado como predição. No caso de macroblocos interpolados, a predição é formada tomando-se a média entre os macroblocos de melhor casamento de cada um dos quadros de referência.

Os dados a serem codificados, referentes a um macrobloco predito ou interpolado, são resultantes da diferença entre esse macrobloco e a sua predição. Tem-se então um macrobloco-diferença que é codificado de maneira semelhante a um macrobloco intracodificado, utilizando-se transformada DCT e codificação de entropia (código de "Huffman" e "run-length"). Os vetores de deslocamento resultantes do processo de EM/CM são codificados e incluídos no "bitstream".

Todos os macroblocos de um quadro I são intracodificados. Nos quadros P pode-se ter macroblocos preditos (em relação ao quadro de referência anterior), ou macroblocos intracodificados, o que equivale a utilizar um macrobloco nulo (todos os elementos iguais a zero) como predição. Finalmente, nos quadros B pode-se ter macroblocos interpolados, macroblocos preditos (em relação ao quadro de referência anterior ou posterior), ou macroblocos intracodificados. A escolha de como é formada a predição para um macrobloco fica a cargo do projetista do codificador e, geralmente, depende da correlação entre o macrobloco atual e o macrobloco de melhor casamento de cada quadro de referência. Nos quadros B e P, existe a possibilidade de não se codificar determinado macrobloco. Nesse caso, o macrobloco é reconstruído no decodificador utilizando-se somente os quadros de referência e considerando nulos os vetores de movimento, se o quadro atual é do tipo P, ou tomando-se os vetores de movimento do último macrobloco codificado, se o quadro atual é do tipo B.

Os dados resultantes desse processo, seja um macrobloco original ou um macrobloco-diferença, é dividido em blocos de tamanho 8x8. A cada bloco é aplicada uma transformação DCT bidimensional. A transformação DCT não resulta em compressão, entretanto descorrelaciona os coeficientes tornando mais eficiente o processo de compressão através de outras técnicas. A DCT 2-D é definida pela Equação (2), e a sua inversa pela Equação (3).

$$F(u, v) = \frac{2C(u)C(v)}{n} \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} f(x, y) \cos\left\{\frac{(2x+1)u\pi}{2n}\right\} \cos\left\{\frac{(2y+1)v\pi}{2n}\right\} \quad (2)$$

$$f(x, y) = \frac{2}{n} \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} C(u)C(v)F(u, v) \cos\left\{\frac{(2x+1)u\pi}{2n}\right\} \cos\left\{\frac{(2y+1)v\pi}{2n}\right\} \quad (3)$$

onde  $C(u), C(v) = \begin{cases} 1/\sqrt{2} & \text{para } u, v = 0 \\ 1 & \text{para } u, v \neq 0 \end{cases}$ .

Após a transformação DCT, os coeficientes resultantes são quantizados. O coeficiente DC (Direct Current- frequência (0,0)) de macroblocos intra é quantizado de forma diferente dos outros coeficientes. O Padrão MPEG-2 só define o processo de quantização inversa, permitindo flexibilidade no projeto do quantizador. A quantização inversa dos coeficientes DC dos blocos intracodificados é dada pela Equação (4), onde  $CQ[x][y]$  são os coeficientes quantizados,  $C'[x][y]$  os coeficientes recuperados, e **intra\_dc\_mult** o passo de quantização. Os coeficientes restantes são recuperados (inversamente quantizados) de acordo com a Equação (5), onde  $W[x][y]$  é a matriz de quantização utilizada.

$$C'[x][y] = \text{intra\_dc\_mult} \times CQ[x][y] \quad (4)$$

$$C'[x][y] = \frac{(2 \times CQ[x][y] + k) \times W[x][y] \times \text{quantiser\_scale}}{32} \quad (5)$$

onde

$$k = \begin{cases} 0 & \text{para blocos intracodificados} \\ \text{Sinal}(CQ[x][y]) & \text{para blocos não intracodificados} \end{cases}$$

e

$$\text{Sinal}(CQ[x][y]) = \begin{cases} 0 & \text{se } CQ[x][y] = 0 \\ 1 & \text{se } CQ[x][y] > 0 \\ -1 & \text{se } CQ[x][y] < 0 \end{cases}$$

Através dos parâmetros **intra\_dc\_mult** e **quantiser\_scale** pode-se alterar dinamicamente o passo de quantização para controlar a relação de compromisso entre taxa de bits transmitidos e qualidade de reprodução da imagem. Pode-se também utilizar esses parâmetros para se obter uma taxa de bits aproximadamente constante.

