



República Federativa do Brasil
Ministério da Economia
Instituto Nacional da Propriedade Industrial

(11) PI 0807189-6 B1



(22) Data do Depósito: 20/02/2008

(45) Data de Concessão: 26/05/2020

(54) Título: MÉTODO, MEIO DE ARMAZENAMENTO E SISTEMA DE PROCESSAMENTO DE MÍDIA DIGITAL

(51) Int.Cl.: H04N 19/186; H04N 19/46; H04N 19/63; H04N 19/70.

(52) CPC: H04N 19/186; H04N 19/46; H04N 19/63; H04N 19/70.

(30) Prioridade Unionista: 21/02/2007 US 60/891.030; 29/06/2007 US 11/772.070.

(73) Titular(es): MICROSOFT TECHNOLOGY LICENSING, LLC.

(72) Inventor(es): SRIDHAR SRINIVASAN; TU, CHENGJIE; REGUNATHAN, SHANKAR; SULLIVAN, GARY J.

(86) Pedido PCT: PCT US2008054481 de 20/02/2008

(87) Publicação PCT: WO 2008/103774 de 28/08/2008

(85) Data do Início da Fase Nacional: 29/07/2009

(57) Resumo: "SINALIZAÇÃO E USO DE INFORMAÇÕES DE POSICIONAMENTO DE ' AMOSTRA DE CRGMINÂNCIA" São descritas regras para a sinalização e interpretação da posição de crominância. Uma regra, chamada de "regra curta", define quinze posições de centralização de crominância distintas e um elemento de sintaxe de quatro bits correspondente. Outra regra, chamada de "regra estendida", define 81 posições de centralização de crominância distintas e elementos de sintaxe de sete bits correspondentes. O método descrito inclui receber dados de mídia digital em um codificador de mídia digital, determinar informações de posição de crominância para os dados de mídia digital recebidos, e representar as informações de posição de crominância com um ou mais elementos de sintaxe em um fluxo de bits codificado. O um ou mais elementos de sintaxe são operáveis para comunicar as informações de posição de crominância a um decodificador de mídia digital. As informações de posição de crominância facilitam a rotação ou inversão da imagem.

Relatório Descritivo da Patente de Invenção para
**"MÉTODO, MEIO DE ARMAZENAMENTO E SISTEMA DE
PROCESSAMENTO DE MÍDIA DIGITAL"**

SUMÁRIO

[0001] No sumário, a descrição detalhada está direcionada a aspectos de codificação e decodificação de dados de mídia digital, e em particular, à codificação e decodificação de dados de mídia digital em codificadores e decodificadores de mídia digital.

[0002] Por exemplo, são descritas regras para a sinalização e interpretação da posição de crominância. Uma regra, chamada de "regra curta", define quinze posições de centralização de crominância distintas e um elemento de sintaxe de quatro bits correspondente. Outra regra, chamada de "regra estendida", define 81 posições de centralização de crominância distintas e elementos de sintaxe de sete bits correspondentes. Também são descritas variações dessas regras e outros aspectos.

[0003] Em um aspecto, um método compreende receber dados de mídia digital em um codificador de mídia digital; determinar informações de posição de crominância para os dados de mídia digital recebidos; representar as informações de posição de crominância com um ou mais elementos de sintaxe em um fluxo de bits codificado, sendo que o elemento de sintaxe é operável para comunicar as informações de posição de crominância a um decodificador de mídia digital, e sendo que as informações de posição de crominância facilitam uma rotação ou inversão da imagem; e gerar o fluxo de bits codificado. Em outro aspecto, as informações de posição de crominância são decodificadas.

[0004] A intenção deste sumário é a de apresentar, de maneira simplificada, uma seleção de conceitos, descritos em detalhes adiante na descrição detalhada. O presente sumário não tem a intenção de

identificar aspectos cruciais ou essenciais da matéria reivindicada e não deve ser usado para limitar o âmbito da matéria reivindicada.

[0005] O que foi dito acima e outros objetivos, aspectos e vantagens ficarão mais evidentes na descrição detalhada seguinte, que prossegue com referência às figuras em anexo.

BREVE DESCRIÇÃO DOS DESENHOS

[0006] A Figura 1 é um diagrama de blocos de um ambiente de computação adequado para implementar as técnicas e ferramentas para sinalização e uso das informações de posição de crominância em uma ou mais das implementações descritas.

[0007] A Figura 2 é um diagrama que mostra uma primeira regra de posição de amostra de crominância exemplificativa.

[0008] A Figura 3 é um diagrama que mostra uma segunda regra de posição de amostra de crominância exemplificativa.

[0009] A Figura 4 é um diagrama que mostra a redução de crominância dos dados de entrelaçamento.

