



PI 00153419

REPÚBLICA FEDERATIVA DO BRASIL
MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA E COMÉRCIO EXTERIOR
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL

CARTA PATENTE Nº PI 0015341-9

O INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL concede a presente PATENTE DE INVENÇÃO, que outorga ao seu titular a propriedade da invenção caracterizada neste título, em todo o território nacional, garantindo os direitos dela decorrentes, previstos na legislação em vigor.

(21) Número do Depósito: PI 0015341-9

(22) Data do Depósito: 08/11/2000

(43) Data da Publicação do Pedido: 17/05/2001

(51) Classificação Internacional: H04N 7/26

(30) Prioridade Unionista: 08/11/1999 US 09/436, 085

(54) Título: "MÉTODO PARA DETERMINAR UMA DESIGNAÇÃO DE TAMANHO DE BLOCO PARA UM BLOCO DE ENTRADA DE PIXELS DE IMAGEM A SER USADA NA COMPRESSÃO DO BLOCO DE ENTRADA, SISTEMA DE COMPRESSÃO DE IMAGENS PARA COMPRIMIR UM BLOCO DE DADOS DE PIXELS, MÉTODO E SISTEMA PARA COMPRESSÃO DE UM BLOCO DE DADOS DE PIXELS DE UMA IMAGEM, BEM COMO, MEIO LEGÍVEL POR COMPUTADOR"..

(73) Titular: QUALCOMM INCORPORATED, Sociedade Norte Americana. Endereço: 5775 Morehouse Drive, San Diego, California 92121-1714, Estados Unidos (US).

(72) Inventor: KADAYAM THYAGARAJAN; MICHAEL J. MERRITT

Prazo de Validade: 10 (dez) anos contados a partir de 06/01/2015, observadas as condições legais.

Expedida em: 6 de Janeiro de 2015.

Assinado digitalmente por:

Liane Elizabeth Caldeira Lage
Diretora de Patentes Substituta



"MÉTODO PARA DETERMINAR UMA DESIGNAÇÃO DE TAMANHO DE BLOCO PARA UM BLOCO DE ENTRADA DE PIXELS DE IMAGEM A SER USADA NA COMPRESSÃO DO BLOCO DE ENTRADA, SISTEMA DE COMPRESSÃO DE IMAGENS PARA COMPRIMIR UM BLOCO DE DADOS DE PIXELS, E, MÉTODO E SISTEMA PARA COMPRESSÃO DE UM BLOCO DE DADOS DE PIXELS DE UMA IMAGEM."

Campo da Invenção

A presente invenção está relacionada ao processamento de imagens. Mais especificamente, a presente invenção está relacionada a um esquema de compressão para sinais de imagem utilizando blocos e sub-blocos de dados de coeficientes de transformada de cosseno discreta codificados de tamanho adaptável.

Descrição da Técnica Correlacionada

No campo de transmissão e recepção de sinais de vídeo, tais como são usados para projetar "filmes", vários aperfeiçoamentos estão sendo feitos nas técnicas de compressão de imagens. Muitos dos sistemas de vídeo atuais e propostos fazem uso de técnicas de codificação digital. A codificação digital propicia uma robustez para os enlaces de comunicações que resiste a problemas tais como desvanecimento por múltiplas trajetórias e jamming ou interferência de sinais, cada um dos quais poderia, de outra forma, degradar seriamente a qualidade das imagens. Ademais, técnicas digitais facilitam o uso de técnicas de criptografia de sinal, que se mostram úteis ou mesmo necessárias para aplicações governamentais e várias aplicações de transmissão comercial recém desenvolvidas.

Vídeo de alta definição é uma área que se beneficia de técnicas de compressão de imagem aperfeiçoadas. Ao ser proposta pela primeira vez, a transmissão através do ar de vídeo de alta definição (ou mesmo a transmissão por cabos ou fibras ópticas) parecia inviável devido às elevadas exigências em termos de largura

de banda. Os típicos sistemas de transmissão sem fio ou outros sendo projetados não acomodavam facilmente largura de banda necessária. No entanto, foi constatado que a compressão de sinais digitais de vídeo pode ser conseguida em um nível que capacita a transmissão usando larguras de banda razoáveis. Tais níveis de compressão de sinal, em conjunto com a transmissão digital do sinal, irá possibilitar a um sistema de vídeo transmitir com menos potência com maior imunidade a problemas do canal, ocupando concomitantemente uma largura de banda mais desejável e útil.

Uma técnica de compressão capaz de oferecer níveis significativos de compressão, enquanto preserva o nível desejado de qualidade para sinais de vídeo, utiliza blocos e sub-blocos de dados de coeficiente de transformada de cosseno discreta (DCT - Discrete Cosine Transform) de tamanho adaptável. Tal técnica será a seguir designada como o método de Transformada de Cosseno Diferencial de Tamanho de Bloco Adaptável (ABSDCT - Adaptive Block Size Differential Cosine Transform). Tal técnica está descrita na Patente US 5,021,891, intitulada "ADAPTIVE BLOCK SIZE IMAGE COMPRESSION METHOD AND SYSTEM", em nome da Requerente da presente invenção e aqui incorporada por referência. As técnicas DCT estão também descritas na Patente US 5,107,345, intitulada "ADAPTIVE BLOCK SIZE IMAGE COMPRESSION METHOD AND SYSTEM", em nome da Requerente da presente invenção e aqui incorporada por referência. Além disso, o uso da técnica ABSDCT em combinação com uma técnica de Transformada Quadtree Diferencial está descrita na Patente US 5,452,104, intitulada "ADAPTIVE BLOCK SIZE IMAGE COMPRESSION METHOD AND SYSTEM", também em nome da Requerente da presente invenção e aqui incorporada por referência. Os sistemas descritos nestas patentes utilizam

o que é denominado como codificação "intra-quadro" (intra-frame), em que cada quadro de dados de imagem é codificado independentemente do conteúdo de qualquer outro quadro. Usando a técnica ABSDCT, a taxa de dados alcançável pode
5 ser reduzida de cerca de 1,5 bilhão de bits por segundo para aproximadamente 50 milhões de bits por segundo sem degradação perceptível da qualidade de imagem.

A técnica ABSDCT pode ser usada para comprimir uma imagem ou sinal preto e branco ou a cores representando
10 a imagem. O sinal de alimentação de cor pode estar em um formato YIQ, com Y sendo a amostra de luminância, ou brilho e I e Q sendo as amostras de crominância, ou cor, para cada bloco de pixels 4x4. Outros formatos conhecidos tais como os formatos YUV ou RGB podem ser também usados. Devido à
15 baixa sensibilidade espacial do olho à cor, a maioria das pesquisas mostrou que uma sub-amostra dos componentes de cor por um fator de quatro nas direções horizontal e vertical é razoável. Assim sendo, um sinal de vídeo pode ser representado por quatro componentes de luminância e
20 dois componentes de crominância.

