

Neste ponto existem duas possibilidades. Se o receptor for do tipo Digital o sinal de TV digitalizado oriundo do decompressor MPEG2 vai direto para *display* do aparelho.

Se o receptor for do tipo analógico, o sinal da saída do decompressor é convertido novamente, sendo modulado para a versão de TV analógica para depois ser enviado para o aparelho de recepção comum. Nesse último caso, o conjunto é denominado *Set Top Box* e pode ser adquirido para ser usado em conjunto com o televisor já existente.

Alguns receptores podem ainda ser dotados de recursos adicionais como equalizadores, antenas adaptativas ou combinador de diversidade que permitem uma melhoria adicional para condições extremas de degradação devido a ruídos e interferências.

---

## 3 PROCESSOS DE COMPRESSÃO

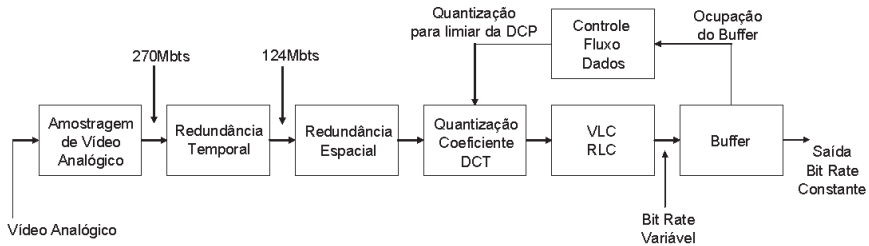
O sinal de vídeo digitalizado constitui uma taxa de *bits* tão elevada que pode superar 1 Gbps, o qual não é possível ser transmitido na largura de banda de 6MHz reservada para os canais de televisão. Os processos de modulação empregados nos sistemas ATSC, DVB-T e ISDB-T (veja o Capítulo 4) permitem transmitir um nível com taxa de *bits* de  $\approx 20$ Mbps na banda de 6MHz dos canais de TV. Assim sendo, tornou-se necessária a redução da taxa de *bits* de 1 Gbps para 20Mbps, a fim de que o sinal para a transmissão na banda seja de 6MHz dos canais de TV. Esse processo de redução de taxa de *bits* é conhecido por “compressão”. Em 1986, um grupo de estudo denominado *Joint Photograph Expert Group* (JPEG), sob os auspícios do *International Standard Organization* (ISO), foi criado com o objetivo de desenvolver e padronizar um método de compressão de sinal digitalizado de imagens estáticas. Anos mais tarde decidiu-se criar um método de compressão de imagens dinâmicas baseado no princípio JPEG.

Esse grupo foi denominado *Moving Pictures Expert Group* (MPEG) e deu origem ao padrão de compressão de vídeo, conhecido como MPEG2, utilizado pelos três sistemas de TV digital: ATSC, DVB-T e ISDB-T. Também existem muitos outros métodos de compressão de vídeo, como o MPEG-4/H264 e o WM9 (*Windows Media 9*). Contudo, por ser o MPEG2 o método mais consagrado para uso em TV Digital, este será descrito neste trabalho. O objetivo é formar uma noção dos princípios que norteiam o processo de compressão, sem entrar em detalhes profundos sobre o funcionamento dos blocos envolvidos.

### 3.1 MPEG2

O princípio básico do processo MPEG2 mostrado no diagrama da Figura 20 é utilizar técnicas que reduzem a qualidade da imagem de modo imperceptível, junta-

mente com técnicas que não afetam a qualidade da imagem, usando principalmente o processo de eliminar redundâncias na informação.

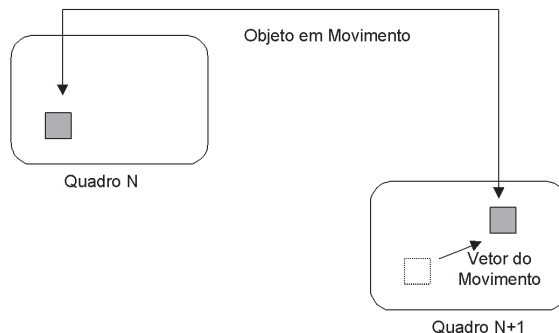


**Figura 20** Diagrama em blocos da codificação MPEG2

### 3.1.1 Amostragem e digitalização de vídeo

O primeiro passo do processo é fazer a amostragem e a digitalização dos sinais de vídeo, como mostrado nos itens 1.4, 1.5.3 e 1.5.4. Como foi visto, a digitalização dos sinais  $Y$ ,  $P_B$ ,  $P_R$  de TV analógica tradicional com formato 4:2:2 ocorreu para a taxa de *bits* de 270Mbps. Já a digitalização dos sinais  $Y$ ,  $P_B$ ,  $P_R$  de HDTV com formato 22:11:11 pode levar a uma taxa de *bits* de 1.485Gbps.