Para os formatos 4:4:4 ou 4:2:2, quatro matrizes de quantização podem ser utilizadas. Por outro lado, no formato 4:2:0 somente duas matrizes de quantização são utilizadas, uma para blocos intracodificados e outra para blocos não intracodificados (e.g., preditos). A Figura 6 apresenta as duas matrizes de quantização recomendadas pelo Padrão MPEG-2, a primeira (Figura 6a) para blocos intracodificados, e a segunda (Figura 6b) para blocos não intracodificados. Além dessas matrizes, podem ser utilizadas outras que, nesse caso, devem ser enviadas junto com os dados codificados (e.g., no cabeçalho de um quadro).

O próximo passo é a codificação dos coeficientes quantizados. Os coeficientes DC dos blocos intra (intracodificados) são codificados diferentemente dos demais. Para os coeficientes DC dos blocos intra, é utilizado um esquema DPCM, no qual codifica-se a diferença entre o coeficiente DC do bloco atual e o coeficiente DC do bloco anterior. A esta diferença é atribuída uma palavra-código que indica o número de bits necessários para se representar esta diferença.

Os coeficientes restantes (coeficientes AC) são codificados através de uma técnica que envolve codificação "run-length" e codificação de Huffman (VLC- "Variable Length Coding"). Após a quantização, vários coeficientes tornam-se nulos, o

que viabiliza a utilização de codificação "run-length". Essa técnica é eficiente quando um símbolo, nesse caso zero (coeficiente nulo), ocorre com muita frequência. Ao invés de se codificar cada um dos símbolos individualmente, codifica-se um número relativo à quantidade de repetições consecutivas desse símbolo. A codificação VLC consiste em representar os símbolos que ocorrem com mais frequência por palavras-código com menor número de bits. Para se aplicar a codificação "run-length", os coeficientes AC de cada bloco são ordenados de acordo com uma das duas matrizes de varredura apresentadas na Figura 6. Geralmente, utiliza-se a varredura em zig-zag, apresentada na Figura 6(a). A varredura alternada, apresentada na Figura 6(b), é utilizada quando a seqüência de vídeo é entrelaçada e o bloco codificado contém informação dos campos par e ímpar. Após ordenados, os coeficientes AC de cada bloco são codificados da seguinte forma: a cada coeficiente não nulo é atribuída uma palavra-código que representa o nível (módulo) do coeficiente e o número de coeficientes nulos entre o coeficiente atual e o último coeficiente não nulo codificado desse bloco. Esta palavra-código é seguida de um bit que representa o sinal do coeficiente. Na tabela VLC, existem duas palavras-código especiais. Uma delas serve para indicar fim de bloco, isto é, indica que não há mais nenhum coeficiente não nulo no bloco atual. A outra representa um código de escape ("scape code"), que indica que são utilizadas palavras de comprimento fixo para representar o comprimento da corrida de zeros (6 bits) e o coeficiente com sinal (12 bits).

O código de escape é utilizado, pois o número de combinações entre corridas de zeros e os níveis de coeficientes é muito grande. Considerando que as corridas de zeros podem variar entre 0 e 63, e o número de níveis pode variar entre 1 e 2048, seriam necessárias 13.107.264 (64x2048) palavras-código para representar todas as combinações possíveis, tornando impraticável esse esquema de codificação. Entretanto, a grande maioria dessas combinações ocorre com uma frequência muito baixa, e são representadas utilizando-se o código de escape como descrito anteriormente. As combinações estatisticamente mais frequentes (um pouco mais que 100, de acordo com as tabelas VLC do Padrão MPEG-2) são representadas individualmente por palavras-código. Mesmo com o uso do código de escape, tipicamente, a codificação VLC proporciona um fator de compressão de pelo menos dois, quando comparada com códigos de comprimento fixo.

Uma certa distribuição de probabilidade está associada ao código VLC utilizado para codificar os coeficientes quantizados [20]. As tabelas de códigos VLC utilizadas são pré-definidas pelo padrão e não podem ser alteradas para se adequarem à real distribuição dos coeficientes a serem codificados.

Entretanto, as matrizes de quantização podem ser alteradas a cada quadro na tentativa de se obter uma maior aproximação entre as distribuições do quadro atual e do código VLC [3]. Nesse caso, deve-se considerar o *overhead* necessário para transmissão dessas matrizes.

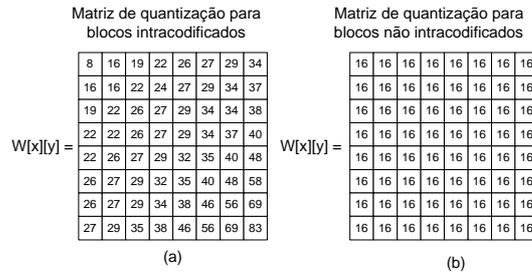


Figura 6. Matrizes de quantização recomendadas pelo Padrão MPEG-2, para blocos intracodificados (a) e para blocos não intracodificados (b).