Pelo uso da ABSDCT, um sinal de vídeo será de modo geral segmentado em blocos de pixels para processamento. Para cada bloco, os componentes de luminância e crominância são passados a um intercalador de
25 blocos. Como exemplo, um bloco de 16 x 16 (pixels) pode ser apresentado ao intercalador de blocos, que ordena ou organiza as amostras de imagem dentro de cada bloco de 16 x 16 para produzir blocos e sub-blocos compostos de dados para análise de transformada de cosseno discreta (DCT). O
30 operador DCT constitui um método para converter um sinal amostrado em tempo para uma representação de frequências do mesmo sinal. Ao converter para uma representação de frequências, as técnicas DCT demonstraram permitir níveis

muito elevados de compressão, uma vez que os quantizadores podem ser projetados para aproveitar as características de distribuição de frequências de uma imagem. Em uma modalidade preferida, uma DCT 16 x 16 é aplicada a uma
5 primeira ordenação, quatro DCTs 8 x 8 são aplicadas a uma segunda ordenação, dezesseis DCTs 4 x 4 são aplicadas a uma terceira ordenação, e sessenta e quatro DCTs 2 x 2 são aplicadas a uma quarta ordenação.

A operação DCT reduz a redundância espacial
10 inerente na fonte de vídeo. Após a DCT ser realizada, a maior parte da energia de sinal de vídeo tende a ficar concentrada em uns poucos coeficientes DCT. Uma transformada adicional, a Transformada Quad-Tree Diferencial (DQT), pode ser usada para reduzir a
15 redundância entre os coeficientes DCT.

Para o bloco 16 x 16 e cada sub-bloco, os valores de coeficientes DCT e o valor DQT (se a DQT for usada) são analisados para determinar o número de bits necessários para codificar o bloco ou sub-bloco. A seguir, o bloco ou a
20 combinação de sub-blocos que requer o menor número de bits para codificação é escolhido para representar o segmento de imagem. Como exemplo, dois sub-blocos 8 x 8, seis sub-blocos 4 x 4 e oito sub-blocos 2 x 2 podem ser escolhidos para representar o segmento de imagem.

25 O bloco ou combinação de sub-blocos escolhido é a seguir apropriadamente disposto em ordem em um bloco 16 x 16. Os valores de coeficientes DCT/DQT podem a seguir passar por ponderação de frequências, quantização, e codificação (tal como codificação de comprimento variável)
30 na preparação para transmissão.

Apesar da técnica ABSDCT acima descrita funcionar notavelmente bem, esta demanda muita computação. Dessa forma, a implementação de hardware compacto da técnica pode

ser difícil. Uma técnica alternativa que tornaria a implementação de hardware mais eficiente é desejada. Um método e sistema para compressão de imagens que são mais eficientes do ponto de vista da computação são providos
5 pela presente invenção da forma descrita a seguir.

Resumo da Invenção

A presente invenção consiste de um sistema e método para compressão de imagem utilizando blocos e sub-blocos de dados de coeficiente de transformada de coseno
10 discreta dimensionados de forma adaptável. Em uma modalidade, um bloco 16 x 16 de pixels é alimentado a um codificador. O codificador compreende um elemento designador de tamanho de bloco, que segmenta o bloco alimentado de pixels para processamento. A designação de
15 tamanho de bloco é baseada nas variâncias do bloco e sub-blocos subdivididos alimentados. De um modo geral, áreas com variâncias maiores serão subdivididas em blocos menores, enquanto áreas com variâncias menores não serão subdivididas, contanto que os valores médios do bloco e
20 sub-blocos fiquem dentro de diferentes faixas predeterminadas. Dessa forma, em primeiro lugar o limite de variância de um bloco é modificado a partir de seu valor nominal, dependendo de seu valor médio, e a seguir a variância do bloco é comparada a um limite e, se a
25 variância for maior que o limite, então o bloco é subdividido.

A designação de tamanho de bloco é provida a um elemento de transformada, que transforma os dados de pixels em dados do domínio da frequência. A transformada é
30 efetuada somente sobre o bloco e sub-blocos selecionados através da designação de tamanho de bloco. Os dados transformados a seguir passam por quantização e serialização. Como exemplo, a varredura em ziguezague pode

ser utilizada para serializar os dados para produzir um fluxo de dados. O fluxo de dados pode a seguir ser codificado por um codificador de comprimento variável em preparação para transmissão. Os dados codificados são
5 enviados através de um canal de transmissão para um decodificador, onde os dados de pixels são reconstruídos em preparação para apresentação.

Breve Descrição dos Desenhos

As características, objetivos e vantagens da
10 presente invenção ficarão mais claros a partir da descrição detalhada apresentada a seguir, quando lida em conjunto com os desenhos, nos quais as mesmas referências numéricas identificam itens correspondentes e nos quais:

Figura 1 - é um diagrama em blocos de um sistema
15 de processamento de imagens que incorpora o sistema e o método de designação de tamanho de bloco com base na variância, de acordo com a presente invenção;

Figura 2 - é um fluxograma ilustrando as etapas
de processamento envolvidas na designação de tamanho de
20 bloco com base em variância; e

Figuras 3a, 3b e 3c - ilustram um exemplo de designação de tamanho de bloco, a decomposição quad-tree correspondente e os dados PQR correspondentes.

Descrição Detalhada das Modalidades Preferidas

25 Para facilitar a transmissão digital de sinais digitais e usufruir dos correspondentes benefícios, é de um modo geral necessário empregar alguma forma de compressão de sinais. Para obter alta definição em uma imagem resultante, é também importante que a alta qualidade da
30 imagem seja mantida. Além disso, a eficiência computacional é desejada para uma implementação compacta de hardware, a qual é importante em muitas aplicações.

A presente invenção propicia um sistema ou equipamento e método para compressão de imagens que levam em consideração a qualidade de imagem e a eficiência computacional ao efetuar a compressão de imagem. A compressão de imagem da presente invenção é baseada nas técnicas de transformada de coseno discreta (DCT). De um modo geral, uma imagem para ser processada no domínio digital seria composta de dados de pixels divididos em um arranjo de blocos não superpostos, com tamanho de $N \times N$. Uma DCT bidimensional pode ser efetuada em cada bloco. A DCT bidimensional é definida pela seguinte relação:

$$X(k, l) = \frac{\alpha(k)\beta(l)}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x(m, n) \cos \left[\frac{(2m+1)\pi k}{2N} \right] \cos \left[\frac{(2n+1)\pi l}{2N} \right],$$

$$0 \leq k, l \leq N-1$$

em que:

$$\alpha(k), \beta(k) = \begin{cases} 1, & \text{se } k = 0 \\ \sqrt{2}, & \text{se } k \neq 0 \end{cases} \quad e$$

$x(m, n)$ é a localização do pixel (m, n) dentro de um bloco $N \times N$; e

$X(k, l)$ é o coeficiente DCT correspondente.

Uma vez que os valores de pixel não são negativos, o componente DCT $X(0,0)$ é sempre positivo e usualmente possui a maior energia. De fato, para imagens típicas, a maior parte da energia da transformada fica concentrada em torno do componente $X(0,0)$. É tal propriedade de compactação de energia que torna a técnica DCT, um método de compressão tão atraente.

A técnica de compressão de imagens da presente invenção utiliza codificação adaptável de contraste para obter redução adicional da taxa de bits. Foi observado que a maioria das imagens naturais é constituída por áreas

neutras (flat) de variação relativamente lenta e áreas ocupadas (busy) tais como contornos de objetos e textura de alto contraste. Os esquemas de codificação adaptável de contraste se aproveitam de tal fator ao designar mais bits para as áreas ocupadas e menos bits para as áreas menos ocupadas.