[00010] A Figura 5 é um diagrama de blocos de um codec baseado em transformação de blocos. A Figura 6 é um diagrama de fluxo de dados de um codificador representativo. A Figura 7 é um diagrama de fluxo de dados de um decodificador representativo.

DESCRIÇÃO DETALHADA

[00011] O presente pedido se refere a técnicas e ferramentas para compressão e descompressão eficiente de dados de mídia digital. Nas várias concretizações descritas, um codificador e/ou decodificador de imagens fixas incorpora técnicas para comprimir e/ou descomprimir dados de imagem.

[00012] Várias alternativas às implementações aqui descritas são possíveis. Por exemplo, as técnicas descritas com referência aos diagramas de fluxo de dados podem ser alteradas mediante a modificação da ordenação dos estágios apresentados nos fluxogramas, a repe-

tição ou omissão de certos estágios, etc. Como outro exemplo, embora algumas implementações sejam descritas com referência a formatos de mídia digital específicos, outros formatos também podem ser usados.

[00013] As várias técnicas e ferramentas podem ser usadas em combinação ou separadamente. Diferentes concretizações implementam uma ou mais das técnicas e ferramentas descritas. Algumas técnicas e ferramentas descritas aqui podem ser usadas em um codificador ou decodificador de imagens fixa, ou em algum outro sistema não limitado especificamente à codificação ou decodificação de imagens fixas.

I. Ambiente de Computação.

[00014] A Figura 1 ilustra exemplo generalizado de um ambiente de computação adequado 100 no qual várias das concretizações descritas podem ser implementadas. O ambiente de computação 100 não tem a intenção de implicar qualquer limitação quanto ao escopo de uso ou funcionalidade, uma vez que as técnicas e ferramentas podem ser implementadas em diversos ambientes de computação de uso geral ou especial.

[00015] Com referência à Figura 1, o ambiente de computação 100 inclui pelo menos uma unidade de processamento 110 e memória 120. Na Figura 1, essa configuração mais básica 130 está incluída dentro de uma linha tracejada. A unidade de processamento 110 executa instruções executáveis por computador e pode ser um processador real ou virtual. Em um sistema de multiprocessamento, várias unidades de processamento executam instruções executáveis por computador para aumentar a capacidade de processamento. A memória 120 pode ser memória volátil (por exemplo, registros, cache, RAM), memória não-volátil (por exemplo, ROM, EEPROM, memória flash, etc.), ou alguma combinação das duas. A memória 120 armazena o software 180 implementando um codificador ou decodificador de mídia digital com uma

ou mais das técnicas e ferramentas descritas.

[00016] Um ambiente de computação pode ter recursos adicionais. Por exemplo, o ambiente de computação 100 inclui o meio de armazenamento 140, um ou mais dispositivos de entrada 150, um ou mais dispositivos de saída 160 e uma ou mais conexões de comunicação 170. Um mecanismo de interconexão (não ilustrado), tal como um barramento, controlador ou rede interconecta os componentes do ambiente de computação 100. Normalmente, o software do sistema operacional (não ilustrado) oferece um ambiente operacional para outros softwares em execução no ambiente de computação 100, e coordena as atividades dos componentes do ambiente de computação 100.

[00017] O armazenamento 140 pode ser removível ou não-removível, e pode incluir discos magnéticos, fitas ou cassetes magnéticos, CD-ROMs, DVDs (inclusive DVDs de alta definição) ou qualquer outro meio que possa ser usado para armazenar informações e que possa ser acessado dentro do ambiente de computação 100. O armazenamento 140 armazena instruções para o software 180 implementando o codificador ou decodificador de mídia digital.

[00018] O(s) dispositivo(s) de entrada 150 pode(m) ser um dispositivo de entrada por toque, tal como um teclado, mouse, caneta ou trackball, um dispositivo de entrada de voz, um dispositivo de digitalização, dispositivo de captura de imagem fixa (por exemplo, uma câmera digital), ou outro dispositivo que alimente informações de entrada ao ambiente de computação 100. Para codificação de áudio ou vídeo, o(s) dispositivo(s) de entrada 150 pode(m) ser uma placa de som, placa de vídeo, placa sintonizadora de TV ou dispositivo similar que aceite entrada de áudio ou vídeo na forma analógica ou digital, ou um CD-ROM ou CD-RW que lê amostras de áudio ou vídeo para o ambiente de computação 100. O(s) dispositivo(s) de saída 160 pode(m) ser um meio de exibição, impressora, alto-falante, gravador de CD/DVD, ou

outro dispositivo que proporcione uma saída para o ambiente de computação 100.

[00019] As conexões de comunicação 170 permitem comunicação por um meio de comunicação para outra entidade de computação. O meio de comunicação transmite informações, como instruções executáveis por computador, entrada ou saída de mídia digital, ou outros dados em um sinal de dados modulado. Um sinal de dados modulado é um sinal que tem uma ou mais de suas características ajustadas ou alteradas de tal maneira a codificar as informações no sinal. A título exemplificativo, sem limitação, os meios de comunicação incluem técnicas com ou sem fio implementadas com um meio elétrico, óptico, RF, infravermelho, acústico ou outra portadora.