- Redundância Temporal:** consiste no aproveitamento da similaridade existente entre os quadros sucessivos que formam uma imagem dinâmica. Um exemplo prático seria a imagem de um jogador de golfe sacando uma bola. Após o saque, a diferença entre os quadros sucessivos seria somente a posição da bola a cada instante. Dessa maneira, o algoritmo enviaria a informação de um quadro completo mais o vetor de deslocamento da imagem efetivamente dinâmica. A diferença entre o quadro  $N$  e o quadro  $N + 1$  da Figura 21 é o deslocamento do bloco. Portanto, em vez de enviar os dados dos dois quadros completos, envia-se a informação do quadro  $N$  completa mais a informação do vetor movimento do quadro  $N + 1$ .



**Figura 21** Sensibilidade do olho humano à frequência espacial

Existe um jogo de compromisso nesse processo, pois caso haja um erro na transmissão do quadro N, por exemplo, o quadro N + 1 não poderá ser formado corretamente. Para minimizar esse problema, durante a transmissão, são enviados alguns quadros com informação completa dentro da seqüência de quadros, seqüência essa denominada Group of Pictures (GOP).

A compensação de movimentos se baseia nas semelhanças entre as imagens subseqüentes e permite a transmissão apenas das diferenças. Esse método pode ser entendido como a subtração de uma imagem pela sua predecessora, ou seja: imagem  $i_{2(t)}$  será subtraída pela imagem  $i_1(t-\tau)$ , o que corresponde a uma subtração em nível de *pixels* no mesmo lugar x, y.

Para garantir que esse processo não gere perdas, o caminho de realimentação do codificador contém um decodificador funcional que irá subtrair exatamente o sinal que um codificador terá à disposição para somar ao sinal recebido.

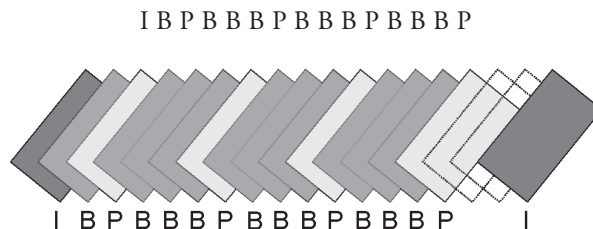
A Figura 22 mostra a seqüência de quadros em que os símbolos têm os seguintes significados:

I – *Intra Frames*: quadros completos, usados no início da seqüência

B – *Bidirectional Frames*: possuem a informação das diferenças entre o quadro atual e o quadro anterior e posterior.

P- *Predicted Frames*: são baseados no anterior, possuem uma alta compressão e trazem mudança em relação ao último quadro, podendo ainda ser usados como referência para o próximo quadro.

Se for transmitida uma imagem com o fator  $N = 15$  e  $M = 3$  (em que N é o número total de quadros e M é o número de quadros tipo B que fica entre dois quadros tipo P), tem-se a transmissão mostrada na Figura 22 com a seguinte seqüência de quadros:



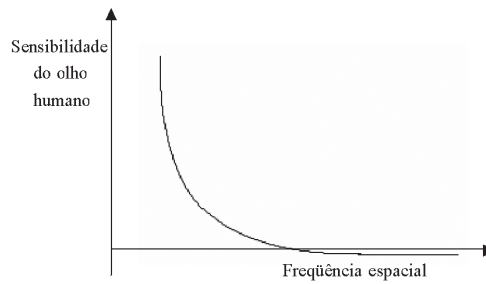
**Figura 22** Transmissão de quadros no sistema MPEG2

Além de introduzir o conceito de estimação de movimentos, o MPEG2 introduz um método para regular a taxa de transmissão através do controle do processo de quantificação. Um *buffer* na saída do codificador irá permitir uma taxa de transmissão

constante. Caso o *buffer* se encontre próximo de exceder sua capacidade de armazenamento, ele irá ordenar que o processo de quantificação diminua a resolução do vídeo, diminuindo assim o montante de dados gerados.