A codificação adaptável de contrastes é também útil para reduzir o efeito bloqueador. Na implementação de outras técnicas DCT de codificação, o efeito bloqueador é talvez o problema mais importante para a qualidade de imagem. Além disso, o efeito bloqueador tende a ser mais perceptível em áreas ocupadas da imagem. No entanto, foi observado que o efeito bloqueador é reduzido quando é usada uma DCT de menor tamanho. O efeito bloqueador se torna virtualmente invisível quando uma DCT 2 x 2 é usada, apesar de que o desempenho em bits por pixel possa sofrer. Dessa forma, a codificação de contraste adaptável pode reduzir o efeito bloqueador ao designar menores tamanhos de blocos (e desse modo mais bits) para as áreas ocupadas e tamanhos de blocos maiores para as áreas relativamente vazias.

Outra característica da presente invenção é a de que ela utiliza a codificação intra-quadro (processamento espacial) em lugar de codificação inter-quadro (processamento espaço-temporal). Uma razão para a adoção da codificação intra-quadro é a elevada complexidade do receptor necessária para processar sinais de codificação inter-quadros. A codificação inter-quadro requer inerentemente múltiplos buffers de quadros além de circuitos de processamento mais complexos. Em muitas aplicações, uma menor complexidade é necessária para uma implementação real.

Uma segunda razão para utilizar a codificação intra-quadro é a de que pode ocorrer uma situação, ou

material de programa, que possa fazer com que um esquema de codificação espaço-temporal entre em colapso e apresente mal desempenho. Como exemplo, os filmes de 24 quadros por segundo podem se inserir em tal categoria, uma vez que o tempo de integração, devido ao obturador mecânico, é relativamente curto. O tempo de integração curto permite um maior grau de alternância ou "aliasing" temporal. A presunção de correlação quadro a quadro se rompe para movimentos rápidos na medida em que se torna tremido.

10 Outra razão para o uso de codificação intra-quadros é a de que um esquema de codificação espaço-temporal é de mais difícil padronização quando estão envolvidas frequências da linha de energia de 50 e 60 Hz. A televisão atualmente transmite sinais em 50Hz ou 60Hz. O uso de um esquema intra-quadro, constituindo uma estratégia digital, pode se adaptar à operação em 50 ou 60 Hz, ou mesmo 24 quadros por segundo, contrabalançando a taxa de quadros com a resolução espacial.

Para finalidades de processamento de imagem, a operação DCT é efetuada nos dados de pixels que são divididos em um arranjo de blocos não-superpostos. Note-se que apesar dos tamanhos de blocos serem aqui descritos como possuindo dimensões de $N \times N$, considera-se que vários tamanhos de blocos podem ser usados. Como exemplo, um tamanho de bloco de $N \times M$ pode ser utilizado, em que tanto N como M são números inteiros, com M sendo maior ou menor que N . Outro aspecto importante é o de que o bloco pode ser dividido em pelo menos um nível de sub-blocos, tal como $N/i \times N/i$, $N/i \times N/j$, $N/i \times M/j$ e assim por diante, em que i e j são números inteiros. Além disso, o tamanho de bloco exemplar, tal como aqui descrito, é um bloco de 16×16 pixels com bloco e sub-blocos correspondentes de coeficientes DCT. É também considerado que vários outros

números inteiros, tais como valores inteiros pares ou ímpares, podem ser usados, por exemplo, 9 x 9.

Fazendo agora referência à Figura 1, um sistema de processamento de imagem 100 que incorpora o sistema de compressão da presente invenção é apresentado. O sistema de processamento de imagem 100 compreende um codificador 102 que comprime um sinal de vídeo recebido. O sinal comprimido é transmitido através de um canal de transmissão 104, e recebido por um decodificador 106. O decodificador 106 decodifica o sinal recebido em amostras de imagem, que podem então ser apresentadas ou mostradas.

De um modo geral, uma imagem é dividida em blocos de pixels para processamento. Um sinal de cor pode ser convertido do espaço RGB para o espaço YC_1C_2 , com Y sendo o componente de luminância, ou brilho, e C_1 e C_2 sendo os componentes de cromaticidade ou cor. Devido à baixa sensibilidade espacial do olho para a cor, muitos sistemas sub-amostram os componentes C_1 e C_2 por um fator de quatro nas direções horizontal e vertical. No entanto, a sub-amostragem não é necessária. Uma imagem de resolução total, conhecida como um formato 4:4:4, pode ser muito útil ou necessária em algumas aplicações tais como aquelas designadas como cobrindo o "cinema digital". Duas representações YC_1C_2 possíveis são a representação YIQ e a representação YUV, ambas bem conhecidas pelos técnicos na área. É também possível empregar uma variação da representação YUV, conhecida como YCbCr.

Em uma modalidade preferida, cada um dos componentes Y, Cb e Cr é processado sem sub-amostragem. Dessa forma, uma alimentação de um bloco 16 x 16 de pixels é provida ao codificador 102. O codificador 102 compreende um elemento designador de tamanho de bloco 108, que efetua a designação de tamanho de bloco em preparação para a

compressão de vídeo. O elemento designador de tamanho de bloco 108 determina a decomposição de bloco do bloco 16 x 16 com base nas características perceptíveis da imagem no bloco. A designação de tamanho de bloco subdivide cada
 5 bloco 16 x 16 em blocos menores de uma forma quad-tree, dependendo da atividade dentro de um bloco 16 x 16. O elemento designador de tamanho de bloco 108 gera dados quad-tree, denominados como dados PQR, cujo comprimento pode ficar entre 1 e 21 bits. Dessa forma, caso a
 10 designação de tamanho de bloco determine que um bloco de 16 x 16 deve ser dividido, o bit R dos dados PQR é determinado e é seguido por quatro bits adicionais de P dados correspondentes aos quatro blocos 8 x 8 divididos. Caso a designação de tamanho de bloco determine que quaisquer dos
 15 blocos 8 x 8 deve ser subdividido, então quatro bits adicionais de dados Q para cada bloco 8 x 8 subdividido são adicionados.

Fazendo agora referência à Figura 2, um fluxograma mostrando detalhes da operação do elemento
 20 designador de tamanho de bloco 108 é provido. O algoritmo usa a variância de um bloco como uma medida na decisão para subdividir um bloco. Começando na etapa 202, um bloco 16 x 16 de pixels é lido. Na etapa 204, a variância v_{16} , do bloco 16 x 16 é computada. A variância é computada da
 25 seguinte forma:

$$var = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x_{i,j}^2 - \left(\frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x_{i,j} \right)^2$$

em que $N = 16$ e $x_{i,j}$ é o pixel na i^a linha, coluna j^a dentro do bloco $N \times N$. Na etapa 206 é primeiramente modificado o limite de variância T_{16} para prover um novo limite T'_{16} caso o valor médio do bloco fique entre dois valores

predeterminados, a seguir a variância do bloco é comparada ao novo limite $T'16$.

Caso a variância $v16$ não seja maior que o limite $T16$, na etapa 208 é escrito ou gravado o endereço de início do bloco 16×16 , e o bit R dos dados PQR é ajustado para 0 para indicar que o bloco 16×16 não está subdividido. O algoritmo a seguir lê o próximo bloco 16×16 de pixels. Caso a variância $v16$ seja maior que o limite $T16$, na etapa 210, o bit R dos dados PQR é ajustado para 1 para indicar que o bloco 16×16 deve ser subdividido em quatro blocos 8×8 .

Os quatro blocos 8×8 , $i = 1:4$, são considerados seqüencialmente para subdivisão adicional, tal como mostrado na etapa 212. Para cada bloco 8×8 , é computada a variância, $v8_i$, na etapa 214. Na etapa 216, primeiramente o limite de variância $T8$ é modificado para prover um novo limite $T'8$ caso o valor médio do bloco fique entre dois valores predeterminados, a seguir a variância do bloco é comparada a este novo limite.