[00020] As técnicas e ferramentas podem ser descritas no contexto geral de meios legíveis por computador. Os meios legíveis por computador são qualquer meio disponível que pode ser acessado dentro de um ambiente de computação. A título de exemplo, sem limitação, com o ambiente de computação 100, os meios legíveis por computador incluem a memória 120, o armazenamento 140, meios de comunicações e combinações de qualquer um dos citados.

[00021] As técnicas e ferramentas podem ser descritas no contexto geral de instruções executáveis por computador, tais como as incluídas nos módulos de programa, sendo executadas em um ambiente de computação em um processador real ou virtual de destino. Geralmente, os módulos de programa incluem rotinas, programas, bibliotecas, objetos, classes, componentes, estruturas de dados etc., que efetuam tarefas específicas ou implementam tipos de dados abstratos específicos. A funcionalidade dos módulos de programa pode ser combinada ou dividida entre os módulos de programa conforme desejado nas várias concretizações. As instruções executáveis por computador para módulos de programa podem ser executadas dentro de um ambiente

de computação local ou distribuído.

[00022] Para fins de apresentação, a descrição detalhada usa termos como "selecionar" e "receber" para descrever operações de computador em um ambiente de computação. Esses termos são abstrações de nível superior para as operações realizadas por um computador, e não devem ser confundidos com os atos realizados por um ser humano. As operações reais do computador que correspondem a esses termos variam, dependendo da implementação.

II. Sinalização e Uso das Informações de Posição da Amostra de Crominância

[00023] As concretizações descritas proporcionam recursos de fluxo de bits de codificador de imagens fixas avançados, inclusive a possibilidade de efetuar rotações cardinais e inversões espelhadas sem uma decodificação e recodificação completa. Esse recurso pode ser suportado por várias técnicas estruturais, como por exemplo:

1. Transformada Biortogonal Sobreposta (LBT) -

a. A simetria das funções de base da LBT permite uma inversão espelhada dos dados especiais dentro do bloco de transformação pela simples negação do sinal dos coeficientes da transformada de simetria ímpar. Isso é verdadeiro para ambas as orientações espaciais, X e Y.

b. A natureza isotrópica das funções de base da LBT permite que os dados espaciais dentro do bloco de transformação sejam transpostos pela simples transposição dos coeficientes da transformada. As posições cardinais podem ser implementadas como combinações de transposição e inversões por efeito espelho.

2. Hierarquias espaciais de blocos, macroblocos e microblocos

a. De modo a realizar uma inversão espelhada dentro de um macrobloco de dados, os blocos de transformada modificados são

varridos na sequência invertida lateralmente (em X e/ou Y, dependendo da necessidade). De modo similar, dentro de um segmento, os macroblocos modificados são varridos na ordem invertida lateralmente, e dentro de uma imagem, os segmentos modificados são varridos na ordem invertida lateralmente.

b. De modo a realizar uma transposição, os blocos, macroblocos e microblocos modificados são transpostos. As posições cardinais podem ser implementadas como combinações de transposição e inversões por efeito espelho.

3. Sinalização de uma área inscrita dentro de uma área de recorte estendida – isso permite que imagens alinhadas sem macroblocos sejam invertidas por efeito espelho ou giradas livremente e o desvio da imagem da grade de macroblocos seja permitido em qualquer direção, e não apenas para a direita e para baixo.

4. Sinalização da posição da amostra de croma – Isso permite que formatos de cores subamostrados de croma, tal como YUV4:2:0 e YUV4:2:2, sejam girados permitindo a especificação independente da localização da amostra de croma. Ela permite ainda que os alinhamentos relativos das posições das amostras de luminância / croma sejam sinalizados ao decodificados, de modo que um filtro de sobreamostragem (upsample) com a fase apropriada possa ser escolhido.

[00024] A sinalização das posições das amostras de croma é abordada em detalhes a seguir. As técnicas de sinalização descritas permitem que as imagens sejam giradas dentro do domínio comprimido sem nenhuma perda de informações e nenhuma alteração significativa no tamanho comprimido. Esse é um recurso de fluxo de bits desejável e introduz benefícios de complexidade.