X. ATUALIZAÇÃO DE QUADROS (“REFRESHING”)

Quando se utiliza codificação diferencial, um dos problemas que podem ocorrer é a propagação de erros. Por exemplo, se ocorre erro na decodificação de um macrobloco que é utilizado como referência para codificar outros macroblocos, o erro vai propagar até que esse macrobloco de referência seja substituído por outro decodificado sem erros. Por isso, o Padrão MPEG-2 recomenda que cada macrobloco que ocupa uma determinada posição espacial deve ser intracodificado, no máximo a cada 132 quadros.

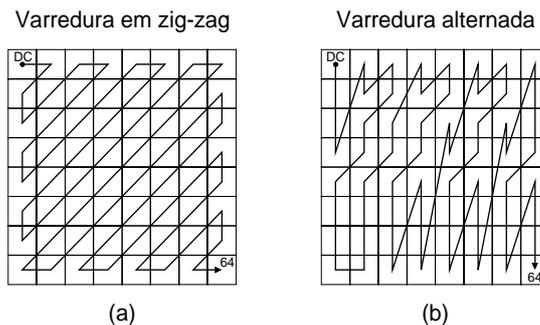


Figura 7. Modos de varredura dos coeficientes do bloco transformado

Para os sistemas HDTV recomenda-se que a atualização de quadros, ou seja, a codificação de quadros intracodificados (quadros I), ocorra em intervalos de no máximo 0,5 segundos. Além disso, recomenda-se que seja enviado o cabeçalho de seqüência antes de cada quadro intracodificado. Esses procedimentos são necessários, pois quando um usuário sintoniza um determinado canal, essas informações (cabeçalho de seqüência e quadro I) serão necessárias para o início do processo de decodificação e reprodução do sinal de vídeo. O desempenho da qualidade de mudança de canal pelo usuário pode ser bastante afetado caso esses procedimentos não sejam observados.

XI. BUFFER DO CANAL

Os padrões HDTV europeu e americano transmitem a taxas de bits constantes. O sistema americano utiliza taxas de 19,28 e 38,57 Mbps para transmissão terrestre e via cabo, respectivamente. Por outro lado, o sistema europeu terrestre (DVB-T) possibilita uma gama de

taxas que variam entre 4,98 a 31,67 Mbps, de acordo com parâmetros de modulação e de codificação de canal utilizados. Essas diferentes taxas possibilitam a utilização desse sistema para outras aplicações além de HDTV, por exemplo televisão digital com definição convencional (SDTV - Standard Definition Television), que utiliza taxas de transmissão de 4 a 6 Mbps.

A utilização de códigos de comprimento variável (VLC) e codificação preditiva no processo de compressão faz com que a taxa de bits seja variável ao longo do tempo em função da estatística do sinal de vídeo original. Como os sistemas HDTV americano e europeu utilizam taxas de bits constantes, é necessário um procedimento para controle de taxa. Esse controle é realizado através da utilização de um buffer na saída do codificador que simula o buffer da entrada do decodificador. O tamanho desse buffer deve ser tal que os atrasos de transmissão decorrentes da inserção desse buffer sejam toleráveis pelo sistema. Mecanismos de realimentação são utilizados para alterar parâmetros dos algoritmos de compressão de forma a evitar que o buffer do codificador não seja sobrecarregado (“overflow”), nem esvaziado abaixo de um determinado limiar (“underflow”). Os parâmetros mais utilizados para esse controle são os relacionados com os passos de quantização dos coeficientes.

## XII. CONCLUSÕES

O objetivo de se alcançar a taxa compatível com 6 MHz de faixa de transmissão é alcançado usando-se combinações de várias técnicas (DPCM, EM/CM, DCT, run-length e VLC). O processo segue o padrão MPEG-2 que permite a viabilização da HDTV comercial para *broadcasting*.