Caso a variância $v8_i$ não seja maior que o limite $T8$, na etapa 218 o endereço de início do bloco 8×8 é escrito, e o bit Q correspondente, Q_i , é ajustado para 0. O próximo bloco 8×8 é a seguir processado. Caso a variância $v8_i$ seja maior que o limite $T8$, na etapa 220, o bit Q correspondente, Q_i , é ajustado para 1 para indicar que o bloco 8×8 deve ser subdividido em quatro blocos 4×4 .

Os quatro blocos 4×4 , $j_i = 1:4$, são considerados seqüencialmente para subdivisão adicional, tal como mostrado na etapa 222. Para cada bloco 4×4 , é computada a variância, $v4_{ij}$, na etapa 224. Na etapa 226, primeiramente é modificado o limite de variância $T4$ para prover um novo limite $T'4$ caso o valor médio do bloco fique

entre dois valores predeterminados, a seguir a variância do bloco é comparada a este novo limite.

Caso a variância v_{4ij} não seja maior que o limite T_4 , na etapa 228, o endereço do bloco 4×4 é escrito e o correspondente bit P , P_{ij} , é ajustado para zero. O próximo bloco 4×4 é a seguir processado. Caso a variância v_{4ij} seja maior que o limite T_4 , na etapa 230, o bit P correspondente, P_{ij} , é ajustado para 1 para indicar que o bloco 4×4 deve ser subdividido em quatro blocos 2×2 . Além disso, o endereço dos quatro blocos 2×2 é escrito.

Os limites T_{16} , T_8 e T_4 podem ser constantes predeterminadas. Isto é conhecido como a decisão "hard" ou rígida. Alternativamente, uma decisão adaptável ou "soft" (suave) pode ser implementada. A decisão suave varia os limites para as variâncias dependendo do valor de pixel médio dos blocos $2N \times 2N$, em que N pode ser 8, 4 ou 2. Dessa forma, funções dos valores médios de pixel podem ser usadas como os limites.

Com o propósito de ilustração, consideremos o exemplo que se segue. Sejam os limites de variância predeterminados para o componente Y 50, 1100 e 880 para os blocos 16×16 , 8×8 e 4×4 , respectivamente. Em outras palavras, $T_{16} = 50$, $T_8 = 1100$ e $T_4 = 880$. Seja a faixa de valores médios 80 e 100. Vamos supor que a variância computada para o bloco 16×16 seja 60. Uma vez que 60 e seu valor médio 90 são maiores que T_{16} , o bloco 16×16 é subdividido em quatro sub-blocos 8×8 . Vamos supor que as variâncias computadas para os blocos 8×8 sejam 1180, 935, 980 e 1210. Uma vez que dois dos blocos 8×8 possuem variâncias que superam T_8 , esses dois blocos são adicionalmente subdivididos para produzir um total de oito sub-blocos 4×4 . Finalmente, vamos supor que as variâncias dos oito blocos 4×4 sejam 620, 630, 670, 610, 590, 525,

930 e 690, com valores médios correspondentes de 90, 120, 110, 115. Uma vez que o valor médio do primeiro bloco 4 x 4 fica na faixa (80, 100), seu limite será reduzido para $T'4 = 200$, que é menor que 880. Portanto, este bloco 4 x 4 será subdividido assim como o sétimo bloco 4 x 4. A designação de tamanho de bloco resultante é apresentada na Figura 3a. A decomposição quad-tree correspondente é apresentada na Figura 3b. Adicionalmente, os dados PQR gerados por tal designação de tamanho de bloco são apresentados na Figura 3c.

Note-se que um procedimento similar é usado para designar tamanhos de blocos para os componentes de cor C_1 e C_2 . Os componentes de cor podem ser decimalizados horizontalmente, verticalmente, ou ambos.

Adicionalmente, note-se que apesar da designação de tamanho de bloco ter sido descrita na forma de uma estratégia de cima para baixo, em que o maior bloco (16 x 16, no presente exemplo) é avaliado em primeiro lugar, poderia ser usada uma estratégia de baixo para cima. A estratégia de baixo para cima iria avaliar os menores blocos (2 x 2 no presente exemplo) em primeiro lugar.

Fazendo novamente referência à Figura 1, o restante do sistema de processamento de imagem 110 será descrito. Os dados PQR, juntamente com os endereços dos blocos selecionados, são providos a um elemento DCT 110. O elemento DCT 110 usa os dados PQR para efetuar transformadas de coseno discretas dos tamanhos apropriados nos blocos selecionados. Somente os blocos selecionados devem passar pelo processamento DCT.

O sistema de processamento de imagem 100 pode opcionalmente incluir o elemento DQT 112 para reduzir a redundância entre os coeficientes DC das DCTs. Um coeficiente DC é encontrado no canto superior esquerdo de

cada bloco DCT. Os coeficientes DC são, de um modo geral, elevados se comparados aos coeficientes AC. A discrepância nos tamanhos torna difícil projetar um codificador de comprimento variável eficiente. Assim sendo, é vantajoso
5 reduzir a redundância entre os coeficientes DC.

O elemento DQT 112 efetua DCTs 2-D sobre os coeficientes DC, tomados 2 x 2 de cada vez. Iniciando com blocos 2 x 2 dentro de blocos 4 x 4, uma DCT 2-D é efetuada sobre os quatro coeficientes DC. Tal DCT 2 x 2 é denominada
10 como a transformada quad-tree diferencial, ou DQT, dos quatro coeficientes DC. A seguir, o coeficiente DC da DQT juntamente com os três coeficientes DC vizinhos com um bloco 8 x 8 são usados para computar o próximo nível de DQT. Finalmente, os coeficientes DC dos quatro blocos 8 x 8
15 dentro de um bloco 16 x 16 são usados para computar a DQT. Dessa forma, em um bloco 16 x 16, existe um coeficiente DC verdadeiro e o restante são coeficientes AC correspondentes à DCT e DQT.

Os coeficientes de transformada (DCT e DQT) são
20 providos a um quantizador 114 para quantização. Em uma modalidade preferida, os coeficientes DCT são quantizados usando máscaras de ponderação de frequência (FWMs - Frequency Weighting Masks) e um fator de escala de quantização. Uma FWM consiste de uma tabela de pesos de
25 frequência das mesmas dimensões que o bloco de coeficientes DCT de alimentação. Os pesos de frequência aplicam diferentes pesos aos diferentes coeficientes DCT. Os pesos são determinados para enfatizar as amostras de alimentação possuindo um conteúdo de frequência que o sistema visual
30 humano é mais sensível, e para des-enfatizar as amostras possuindo conteúdo de frequências às quais o sistema visual é menos sensível. Os pesos podem ser também designados com base em fatores tais como distâncias de observação, e etc.

Os pesos são selecionados com base em dados empíricos. Um método para designar as máscaras de ponderação para coeficientes DCT 8 x 8 é descrito na ISO/IEC JTC₁ CD 10918, "DIGITAL COMPRESSION AND ENCODING OF CONTINUOUS-TONE STILL IMAGES - PART 1: REQUIREMENTS AND GUIDELINES", International Standards Organization, 1994, a qual é aqui incorporada por referência. De um modo geral, duas FWMs são designadas, uma para o componente de luminância e outra para os componentes de crominância. As tabelas de FWM para os blocos de tamanho 2 x 2 e 4 x 4 são obtidas por decimação e a 16 x 16 por interpolação daquela para o bloco 8 x 8. O fator de escala controla a qualidade e taxa de bits dos coeficientes quantizados.