A. Centralização da Croma

[00025] Uma imagem consiste de vários planos de dados. No espa-

ção primário, uma imagem é geralmente formada de 3 planos de cores correspondendo, respectivamente, aos canais Vermelho, Verde e Azul (R, G e B). No espaço de cores interno usado na maioria dos codecs populares, uma imagem é constituída de 3 planos de cores convertidos, geralmente chamados de Y, U e V. O componente Y é chamado de plano luma ou luminância, que corresponde aproximadamente a uma representação da imagem em escala de cinza. Os componentes U e V são chamados de croma, crominância ou planos de diferença de cor. A nomenclatura Y, U, V é usada aqui num sentido genérico, sendo entendido que as técnicas e ferramentas descritas são aplicáveis a uma variedade de formatos de cores do tipo "YUV", tal como YCbCr, YCoCg, etc. Um formato de cor chamado de YUV 4:4:4 tem uma amostra U e uma amostra V para cada amostra Y.

[00026] O olho humano é muito sensível à variação de intensidade e à resolução do canal de luminância. Ele é relativamente menos sensível à crominância. Isso possibilita uma maneira simples de reduzir a contagem de pixels nos dados pela redução ou diminuição da resolução dos componentes de crominância (U e V).

[00027] Duas técnicas de sub-amostragem de crominância são populares:

1. YUV 4:2:2 – aqui, a resolução espacial de U e V na direção X é reduzida por um fator de 2 (geralmente com algum filtro de suavização).

2. YUV 4:2:0 – nesta, a resolução espacial de U e V em ambas as direções X e Y é reduzida por um fator de 2.

[00028] Para o caso da YUV 4:2:2, cada amostra de crominância corresponde a duas amostras de luminância. De forma similar, para o caso da YUV 4:2:0, cada amostra de crominância corresponde a quatro amostras de luminância. A subamostragem da crominância é geralmente realizada após a filtragem das amostras com um filtro de su-

avização.

[00029] A fase desse filtro determina a posição relativa das amostras de croma e luminância.

[00030] Durante a conversão de qualquer desses formatos de volta ao YUV 4:4:4 com fins de exibição ou impressão, as informações sobre as posições relativas da amostra devem estar disponíveis de modo que o filtro de sobreamostragem apropriado possa ser utilizado.

[00031] Uma solução para esse problema consiste em direcionar ou sinalizar o filtro de sobreamostragem exato que deve ser utilizado. No entanto, essa abordagem impõe exigências adicionais sobre o sistema e pode não ser compatível com o restante da indústria.

[00032] Uma solução mais simples e flexível para indicar como reconstruir os dados em resolução total a partir de uma versão subamostrada consiste em sinalizar informações de "posição" com respeito ao alinhamento das amostras de croma e luminância. Essa abordagem permite ao codificador utilizar qualquer filtro de sobreamostragem cuja fase seja correspondida às informações de posição.

[00033] Embora essa abordagem não especifique uma regra de reconstrução única (isto é, filtro de sobreamostragem único), ela tem um desempenho suficientemente satisfatório e ampla aceitação.

[00034] A "posição" de um ponto de dado subamostrado é a localização ou fase desse valor dentro de uma grade de resolução total. As informações de posição são usadas para escolher entre os filtros de sobreamostragem que são compatíveis com a restrição de fase. As informações de posição são bidimensionais em geral – um desvio é especificado tanto na direção horizontal quanto vertical.

[00035] As Figuras 2 e 4 mostram exemplos de duas regras de posição de croma comuns usadas para YUV 4:2:0. Na Figura 2, fase = (0,0), e, na Figura 3, fase = (0,5, 0,5) em unidades de pixel de luminância.

B. Centralização da cromaância com rotação/inversões de imagem

[00036] Os dois exemplos apresentados nas Figuras 2 e 3 são os casos mais comuns para a subamostragem em YUV 4:2:0 da cromaância. Essas duas regras de centralização geralmente são suficientes para dados de vídeo, mas normalmente são insuficientes para dados de imagem. Uma diferença entre o vídeo e as imagens é que o vídeo raramente é girado ou invertido por efeito espelho, ao passo que as imagens são giradas e/ou invertidas por efeito espelho com muita frequência.

[00037] Para entender porque as duas regras de centralização são geralmente insuficientes para dados de imagem, considere os seguintes casos.

[00038] Caso 1: Considere uma inversão por efeito espelho ao longo da direção horizontal para o exemplo de centralização 1. Agora, a amostra de cromaância está co-localizada não com a amostra de luminância superior-esquerda, mas sim com a amostra de luminância superior-direita. A fase correspondente da cromaância é (1,0) em unidades de pixel de luminância, o que não é definido pelas regras apresentadas nas Figuras 2 e 3.

[00039] Caso 2: De modo similar, uma inversão espelhada ao longo da direção vertical de uma imagem com a posição de cromaância apresentada no exemplo 1 resulta em uma posição de cromaância com fase (0,1) em unidades de pixel de luminância, o que não é definido pelas regras apresentadas nas Figuras 2 e 3.