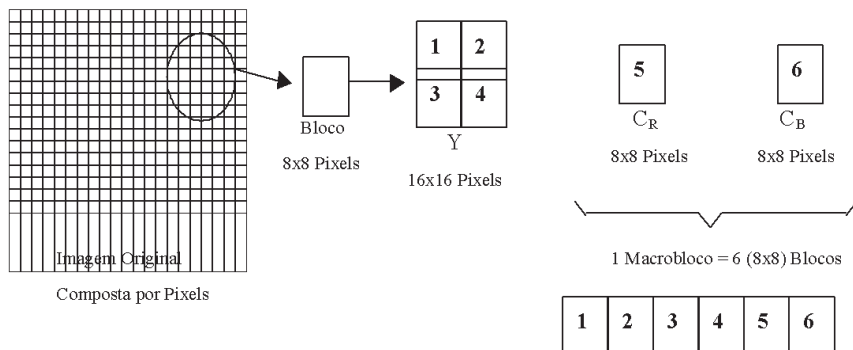
- **Redundância Espacial:** a redundância espacial ou frequência espacial consiste na semelhança dos *pixels* adjacentes de uma imagem. Um exemplo é um avião passando no céu sem nuvens, em que a informação relevante a ser transmitida é o avião, e o fundo é a parte da imagem azul cujo conteúdo de imagem é praticamente uniforme.

A Figura 23 mostra que o olho humano possui baixa sensibilidade para frequências espaciais altas.



**Figura 23** Sensibilidade do olho humano à frequência espacial

Baseando-se nesse conceito físico de que as altas frequências podem ser eliminadas sem causar degradações perceptíveis na imagem, a idéia do MPEG2 é converter a amplitude espacial da imagem em frequência espacial através da DCT (transformada discreta de co-seno). A imagem é subdividida em blocos de  $8 \times 8$  *pixels*, como mostra a Figura 24.



**Figura 24** Amostragem da figura a ser comprimida

Cada bloco é composto por 64 valores que podem variar (quando digitalizado em 8 *bits*, tem-se  $2^8 = 256$ ) de 0 a 255 para a luminância ou -128 a +127 para a crominância. A esse bloco é aplicada a DCT, que resultará em 64 coeficientes que representam a frequência espacial, sendo que o primeiro coeficiente no topo e à esquerda representa o componente DC, e o último coeficiente localizado abaixo é a maior frequência espacial contida na informação.

A DCT obedece à equação 8:

$$Z_{k,l} = \frac{1}{4} C_k C_l \sum_{i=0}^{i=7} \sum_{j=0}^{j=7} x(i,j) \cos \frac{\pi(2i-1)k}{16} \cos \frac{\pi(2j+1)l}{16} \quad (8)$$

Dependendo do número de detalhes contidos no bloco original, boa parte dos coeficientes assumirá valores próximos de zero, sendo esses valores passíveis de ser descartados.

A aplicação do DCT reduz consideravelmente a taxa da informação a ser transmitida, e é considerada uma compressão sem perdas e um processo completamente reversível. Aproveitando-se das características de reduzida sensibilidade do olho humano às altas frequências espaciais, é possível, sem degradação perceptível da imagem, eliminarem-se alguns coeficientes resultantes do processo DCT. Esse processo é denominado limiarização (*thresholding*). Após a limiarização, os coeficientes restantes são quantizados com uma precisão que decresce inversamente proporcional à medida que há um acréscimo das altas frequências espaciais, como mostrado na Figura 23. Após a aplicação da limiarização e quantização, o processo torna-se do tipo “com perdas”, pois parte das informações foi completamente descartada e não pode mais ser recuperada.

- **Variable Length Code (VLC) e Run Length Code (RLC):** após o sinal passar pela quantização, ele é codificado pelo processo denominado RLC, muito empregado em *softwares* de compressão como o ZIP, por possuir grande eficiência em mensagens com elementos semelhantes repetidos.

Exemplificando, tomando-se a palavra ARARAQUARA, em vez de transmitir a palavra inteira, transmite-se o posicionamento das letras e o conteúdo da palavra: 5A, 3R, 1Q e 1U. Quanto maior o número de elementos repetidos, maior é a eficiência da codificação.