Dessa forma, cada coeficiente DCT é quantizado de acordo com a relação:

$$DCT_q(i,j) = \left\lfloor \frac{8 * DCT(i,j)}{fwm(i,j) * q} \pm \frac{1}{2} \right\rfloor$$

em que $DCT(i,j)$ é o coeficiente DCT de alimentação (entrada), $fwm(i,j)$ é a máscara de ponderação de frequência, q é o fator de escala e $DCT_q(i,j)$ é o coeficiente quantizado. Note-se que, dependendo do sinal do coeficiente DCT, o primeiro termo dentro das chaves é arredondado para cima ou para baixo. Os coeficientes DQT são também quantizados usando uma máscara de ponderação adequada. No entanto, múltiplas tabelas ou máscaras podem ser usadas, e aplicadas a cada um dos componentes Y, Cb e Cr.

Os coeficientes quantizados são providos a um serializador de varredura em ziguezague 116.

O serializador 116 varre os blocos de coeficientes quantizados em forma de ziguezague para produzir um fluxo serializado de coeficientes quantizados. Vários padrões diferentes de varredura em ziguezague, como

também outros padrões diferentes de ziguezague podem ser também escolhidos. Uma técnica preferida emprega tamanhos de blocos 8 x 8 para varredura em ziguezague, apesar de outros tamanhos poderem ser usados.

5 Note-se que o serializador de varredura em ziguezague 116 pode ser posicionado antes ou após o quantizador 114. Os resultados finais são iguais.

 Em qualquer dos casos, o fluxo de coeficientes quantizados é provido a um codificador de comprimento
10 variável 118. O codificador de comprimento variável 118 pode fazer uso de codificação de run-length de zeros seguida por codificação Huffman. Tal técnica é descrita em detalhes nas Patentes US 5,021,891, 5,107,345 e 5,452,104, e é aqui resumida. Um codificador de run-length tomaria os
15 coeficientes quantizados e separaria os coeficientes iguais a zero dos diferentes de zero. Os valores de zero são designados como valores de run-length e são codificados por Huffman. Os valores diferentes de zero são codificados por Huffman separadamente.

20 Uma codificação Huffman modificada dos coeficientes quantizados é também possível e é usada na modalidade preferida. Neste caso, após a varredura em ziguezague, um codificador de run-length irá determinar os pares run-length/tamanho dentro de cada bloco 8 x 8. Tais
25 pares run-length/tamanho são a seguir codificados por Huffman.

 Os códigos Huffman são projetados a partir das estatísticas medidas ou teóricas de uma imagem. Foi observado que a maioria das imagens naturais é constituída
30 por áreas vazias ou de variação relativamente lenta, e áreas ocupadas tais como contornos de objetos e textura de alto contraste. Os codificadores Huffman com transformadas do domínio da frequência tais como a DCT exploram tais

características ao designar mais bits para as áreas ocupadas e menos bits para as áreas vazias. De um modo geral, os codificadores Huffman fazem uso de tabelas de consulta para codificar os valores de run-length e
5 diferentes de zero. Múltiplas tabelas são geralmente usadas, com três tabelas sendo preferidas na presente invenção, apesar de 1 ou 2 podem ser usados, como desejado.

O sinal de imagem comprimido gerado pelo codificador 102 é transmitido para o decodificador 106
10 através do canal de transmissão 104. Os dados PQR, que contêm as informações de designação de tamanho de bloco, são também providos ao decodificador 106. O decodificador 106 compreende um decodificador de comprimento variável 120, que decodifica os valores de run-length e os valores
15 diferentes de zero.

A saída do decodificador de comprimento variável 120 é provida a um serializador de varredura em ziguezague 122 que ordena os coeficientes de acordo com o esquema de varredura empregado. O serializador de varredura em
20 ziguezague 122 recebe os dados PQR para auxiliar na ordenação apropriada dos coeficientes em um bloco de coeficiente composto.

O bloco composto é provido a um quantizador inverso 124, para desfazer o processamento devido ao uso
25 das máscaras de ponderação de frequência.

O bloco de coeficiente é a seguir provido a um elemento IDQT 126, seguido por um elemento IDCT 128, caso a transformada quad-tree diferencial tenha sido aplicada. Caso contrário, o bloco de coeficiente é provido
30 diretamente ao elemento IDCT 128. O elemento IDQT 126 e o elemento IDCT 128 transformam inversamente os coeficientes para produzir um bloco de dados de pixel. Os dados de pixel

podem a seguir ter que serem interpolados, convertidos à forma RGB, e a seguir armazenados para futura apresentação.

Assim sendo, foram apresentados um sistema e um método para compressão de imagem que efetua a designação de tamanho de bloco com base na variância de pixels. A designação de tamanho de bloco com base na variância oferece várias vantagens. Devido ao fato de que a transformada de cosseno discreta é efetuada após serem determinados os tamanhos de blocos, se consegue uma computação eficiente. A transformada de alta demanda de recursos computacionais deve ser efetuada somente sobre os blocos selecionados. Além disso, o processo de seleção de blocos é eficiente, uma vez que a variância dos valores de pixels é matematicamente simples de calcular. Mais outra vantagem da designação de tamanho de bloco com base na variância é a de que ela é baseada em termos de percepção. A variância de pixels é uma medida da atividade em um bloco, e provê indicação da presença de bordas, texturas, etc. Ela tende a captar os detalhes de um bloco de forma muito melhor que medidas tais como a média dos valores de pixels. Dessa forma, o esquema baseado em variância da presente invenção designa blocos menores para regiões com mais bordas e blocos maiores para as regiões mais neutras. Como resultado, pode ser conseguida qualidade excepcional nas imagens reconstruídas.

Mais uma importante vantagem é a de que uma vez que a designação de tamanho de bloco é efetuada antes da quantização, uma maior flexibilidade é permitida no controle da taxa de bits e qualidade. Uma vez que o limite de variância é adaptado à média local, pequenos blocos são designados mesmo em áreas relativamente escuras. Isto preserva os detalhes em todas as áreas que estão acima do limite de visibilidade meramente notável. Ademais, a

compressão de imagem com base em variância propicia uma degradação suave da qualidade de imagem quando o fator de escala de quantização é mudado de valores baixos para altos, ao contrário de métodos tais como o MPEG. Isto é particularmente crucial para aplicações tais como na área do cinema digital.

Com a alta demanda pelo vídeo digital, a pirataria constitui séria ameaça. A possibilidade de marca d'água digital constitui importante exigência para desestimular a violação de copyrights e perda de rendas. Como a marca d'água é efetuada em áreas de uma imagem que são perceptivelmente significativas, a designação de tamanho de bloco baseada em variância é um candidato natural para se efetuar a marca d'água.

A descrição acima das modalidades preferidas é provida para permitir que os técnicos na área efetivem ou façam uso da presente invenção. As diferentes modificações a estas modalidades ficarão prontamente claras para os técnicos na área e os princípios genéricos aqui definidos podem ser aplicados a outras modalidades sem o uso das faculdades inventivas. Dessa forma, a presente invenção não deve ser limitada às modalidades aqui apresentadas, devendo receber o escopo mais amplo, consistente com os princípios e as novas características aqui descritos.