[00040] Os casos anteriores mostram a utilidade de se definir regras de centralização de cromaância adicionais como informações laterais a um fluxo de bits para auxiliar no processo de reconstrução correta quando a imagem for submetida às operações básicas de rotações cardinais e inversões espelhadas.

C. Centralização da crominância com dados de entrelaçamento

[00041] Outra complicação é introduzida pelo vídeo entrelaçado. Um quadro de vídeo entrelaçado contém dois campos – o campo superior e o campo inferior. Um campo de vídeo pode ser armazenado em sua resolução total, sem nenhuma subamostragem de crominância. Mais geralmente, ele é transportado em uma forma com crominância subamostrada, tal como YUV 4:2:2, em que a crominância é subamostrada na direção X por um fator de 2, e corresponde à resolução de luminância na direção Y. Nos codecs de vídeo recentes, um campo de vídeo entrelaçado é definido no espaço YUV 4:2:0, de modo que sua crominância seja subamostrada por um fator de 2 em ambas as direções, X e Y.

[00042] Essa operação geralmente resulta em uma centralização de crominância com um desvio de fase de 0,25 ou 0,75 (em unidades de pixel de luminância) na direção vertical, dependendo de se esse dado é um dado de campo superior ou inferior, respectivamente. Tal centralização pode ser usada para assegurar o seguinte:

1. Que linhas alternadas de crominância no quadro sejam produzidas por campos alternados.

[00043] Que a centralização da crominância seja uniforme ao longo das linhas sucessivas do quadro.

[00044] A subamostragem da crominância dos dados de entrelaçamento é ilustrada na Figura 4. A subamostragem no eixo geométrico X pode ter qualquer fase e não é relevante para esta discussão. Portanto, a figura ilustra apenas a centralização e os deslocamentos no eixo geométrico Y.

D. Posições de crominância

[00045] Tendo em mente o que foi dito anteriormente, definidos duas regras para a posição de crominância. A primeira regra, chamada

de regra curta, define 15 fases de centralização de cromaticidade. Essa regra é sinalizada usando uma palavra de 4 bits dentro de um fluxo de bits de imagem. A Tabela 1 enumera os valores e fases correspondentes do elemento de sintaxe CHROMA_CENTERING_SHORT em uma implementação. No exemplo apresentado na Tabela 1, CHROMA_CENTERING_SHORT pode assumir valores entre 0 e 15, mas o valor 14 é reservado e não é usado.

[00046] O CHROMA_CENTERING_SHORT pode ser sinalizado, por exemplo, em um cabeçalho de imagem ou em um cabeçalho de plano de imagem.

CHROMA_CENTERING_SHORT	Fase X = 0	Fase X = 0,5	Fase X = 1
Fase Y = 0	0	10	1
Fase Y = 0,25	6	12	7
Fase Y = 0,5	4	15	5
Fase Y = 0,75	8	13	9
Fase Y = 1	2	11	3

Tabela 1: Enumeração dos valores de CHROMACENTERINGSHORT e das fases correspondentes.

[00047] Uma segunda regra de centralização de cromaticidade mais abrangente, chamada de regra *estendida*, também é descrita. Essa regra tem a vantagem de permitir que uma imagem seja traduzida, seja perda de dados, por qualquer número inteiro de pixels. Isso serve, além disso, para giros e inversões por efeito espelho.

[00048] Em uma implementação, a regra estendida é sinalizada com uma palavra de sete bits (CHROMA_CENTERING_LONG) dentro do fluxo de bits da imagem, e a enumeração das fases correspondendo ao elemento de sintaxe CHROMA_CENTERING_LONG se dá como segue.

[00049] $\text{CHROMA_CENTERING_LONG} = \text{CHROMA_CENTERING_X} + \text{CHROMA_CENTERING_Y} * 9$, em que CHROMA_CENTERING_X e

CHROMA_CENTERING_Y são elementos de sintaxe definindo a fase nas direções X e Y, como apresentado adiante na Tabela 2. CHROMA_CENTERING_X e CHROMA_CENTERING_Y assumem valores entre 0 e 8. Portanto, CHROMA_CENTERING_LONG pode assumir valores entre 0 e 80. Os valores fora dessa faixa são reservados. CHROMA_CENTERING_LONG, CHROMA_CENTERING_X e/ou CHROMA_CENTERING_Y podem ser sinalizados, por exemplo, em um cabeçalho de imagem ou em um cabeçalho de plano de imagem.

CHROMA_CENTERING_X ou Y	Fase X ou Y
8	-0,5
7	-0,25
0	0
1	0,25
2	0,5
3	0,75
4	1,0
5	1,25
6	1,5

Tabela 2: Enumeração dos valores de CHROMA_CENTERING_X e CHROMA_CENTERING_Y e da fase correspondente

[00050] É possível usar outros mapeamentos no lugar das Tabelas 1 e 2. Também é possível usar outras codificações dos elementos CHROMA_CENTERING, tais como códigos de comprimento variável.