Em conjunto, é aplicado o código VLR, também denominado algoritmo de Huffman, cujo princípio básico é atribuir símbolos de menor comprimento às informações mais frequentes. Aplicando o processo à palavra ARARAQUARA, tem-se:

5 caracteres A  
3 caracteres R  
1 caractere Q  
1 caractere U

Se possuírem disponíveis apenas quatro símbolos de diferentes comprimentos será atribuída a seguinte codificação:

Símbolo “0” para letra A  
Símbolo “01” para letra R  
Símbolo “111” para letra Q  
Símbolo “00000” para letra U

É claro que os métodos empregados no MPEG2 são mais complexos do que o simples exemplo mostrado acima. Entretanto, ele serve para mostrar o princípio empregado para comprimir os dados de uma transmissão de vídeo.

- **Buffer:** é o dispositivo destinado a controlar o fluxo de dados do MPEG2, garantindo que a taxa de *bits* na saída do compressor seja sempre constante. O controle dessa taxa é feito durante o processo de quantização. Para isso, o circuito de controle de fluxo realimenta com informação sobre o status do fluxo de saída, para que o circuito de quantização do DCT possa sempre adequar o nível de quantização.

### 3.1.2 Considerações gerais

Após todas essas etapas, a taxa de compressão obtida pode ser superior a 50:1, o que significa que um sinal de vídeo de HDTV digitalizada que possui uma taxa superior a 1 Gbps passa a ter uma taxa de  $\approx 20$ Mbps depois de comprimido.

Existem vários outros métodos de compressão, como o MPEG4, o WM9 e o H264, mas como no momento estes não estão sendo usados para transmissão de TV terrestre, não serão discutidos neste trabalho.

## 3.2 Compressão de áudio

A gama de frequências audíveis pelo ouvido humano vai de 16Hz a 20kHz, e o limite superior varia de pessoa para pessoa. Para obter a qualidade de CD, obedecendo ao critério de Nyquist, é preciso amostrar o sinal de áudio pelo menos 44.000 vezes por segundo. Utilizando-se código de 16 *bits*, tem-se 704 Kbits; se a transmissão for para música estéreo, a taxa será de 1,4Mbps. Surge então a necessidade de um algoritmo de compressão que reduza essa taxa de *bits*, para possibilitar uma transmissão de maneira eficiente.

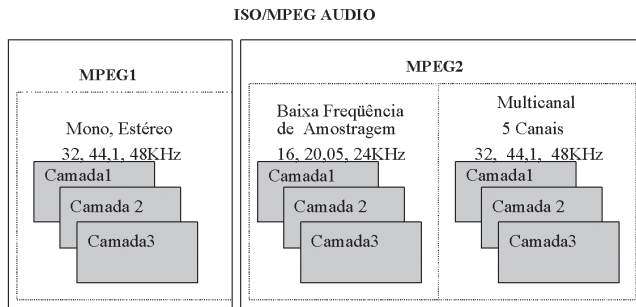
O codificador de áudio MPEG consegue reduzir a taxa de dados de um CD em aproximadamente 12 vezes, sem perda significativa da qualidade.

O principal fator do qual o processo se vale para essa redução consiste no uso da característica do ouvido humano de não ser um aparelho perfeito para captar sons; o MPEG aproveita essas imperfeições para comprimir o áudio.

Ao se ouvir uma orquestra tocando simultaneamente violinos, trombetas e os instrumentos de percussão, os ouvintes não são capazes de captar todos os sons simultaneamente porque alguns dos sons estão camuflados pelos sons dos instrumentos mais fortes, uma vez que o ouvido humano é mais sensível aos sons mais graves do que aos sons mais agudos.

Uma gravação em CD dessa orquestra contém todos os sons dos instrumentos da orquestra. Mas, se as partes que as pessoas não são capazes de ouvir forem removidas, consegue-se ouvir a mesma orquestra com a mesma qualidade do CD. De maneira geral, esse é o funcionamento do sistema de compressão do MPEG Áudio.

O MPEG é o primeiro padrão internacional no domínio de compressão de áudio de alta fidelidade. Em particular, o MPEG define a sintaxe da seqüência de *bits* de áudio codificado a baixas taxas de *bits*. O algoritmo de codificação não é definido pela norma MPEG.



**Figura 25** Hierarquia entre as camadas 1, 2 e 3 do MPEG áudio

A Figura 25 mostra as hierarquias das camadas do MPEG Áudio.