REIVINDICAÇÕES

1. Método para determinar uma designação de tamanho de bloco para um bloco de entrada de pixels de imagem a ser usada na compressão do bloco de entrada, o
5 método compreendendo as etapas de:

ler um bloco de dados de pixels;

gerar uma designação de tamanho de bloco para o bloco de dados de pixels com base nas variâncias de valores de pixels do bloco dos dados de pixels e dos blocos
10 subdivididos do bloco dos dados de pixels, a etapa de gerar compreende também as etapas de:

determinar uma variância de valores de pixels para o bloco de dados de pixels;

comparar a variância com um limite;

15 efetuar uma decisão para subdividir o bloco caso a variância seja maior que o limite;

caso a decisão seja para subdividir o bloco, então repetir as etapas de determinar, comparar, e efetuar para cada bloco subdividido até que um critério
20 predeterminado seja atendido;

designar como a designação de tamanho de bloco cada bloco que não for adicionalmente subdividido; e
prover uma estrutura de dados contendo informações sobre a designação de tamanho de bloco;

25 o método sendo **CARACTERIZADO** pelo fato de que:

na etapa de comparar a variância com um limite, o limite é definido com base na função da média dos valores de pixels sendo avaliados; e

na etapa de determinar a variância de valores de pixels para o bloco de dados de pixels, a variância é
30 calculada sobre todos os pixels do bloco.

2. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o limite muda para cada nível de subdivisão.

3. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o critério predeterminado para não mais repetir as etapas de determinar, comparar e efetuar, é baseado em um tamanho de bloco mínimo pré-selecionado de dados de pixels.

4. Sistema de compressão de imagens para comprimir um bloco de dados de pixels, o sistema compreendendo:

dispositivo para designação de tamanho de bloco para selecionar o bloco ou blocos subdivididos do bloco a ser comprimido com base em variâncias de valores de pixels do bloco de dados de pixels e dos blocos subdivididos do bloco de dados de pixels, em que o dispositivo para designação de tamanho de bloco determina uma variância de valores de pixels para o bloco de dados de pixels, compara a variância com um limite, efetua uma decisão para subdividir o bloco caso a variância seja maior que o limite, caso a decisão seja para subdividir o bloco, então repete a determinação da variância, a comparação com o limite e a decisão para subdividir para cada bloco subdividido até que um critério predeterminado seja atendido, e designa como a designação de tamanho de bloco, cada bloco que não for adicionalmente subdividido;

dispositivo transformador para transformar dados de pixels do bloco ou blocos subdivididos selecionados em dados no domínio da frequência;

dispositivo quantizador para quantizar os dados do domínio da frequência;

dispositivo serializador para varrer os dados quantizados em um fluxo de dados em série; e

dispositivo codificador de comprimento variável para codificar o fluxo de dados em série em preparação para transmissão;

o sistema sendo **CARACTERIZADO** pelo fato de que o dispositivo para designação de tamanho de bloco compreende:

dispositivo para comparar a variância com um limite, o limite sendo definido com base na função da média
5 dos valores de pixels do bloco sendo avaliados; e

dispositivo para determinar a variância de pixels, a variância sendo calculada sobre todos os pixels do bloco de dados de pixels.

5. Sistema, de acordo com a reivindicação 4,
10 **CARACTERIZADO** pelo fato de que o limite muda para cada nível de subdivisão.

6. Sistema, de acordo com a reivindicação 4,
CARACTERIZADO pelo fato de que o critério predeterminado para não mais subdividir é baseado em um tamanho de bloco
15 mínimo pré-selecionado de dados de pixels sendo atingido.

7. Sistema, de acordo com a reivindicação 4,
CARACTERIZADO pelo fato de que o dispositivo transformador efetua uma transformada de coseno discreta.

8. Sistema, de acordo com a reivindicação 4,
20 **CARACTERIZADO** pelo fato de que o dispositivo transformador efetua uma transformada de coseno discreta seguida por uma transformada diferencial quad-tree.

9. Sistema, de acordo com a reivindicação 4,
CARACTERIZADO pelo fato de que o dispositivo serializador
25 compreende um scanner ziguezague.

10. Sistema, de acordo com a reivindicação 9,
CARACTERIZADO pelo fato de que o scanner ziguezague emprega um tamanho de bloco de 8 x 8 para a varredura ziguezague.

11. Sistema, de acordo com a reivindicação 4,
30 **CARACTERIZADO** pelo fato de que em que o dispositivo codificador de comprimento variável compreende um codificador Huffman.

12. Sistema, de acordo com a reivindicação 11,
CARACTERIZADO pelo fato de que o codificador Huffman

emprega múltiplas tabelas de consulta para codificar valores de comprimento corridos e diferentes de zero.

13. Sistema, de acordo com a reivindicação 12, **CARACTERIZADO** pelo fato de que existem três tabelas de
5 consulta.

14. Método para compressão de um bloco de dados de pixels de uma imagem, o método compreendendo as etapas de:

ler um bloco de dados de pixels;
10 gerar uma designação de tamanho de bloco para o bloco de dados de pixel com base nas variâncias dos valores de pixels do bloco de dados de pixels e dos blocos subdivididos do bloco de dados de pixels, a etapa de gerar compreendendo também as etapas de:

15 determinar uma variância de valores de pixels para o bloco de dados de pixels;

comparar a variância com um limite;

efetuar uma decisão para subdividir o bloco caso a variância seja maior que o limite;

20 caso a decisão seja para subdividir o bloco, então repetir as etapas de determinar, comparar, e efetuar para cada bloco subdividido até que um ou mais dos critérios predeterminados seja atendido; e

designar, como a designação de tamanho de
25 bloco, cada bloco que não for adicionalmente subdividido;

prover uma estrutura de dados contendo informações sobre a designação de tamanho de bloco;

transformar os dados de pixels de blocos selecionados como indicado pela estrutura de dados em uma
30 representação no domínio da frequência;

quantizar os dados no domínio da frequência com base nas características da percepção humana da imagem;

varrer os dados quantizados em um fluxo de dados em série; e

codificar o fluxo de dados em série em preparação para transmissão;

o método sendo **CARACTERIZADO** pelo fato de que:

na etapa de comparar a variância com o limite, o
5 limite é definido com base na função da média dos valores de pixels sendo avaliados; e

na etapa de determinar a variância de valores de pixels para o bloco de dados de pixels, a variância é calculada sobre todos os pixels do bloco.

10 15. Método, de acordo com a reivindicação 14, **CARACTERIZADO** pelo fato de que o limite muda para cada nível de subdivisão.

16. Método, de acordo com a reivindicação 14, **CARACTERIZADO** pelo fato de que o critério predeterminado
15 para não mais repetir as etapas de determinar, comparar e efetuar, está baseado em um tamanho de bloco mínimo pré-selecionado de dados de pixels.

17. Método, de acordo com a reivindicação 14, **CARACTERIZADO** pelo fato de que uma transformada de coseno
20 discreta é efetuada durante a etapa de transformar.

18. Método, de acordo com a reivindicação 14, **CARACTERIZADO** pelo fato de que é efetuada uma transformada de coseno discreta seguida por uma transformada diferencial quad-tree durante a etapa de transformar.

25 19. Método, de acordo com a reivindicação 14, **CARACTERIZADO** pelo fato de que é efetuada varredura ziguezague durante a etapa de varrer.