III. Codificação Baseada em Transformada de Bloco

[00051] A codificação de transformada é uma técnica de compressão usada em diversos sistemas de compressão de mídia digital (por exemplo, áudio, imagem e vídeo). As imagens e vídeos digitais não comprimidos são geralmente representados ou capturados como amostras de elementos de imagem ou cores em localizações em um

quadro de imagem ou vídeo disposto em uma grade bidimensional (2D). Isso é chamado de representação no domínio espacial da imagem ou vídeo. Por exemplo, um típico formato para imagens consiste de um fluxo de amostras de elementos de imagem colorida de 24 bits dispostas como uma grade. Cada amostra é um número que representa componentes de cor em uma localização de pixel na grade dentro de um espaço de cor, tal como RGB ou YIQ, dentre outros. Vários sistemas de vídeo e imagem podem usar diferentes resoluções de cores, espaciais e de tempo na amostragem. De forma similar, o áudio digital é normalmente representado como um fluxo de sinais de áudio com amostragem temporal. Por exemplo, um típico formato de áudio consiste de um fluxo de amostras de amplitude de 16 bits de um sinal de áudio obtido em intervalos de tempo regulares.

[00052] Os sinais de áudio digital, imagem e vídeo não comprimidos podem consumir considerável capacidade de armazenamento e transmissão. A codificação de transformada reduz o tamanho dos áudios, imagens e vídeos digitais pela transformação da representação no domínio espacial do sinal em uma representação no domínio da frequência (ou outro domínio de transformada similar), e então reduzindo a resolução de certos componentes de frequência, geralmente menos perceptíveis, da representação no domínio da transformada. Isso geralmente produz uma degradação muito menos perceptível do sinal digital se comparado à redução da cor ou resolução espacial das imagens ou vídeos no domínio espacial, ou do áudio no domínio do tempo.

[00053] Mais especificamente, um típico sistema codificador/decodificador baseado em transformada de bloco 500 (também chamado de "codec") apresentado na Figura 5 divide os pixels da imagem digital não comprimida em dois blocos dimensionais de tamanho fixo ($X_1, \dots, X_n$), cada bloco possivelmente em superposição com ou

tros blocos. Em um codificador 510, uma transformada linear 520-521 que efetua a análise de frequência espacial é aplicada a cada bloco, o que converte as amostras espaçadas dentro do bloco em um conjunto de coeficientes de frequência (ou transformada) geralmente representando a potência do sinal digital em bandas de frequência correspondentes ao longo do intervalo de bloco. Para compressão, os coeficientes de transformada podem ser seletivamente quantizados 530 (isto é, ter sua resolução reduzida, tal como pela eliminação dos bits menos significativos dos valores de coeficiente ou então mapeando valores em um conjunto de números de resolução superior para uma resolução inferior) e também codificados por entropia ou comprimento variável 530 em um fluxo de dados comprimido. Na decodificação, os coeficientes de transformada irão inversamente transformar 570-571 para reconstruir aproximadamente o sinal original de imagem/vídeo com amostragem espacial/de cor (blocos reconstruídos $\hat{x}_1, \dots, \hat{x}_n$).

[00054] A transformada de bloco 520-521 pode ser definida como uma operação matemática em um vetor x de tamanho N . Na maioria das vezes, a operação é uma multiplicação linear, produzindo a saída do domínio de transformada $y = M x$, M , sendo a matriz de transformada. Quando os dados de entrada são arbitrariamente longos, eles são segmentados em N vetores dimensionados e uma transformada de bloco é aplicada a cada segmento.

[00055] Para fins de compressão de dados, são escolhidas transformadas de bloco reversíveis. Em outras palavras, a matriz M é invertível. Em múltiplas dimensões (por exemplo, para imagem e vídeo), as transformadas de bloco são geralmente implementadas como operações separáveis. A multiplicação da matriz é aplicada de modo separável ao longo de cada dimensão dos dados (isto é, tanto fileiras quanto colunas).

[00056] Para compressão, os coeficientes de transformada (com-

ponentes do vetor y) podem ser seletivamente quantizados (isto é, ter sua resolução reduzida, tal como pela eliminação dos bits menos significativos dos valores de coeficiente ou então mapeando valores em um conjunto de números de resolução superior para uma resolução inferior) e também codificados por entropia ou comprimento variável em um fluxo de dados comprimido.

[00057] Na decodificação no decodificador 550, o inverso dessas operações (desquantização/decodificação por entropia 560 e transformada de bloco inversa 570-571 é aplicado ao lado do decodificador 550, como mostra a Figura 5. Durante a reconstrução dos dados, a matriz inversa M^{-1} (transformada inversa 570-571) é aplicada como um multiplicador aos dados de domínio da transformada. Quando aplicada aos dados de domínio da transformada, a transformada inversa reconstrói aproximadamente a mídia digital original no domínio do tempo ou no domínio espacial.