### 3.2.1 MPEG1 Áudio

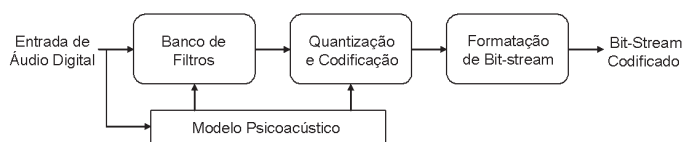
O MPEG1 providencia o esquema de compressão para fontes de áudio estéreo e mono com qualidade semelhante a CD Áudio. A parte do algoritmo correspondente ao áudio é dividida em três camadas: (1) **MPEG1 Áudio Camada 1**, que oferece menor taxa de compressão, e é usado em sistemas de áudio de consumo. A sua vantagem é menor custo de implementação; (2) **MPEG1 Áudio Camada 2**, que oferece maior taxa de compressão. É encontrado em aplicações profissionais ou em aplicações de consumo mais sofisticado; e (3) **MPEG1 Áudio Camada 3**, que oferece maior taxa de compressão e é encontrado em Rede Digital de Serviços Integrados (RDSI). O seu custo é maior e apresenta maior complexidade de codificação e decodificação.

A idéia principal é que uma aplicação do MPEG1 viabiliza usar a camada que oferece o melhor compromisso entre a necessidade de computação e o desempenho de compressão. O áudio pode ser codificado em qualquer uma das camadas.

O MPEG1 Áudio destina-se a transformar sinais de áudio amostrados a taxas de 32KHz, 44,1KHz ou 48KHz e codificá-los para uma taxa de 32Kbps a 192Kbps.

### 3.2.2 MPEG2 Áudio

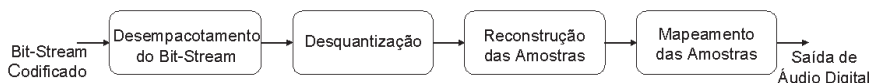
O MPEG2 Áudio foi desenvolvido com base no MPEG1, fazendo com que os benefícios deste proporcionassem maior qualidade e maior eficiência, propiciando um padrão de qualidade para satisfazer os requisitos necessários para as redes digitais.



**Figura 26** Codificador MPEG Áudio

Além dos canais mono e estéreo que o MPEG1 suporta, foi adicionado suporte para múltiplos canais. Assim, o MPEG2 tem a possibilidade de suportar até cinco canais de som (esquerdo, direito, centro e dois canais de *surround* e mais um canal de graves), ou, em extensão, sete canais mais um canal de graves. Isso possibilita ter em casa o som com o mesmo desempenho de algumas salas de cinema, permitindo a reprodução de som estereofônico mais realista. As figuras 26 e 27 mostram a estrutura de MPEG Áudio.

O MPEG2 possui outras funções, como a característica “multilíngüe”, permitindo, por exemplo, que um filme possa ser ouvido em diferentes linguagens, de acordo com a escolha do telespectador.



**Figura 27** Decodificador MPEG Áudio

Para efetuar a compressão de áudio, existem duas alternativas: a primeira consiste em reduzir a frequência de amostragem, e a segunda, em reduzir a resolução da amostra para um valor inferior. Para se obter uma boa qualidade, não se pode fazer grandes alterações na frequência de amostragem, pois o ouvido humano escuta até 20kHz. Pelo teorema de Nyquist, é necessário ter uma taxa de amostragem equivalente a duas vezes a frequência que se deseja transmitir; portanto, é preciso ir até 44,1K amostras por segundo.

A razão de escolher código de 16 bits deve-se ao desejo de obter boa relação sinal-ruído, de quantificação proveniente do processo de digitalização do sinal. Para cada bit que se adiciona ao código há uma melhoria de 6dB na relação sinal-ruído. Ao se utilizar código de 8 bits por amostra, pode-se ainda perceber ruído de fundo nos momentos de silêncio de música. Mas, para esse nível de codificação, se o nível do som é elevado, não se ouve nenhum ruído. Esse efeito é chamado de Efeito Máscara, e é a chave para o método de codificação utilizado pelo MPEG2.

Com o sistema de CD Áudio consegue-se uma relação sinal-ruído de 90dB, o que está dentro da gama dinâmica do ouvido humano, pois nesta faixa não se ouve nenhum ruído.

### 3.2.3 Modelo psicoacústico, efeito máscara e bandas críticas

O modelo psicoacústico baseia-se no efeito da sensibilidade do ouvido humano, que tem uma gama dinâmica de 96dB em níveis de som, cuja frequência varia de 20Hz a 20kHz. Entretanto, a sensibilidade não é uniforme para toda gama de frequência, sendo mais sensível na faixa de 1kHz a 3kHz.