20. Método, de acordo com a reivindicação 19, **CARACTERIZADO** pelo fato de que a varredura ziguezague é
30 efetuada usando um tamanho de bloco 8 x 8.

21. Método, de acordo com a reivindicação 14, **CARACTERIZADO** pelo fato de que uma codificação Huffman é efetuada durante a etapa de codificar.

22. Método, de acordo com a reivindicação 21, **CARACTERIZADO** pelo fato de que a codificação Huffman emprega múltiplas tabelas de consulta para codificar valores de comprimento corridos e diferentes de zero.

5 23. Método, de acordo com a reivindicação 22, **CARACTERIZADO** pelo fato de que existem três tabelas de consulta.

24. Sistema para compressão de um bloco de dados de pixels de uma imagem, o sistema compreendendo:

10 dispositivo de leitura para ler um bloco de dados de pixels;

dispositivo gerador para gerar uma designação de tamanho de bloco para o bloco de dados de pixels com base nas variâncias dos valores de pixels do bloco de dados de pixels e dos blocos subdivididos do bloco de dados de pixels, o dispositivo gerador compreende também:

dispositivo de determinação para determinar uma variância de valores de pixels para o bloco de dados de pixels;

20 dispositivo de comparação para comparar as variâncias com um limite;

dispositivo de efetuar decisão para efetuar uma decisão para subdividir o bloco caso a variância seja maior que o limite;

25 dispositivo de repetição para, caso o dispositivo de efetuar decisão decida subdividir o bloco, então repetir as etapas de determinar, comparar, e efetuar a decisão para cada bloco subdividido até que um critério predeterminado seja atendido; e

30 dispositivo de designação para designar, como a designação de tamanho de bloco, cada bloco que não for adicionalmente subdividido;

dispositivo provedor para prover uma estrutura de dados contendo informação sobre a designação de tamanho de bloco;

5 dispositivo transformador para transformar os dados de pixels de blocos selecionados como indicado pela estrutura de dados em uma representação no domínio da frequência;

10 dispositivo quantizador para quantizar os dados no domínio da frequência com base nas características da percepção humana da imagem;

dispositivo serializador para varrer os dados quantizados em um fluxo de dados em série; e

dispositivo de codificação para codificar o fluxo de dados em série em preparação para transmissão;

15 o sistema sendo **CARACTERIZADO** pelo fato de que:

quando o dispositivo de comparação compara a variância com um limite, o limite é definido com base na função da média dos valores de pixels sendo avaliados; e

20 quando o dispositivo de determinação determina a variância de pixels, a variância é calculada sobre todos os pixels do bloco de dados de pixels.

25. Sistema, de acordo com a reivindicação 24, **CARACTERIZADO** pelo fato de que o limite muda para cada nível de subdivisão.

25 26. Sistema, de acordo com a reivindicação 24, **CARACTERIZADO** pelo fato de que o critério predeterminado é baseado em um tamanho de bloco mínimo pré-selecionado de dados de pixels.

30 27. Sistema, de acordo com a reivindicação 24, **CARACTERIZADO** pelo fato de que o dispositivo de transformação utiliza a transformada de cosseno discreta.

28. Sistema, de acordo com a reivindicação 27, **CARACTERIZADO** pelo fato de que o dispositivo de

transformação utiliza uma transformada de cosseno discreta seguida por uma transformada diferencial quad-tree.

29. Sistema, de acordo com a reivindicação 24,
5 **CARACTERIZADO** pelo fato de que o dispositivo serializador de varredura utiliza varredura ziguezague.

30. Sistema, de acordo com a reivindicação 24,
CARACTERIZADO pelo fato de que o dispositivo de codificação utiliza codificação Huffman.

31. Sistema, de acordo com a reivindicação 30,
10 **CARACTERIZADO** pelo fato de que a codificação Huffman utiliza múltiplas tabelas de consulta para codificar valores de comprimento corridos e diferentes de zero.