[00058] Em diversas aplicações de codificação baseadas em transformada de bloco, e transformada é preferencialmente reversível para suportar tanto a compressão com perdas quanto sem perdas, dependendo do fator de quantização. Sem nenhuma quantização (geralmente representada como um fator de quantização de 1) por exemplo, um codec utilizando uma transformada reversível pode reproduzir exatamente os dados de entrada na decodificação. No entanto, a necessidade de reversibilidade nessas aplicações restringe a escolha das transformadas baseadas nas quais o codec pode ser projetado.

[00059] Muitos sistemas de compressão de imagem e vídeo, tal como MPEG e Windows Media, dentre outros, utilizam transformadas baseando-se na Transformada Discreta de Cosseno (DCT). O DCT é conhecido por ter propriedades de compactação de energia favoráveis, que resultam numa compressão de dados quase ideal. Nesses sistemas de compressão, a DCT inversa (IDCT) é empregada nos ciclos de

reconstrução tanto no decodificador quanto no decodificador do sistema de compressão para reconstruir os blocos de imagem individuais.

IV. Exemplo de Implementação do Codificador/Decodificador

[00060] As Figuras 6 e 7 são um diagrama generalizado do processo empregado em um decodificador de dados 700 e em um codificador de dados 600 bidimensional (2D) representativo. Os diagramas apresentam uma ilustração generalizada ou simplificada de um sistema de compressão / descompressão que pode ser usado para implementar as técnicas e ferramentas descritas. Em sistemas de compressão / descompressão alternativos, podem ser usados menos processos ou processos adicionais aos ilustrados neste codificador e decodificador representativo para a compressão de dados 2D. Por exemplo, alguns codificadores/decodificadores também podem incluir conversão de cor, formatos de cor, codificação escalonável, codificação sem perdas, modos de macrobloco, etc. O sistema de compressão (codificador e decodificador) pode oferecer compressão com perda e/ou sem perda dos dados 2D, dependendo da quantização que pode ser baseada em um parâmetro de quantização variando de sem perdas para com perdas.

[00061] O codificador de dados 2D 600 produz um fluxo de bits comprimido 620 que é uma representação mais compacta (para típica entrada) de dados 2D 610 apresentados na forma de entrada ao codificador. Por exemplo, a entrada de dados 2D pode ser uma imagem, um quadro de uma sequência de vídeo ou outros dados com duas dimensões. O codificador de dados 2D divide um quadro dos dados de entrada em blocos (ilustrados em geral na Figura 6 como o particionamento 630), que, na implementação ilustrada, são blocos de pixels 4x4 não sobrepostos que formam um padrão regular ao longo do plano do quadro. Esses blocos são agrupados em aglomerados, chamados de macroblocos, que são os pixels com tamanho 16x16 neste codificador

representativo. Por sua vez, os macroblocos são agrupados em estruturas regulares chamadas de microblocos. Os microblocos também formam um padrão regular na imagem, de modo que os microblocos em uma fileira horizontal sejam de altura uniforme e estejam alinhados, e os microblocos em uma coluna vertical sejam de largura uniforme e estejam alinhados. No codificador representativo, os microblocos podem ser qualquer tamanho arbitrário que seja um múltiplo de 16 na direção horizontal e/ou vertical. Implementações alternativas do codificador podem dividir a imagem em blocos, macroblocos, microblocos, ou outras unidades de tamanhos e estruturas diferentes.

[00062] Um operador de "sobreposição direta" 640 é aplicado a cada borda entre os blocos, após o que cada bloco de 4x4 é transformado usando uma transformada de bloco 650. Essa transformada de bloco 650 pode ser a transformada 2D reversível, em escala livre, descrita por Srinivasan, Pedido de Patente U.S. Nº 11/015,707, intitulado "Reversible Transform For Lossy And Lossless 2-D Data Compression", depositado em 17 de dezembro de 2004. O operador de sobreposição 640 pode ser o operador de sobreposição reversível descrito por Tu e col., Pedido de Patente U.S. Nº 11/015,148, intitulado, "Reversible Overlap Operator for Efficient Lossless Data Compression", depositado em 17 de dezembro de 2004; e por Tu e col., Pedido de Patente U.S. Nº 11/035,991, intitulado, "Reversible 2-Dimensional Pre-/Post-Filtering For Lapped Biorthogonal Transform", depositado em 14 de janeiro de 2005. Como alternativa, a transformada discreta de cosseno ou outras transformadas de bloco e operadores de sobreposição podem ser utilizados. Após a transformada, o coeficiente DC 660 de cada bloco de transformada 4x4 é submetido a uma cadeia de processamento similar (inclinação, sobreposição direta, seguido de uma transformada de bloco de 4x4). Os coeficientes de transformada DC resultantes e os coeficientes de transformada AC 662 são quantizados 670,

codificados por entropia 680 e empacotados 690.