## XIII. AGRADECIMENTOS

Os autores agradecem o apoio e suporte fornecidos pelos seguintes órgãos de pesquisa: **FAPESP** - Fundação de Amparo à Pesquisa do Estado de São Paulo, **CNPq** - Conselho Nacional de Desenvolvimento Científico e Tecnológico, **CAPES** - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; **CPqD** - Fundação Centro de Pesquisa e Desenvolvimento em Telecomunicações e **FAEP** - Fundo de Apoio ao Ensino e à Pesquisa da UNICAMP.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] K. Challapali, *et al.* “**The Grand Alliance System for US HDTV**,” Proceeding of the IEEE, USA, vol. 83, no. 2, p. 158-174, Fevereiro 1995
- [2] ATSC Standard A/53. Digital Television Standard. 1995.
- [3] ATSC Standard A/54. Guide to the Use of the ATSC Digital Television Standard. 1995.
- [4] <http://www.atsc.org>
- [5] European Telecommunication Standard ETS 300 744, “**Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital Terrestrial television (DVB-T)**,” ETSI, março 1997.
- [6] <http://www.etsi.org>
- [7] <http://www.dvb.org>
- [8] Willian Y. ZOU, e Yiyan WU, “**COFDM: an Overview**,” IEEE Transactions on Broadcasting, vol. 41, no. 1, março 1995.
- [9] <http://www.mpeg.org>
- [10] ITU-T Recommendation H.262, ISO/IEC 13182-2, “**Information Technology - Generic Coding of Moving Pictures and Associated Audio Information: Video**,” agosto, 1995.
- [11] E. Petejan, “**The HDTV Grand Alliance System**,” Proceeding of the IEEE, USA, vol. 83, no. 7, p. 1094-1105, Julho 1995.
- [12] R. Schafer, “**The Status of HDTV in Europe**,” IEEE Communications Magazine, USA, p. 120-125, junho 1996.
- [13] L. Chiariglione, “**MPEG and multimedia communications**,” IEEE Trans. Circuits Syst. Video Technol., vol. 7, no. 1, pp. 5-18, fevereiro 1997.
- [14] L. Chiariglione, “**The development of an integrated audiovisual coding standard: MPEG**,” Proc. of IEEE, vol. 83, no. 2, pp. 151-157, fevereiro 1995.
- [15] T. Sikora, “**MPEG Digital Video-Coding Standards**,” IEEE Sig. Proc. Mag., pp. 82-100, setembro 1997.
- [16] L. R. Mendes, e Max H. M. Costa, “**Transmissão de vídeo utilizando o padrão MPEG-2 em duas camadas**,” XV Simpósio Brasileiro de Telecomunicações, pp. 585-589, setembro 1997, Recife PE.
- [17] L. R. Mendes, R. T. del Pietro, Max H. M. Costa, “**Two-scale Transmission of DCT-coded video over lossy packet networks**,” Proc. of SBT/IEEE International Telecommunications Symposium, pp. 510-515, agosto 1998, São Paulo SP.
- [18] B. G. Haskell, A. Puri, e A. N. Netrevali, **Digital Video: An Introduction to MPEG-2**. New York: Chapman & Hall, 1997.
- [19] J. L. Mitchel et al., **MPEG Video Compression Standard**. New York: Chapman & Hall, 1997.
- [20] N. S. Jayant, e P. Noll, **Digital Coding of Waveforms - Principles and Applications to Speech and Video**. Englewood Cliffs: Prentice-Hall, 1984.

**Luiz Rômulo Mendes** recebeu o título de Engenheiro Eletricista com ênfase em Eletrônica pela Universidade Federal de Uberlândia, e Mestre em Engenharia Elétrica e Telecomunicações pela UNICAMP. Atualmente é estudante de Doutorado do DECOM/FEEC/UNICAMP, na Faculdade de Engenharia Elétrica e de Computação da UNICAMP. Tem interesses nas áreas de processamento digital de sinais de áudio e vídeo e redes de alta velocidade.

**Yuzo Iano** recebeu os títulos de Engenheiro Eletrônico, Mestre e Doutor em Engenharia Elétrica através da UNICAMP. Atualmente é Professor Adjunto do DECOM/FEEC/ UNICAMP. É o responsável pelo Laboratório de Comunicações Visuais do DECOM. Seus interesses estão em processamento digital de sinais de áudio e vídeo. Trabalha com pesquisas em televisão e em especial com a HDTV desde 1986.

**Vicente I. Becera Sablón** recebeu os títulos de Engenheiro Eletrônico, na UCLV (Universidad Central de Las Villas, Cuba), de Mestre em Engenharia Eletrônica e Telecomunicações na UNICAMP. Atualmente é Professor Assistente da Faculdade de Matemática-Física e Computação da UCLV. É estudante de Doutorado do Departamento de

Comunicações da FEEC/UNICAMP. Trabalha em processamento digital de sinais de áudio e vídeo.

**Renê Togni Del Pietro** recebeu os títulos de Engenheiro Eletrônico e Mestre em Engenharia Eletrônica e Telecomunicações pela UNICAMP. (Universidade Estadual de Campinas). Atualmente é estudante de Doutorado do Departamento de Comunicações da Faculdade de Engenharia Elétrica e de Computação da UNICAMP (DECOM/FEEC/ UNICAMP). Seus interesses atuais estão em processamento digital de sinais de áudio e vídeo.

e-mail: romulo@decom.fee.unicamp.br;  
yuzo@decom.fee.unicamp.br;  
vicente@decom.fee.unicamp.br;  
togni@decom.fee.unicamp.br