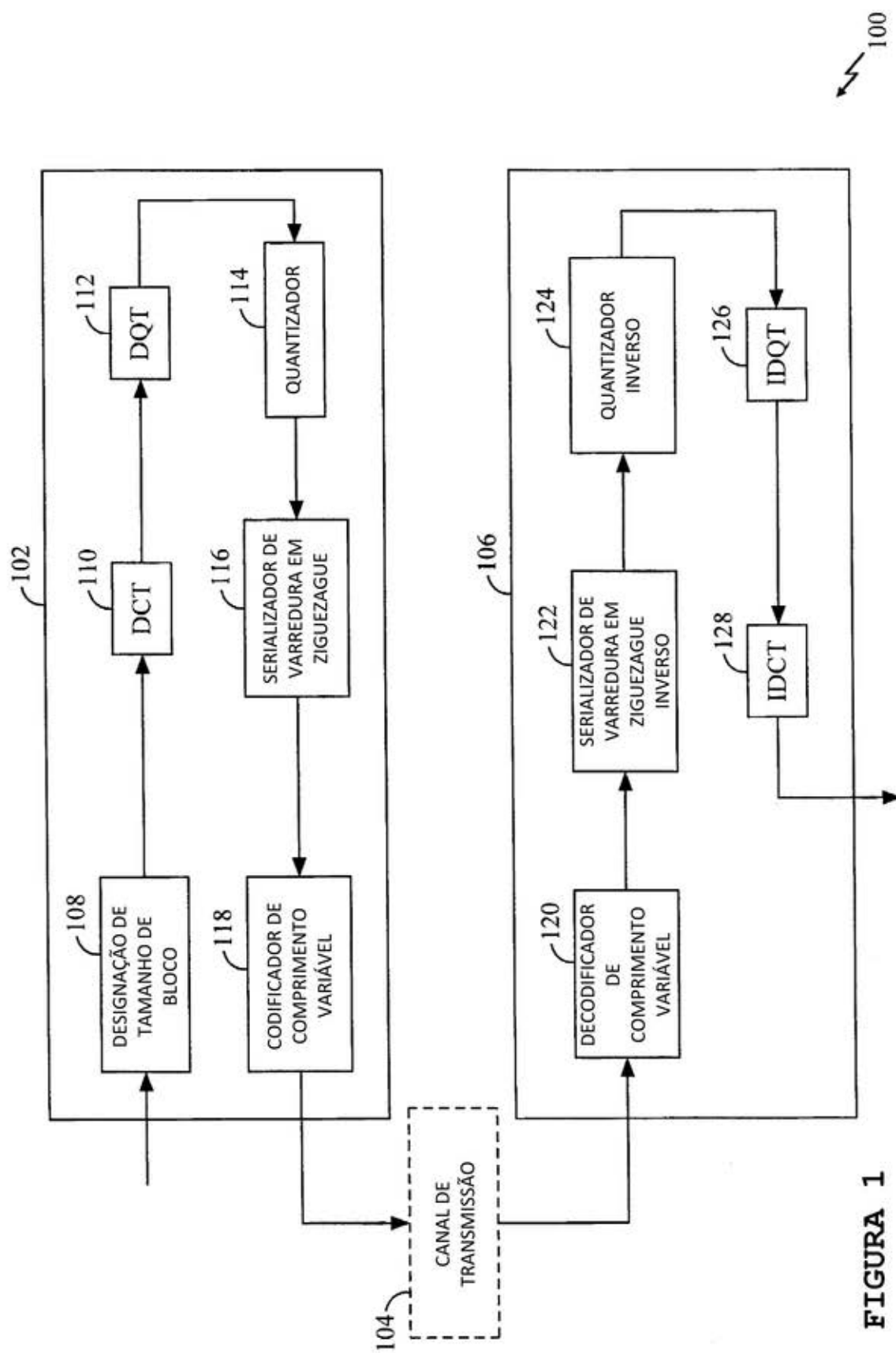


FIGURA 1

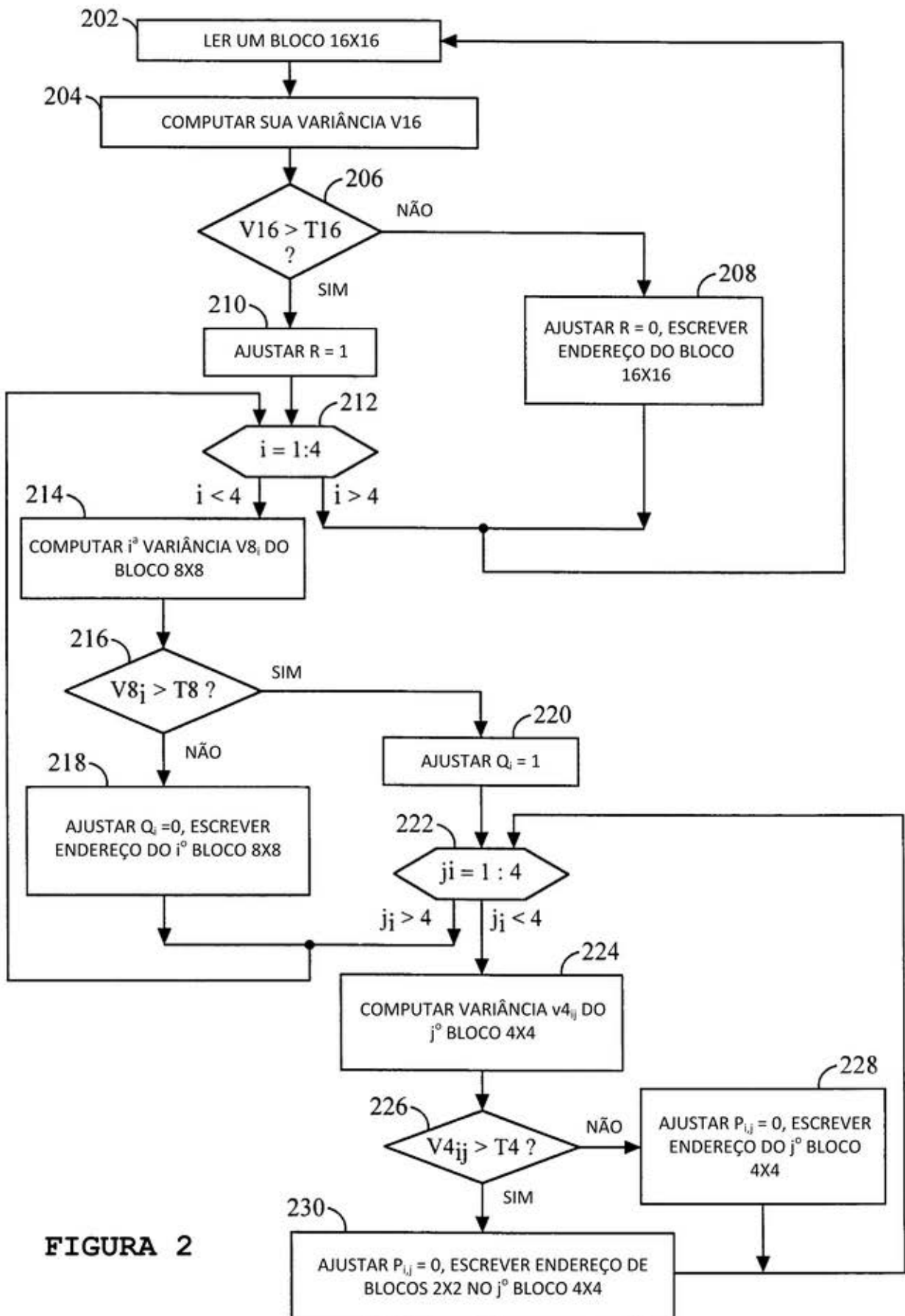


FIGURA 2

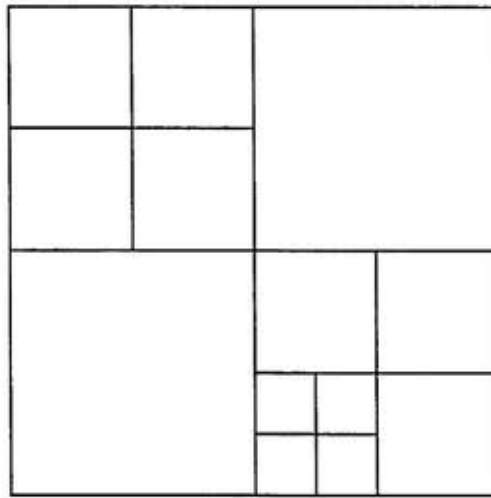


FIGURA 3A

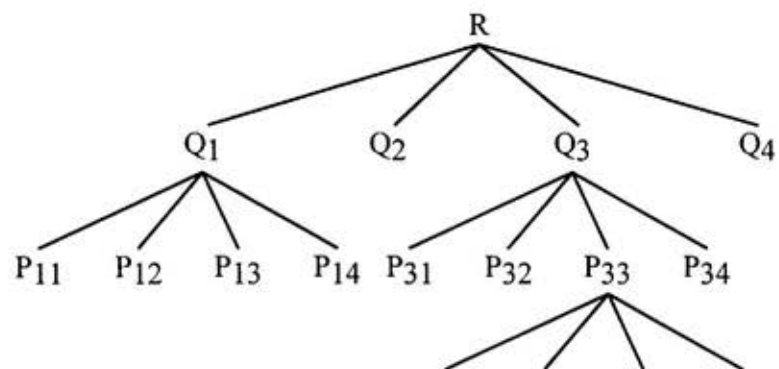


FIGURA 3B

1	1	0	1	0	0	0	0	0	0	0	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---

DADOS PQR

FIGURA 3C

RESUMO

"MÉTODO PARA DETERMINAR UMA DESIGNAÇÃO DE TAMANHO DE BLOCO PARA UM BLOCO DE ENTRADA DE PIXELS DE IMAGEM A SER USADA NA COMPRESSÃO DO BLOCO DE ENTRADA, SISTEMA DE COMPRESSÃO DE IMAGENS
5 PARA COMPRIMIR UM BLOCO DE DADOS DE PIXELS, E, MÉTODO E SISTEMA PARA COMPRESSÃO DE UM BLOCO DE DADOS DE PIXELS DE UMA IMAGEM."

Sistema e método para compressão de imagem utilizando blocos e sub-blocos de dados de coeficiente de transformada de coseno discreta dimensionados de forma adaptável são descritos. Um elemento designador de tamanho
10 de bloco no codificador seleciona o bloco ou sub-bloco de um bloco de alimentação de pixels a ser processado. A seleção é baseada na variância dos valores de pixels. Os blocos com variâncias maiores que um limite são subdivididos, enquanto blocos com variâncias menores que um
15 limite não são subdivididos. Um elemento de transformada transforma os valores de pixel dos blocos selecionados no domínio da frequência. Os valores do domínio da frequência podem então ser quantizados, serializados, e codificados
20 com comprimentos variáveis em preparação para a transmissão.