Quando dois tons de frequências próximas estão simultaneamente presentes, o ouvido humano percebe melhor o som com nível mais elevado. Quando os dois tons têm frequências mais afastadas uma da outra, o ouvido humano percebe mais nitidamente a presença simultânea dos dois. A esse efeito dá-se o nome de Efeito Máscara – efeito que permite aumentar o nível de ruído de fundo à volta do som forte, uma vez que o ruído já será mascarado.

A largura de banda das curvas dos sinais que causam o Efeito Máscara não é proporcional à frequência. Essa largura, chamada de Dimensão de Banda Crítica, é de cerca de 100Hz para baixas frequências, aumentando muito para frequências mais elevadas.

- MPEG2 Áudio Camada 1: o mapeamento tempo-frequência é constituído por um banco de filtros polifásicos com 32 sub-bandas, os quais têm a característica de combinar a baixa complexidade computacional com opções de implementação flexíveis. Essas sub-bandas são igualmente espaçadas em frequência.

O modelo psicoacústico usa uma *Fast Fourier Transformer* (FFT) de 512 pontos ( $2^8$ ), de modo a obter uma informação espectral detalhada do sinal. A saída da FFT é usada para determinar as máscaras tonais e monotonais do sinal (ruído). Cada máscara produz um limiar de mascaramento dependente da sua tonalidade, intensidade e frequência. Para cada sub-banda, os limiares de mascaramento são combinados de modo a formar um limiar global de mascaramento. Esse limiar é comparado com o nível do sinal para cada sub-banda, produzindo uma relação sinal-máscara, ou *Signal Mask Rate* (SMR), que é a entrada do quantizador.

O quantizador/codificador examina as amostras de cada sub-banda, procura o valor máximo absoluto e quantiza a 6 bits. Esse valor é chamado de fator de escala da sub-banda. A seguir, determina a alocação de bits para cada sub-banda. É possível mesmo que em bandas muito mascaradas seja obtido o resultado “0 bit”, de modo que nenhuma amostra seja codificada. Finalmente, as amostras da sub-banda são linearmente quantizadas. O empacotador da trama tem um trabalho muito simples. Cada trama começa com uma informação de cabeçalho que será usada para sincronização e contabilização e um código cíclico (CRC), usado para detecção e correção de erros. Cada uma das 32 sub-bandas usa 4 bits para alocação de bits e 6 bits para fator de escala. Os bits restantes da trama são usados para amostras da sub-banda, com uma cauda adicional para informação extra.

- MPEG2 Áudio Camada 2: o mapeamento tempo-freqüência é igual ao da Camada 1. Usa um banco de filtros polifásicos de 32 sub-bandas. O modelo psicoacústico é semelhante ao da Camada 1 porém usa uma *Fast Fourier Transformer* (FFT) de 2.014 pontos para maior resolução de freqüência. Usa também o mesmo procedimento que a Camada 1 para produzir as relações sinal-máscara para cada uma das 32 sub-bandas.

O quantizador/codificador é também semelhante ao da Camada 1, gerando fatores de escala de 6 bits para cada sub-banda. As tramas da Camada 2 são três vezes mais longas que as tramas da Camada 1, assim a Camada 2 permite a cada banda uma seqüência de três fatores de escala sucessivos, e o codificador usa um, dois ou três, dependendo do quanto eles diferem uns dos outros. Isso resulta num fator de dois na redução do débito binário para os fatores de escala, quando comparado com os da Camada 1. As alocações de bits são calculadas de maneira similar à da Camada 1. O empacotador da trama usa uma estrutura de cabeçalho e código cíclico (CRC) similar ao da Camada 1.

O número de *bits* usados para descrever as alocações de bits varia com a sub-banda:

4 para as sub-bandas baixas  
3 para as sub-bandas intermediárias  
2 para as sub-bandas altas

Os fatores de escala (um, dois ou três, dependendo dos dados) são codificados em conjunto com um código de 2 bits que descreve qual combinação de fatores de escala se trata.

As amostras da sub-banda são quantificadas de acordo com a relação de bits, sendo então combinadas em grupos de três denominados grânulos. Cada grânulo é codificado com uma palavra de código. Isso permite que a Camada 2 capture muito mais redundância do sinal que a Camada 1.