[00063] O decodificador realiza o processo inverso. No lado do decodificador, os bits do coeficiente de transformada são extraídos 710 de seus respectivos pacotes, dos quais os próprios coeficientes são decodificados 720 e desquantizados 730. Os coeficientes DC 740 são regenerados pela aplicação de uma transformada inversa, e o plano dos coeficientes DC é "sobreposto por inversão" usando um operador de suavização adequado aplicado às bordas do bloco DC. Subsequentemente, todos os dados são regenerados pela aplicação da transformada inversa 4x4 750 aos coeficientes DC, e os coeficientes AC 742 são decodificados do fluxo de bits. Finalmente, as bordas de bloco nos planos de imagem resultante são filtrados por sobreposição inversa 760. Isso produz uma saída de dados 2D reconstruída 790.

[00064] Em uma implementação exemplificativa, o codificador 600 (Figura 6) comprime uma imagem de entrada no fluxo de bits comprimido 620 (por exemplo, um arquivo) e o decodificador 700 (Figura 7) reconstrói a entrada original ou uma aproximação dela, baseado em se a codificação com perdas ou sem perdas é empregada. O processo de codificação envolve a aplicação de uma transformada sobreposta direta (LT) discutida a seguir, que é implementada com a pré/pós-filtragem bidimensional também descrita em mais detalhes a seguir. O processo de decodificação envolve a aplicação da transformada sobreposta inversa (ILT) usando a pré/pós-filtragem bidimensional reversível.

[00065] A LT e a ILT ilustrada são inversos uma da outra, num sentido exato, e, portanto, podem ser chamadas coletivamente de transformada sobreposta reversível. Como uma transformada reversível, o par LT/ILT pode ser usado para compressão de imagem sem perdas.

[00066] Os dados de entrada 610 comprimidos pelo codificador 600/decodificador 700 ilustrado podem ser imagens de vários formatos de cores (por exemplo, formatos de imagem colorida RGB/YUV 4:4:4,

YUV 4:2:2 ou YUV 4:2:0). Geralmente, a imagem de entrada tem um componente de luminância (Y). Caso seja uma imagem RGB/YUV 4:4:4, YUV 4:2:2 ou YUV 4:2:0, a imagem também tem componentes de croma, tal como um componente U e um componente V. Os planos de cor ou componentes separados da imagem podem ter diferentes resoluções espaciais. No caso de uma imagem de entrada no formato de cores YUV 4:2:0, por exemplo, os componentes U e V possuem metade da largura e altura do componente Y.

[00067] Como discutido acima, o codificador 600 divide a imagem de entrada em macroblocos. Em uma implementação exemplificativa, o codificador 600 divide a imagem de entrada em áreas de 16x16 pixels (chamadas de "macroblocos") no canal Y (que podem ser áreas de 16x16, 16x8 ou 8x8 nos canais U e V, dependendo do formato de cor). Cada plano de cor de macrobloco é dividido em regiões ou blocos de 4x4 pixels. Portanto, um macrobloco é composto para os vários formatos de cores da seguinte maneira para essa implementação exemplificativa do codificador:

- Para uma imagem em escala de cinza, cada macrobloco contém 16 blocos de luminância (Y) de 4x4.
- Para uma imagem colorida no formato YUV 4:2:0, cada macrobloco contém 16 blocos Y de 4x4, e 4 blocos de croma (U e V) de 4x4.
- Para uma imagem colorida no formato YUV 4:2:2, cada macrobloco contém 16 blocos Y de 4x4, e 8 blocos de croma (U e V) de 4x4.
- Para uma imagem colorida no formato RGB ou YUV 4:4:4, cada macrobloco contém 16 blocos de canais Y, U e V.

[00068] Sendo assim, após a transformada, um macrobloco neste codificador 600/decodificador 700 representativo possui três subbandas de frequência: uma subbanda DC (macrobloco DC), uma subbanda

da passa-baixa (macrobloco passa-baixa) e uma subbanda passa-alta (macrobloco passa-alta). No sistema representativo, as subbandas passa-baixa e/ou passa-alta são opcionais no fluxo de bits—essas subbandas podem ser totalmente eliminadas.

[00069] Além disso, os dados comprimidos podem ser empacotados no fluxo de bits em uma dentre duas ordenações: ordem espacial e ordem de frequência. Para a ordem espacial, diferentes subbandas do mesmo macrobloco dentro de um microbloco são ordenadas juntas, e o fluxo de bits resultante de cada microbloco é gravado em um pacote. Para a ordem de frequência, a mesma subbanda de diferentes macroblocos dentro de um microbloco é agrupada, e, assim, o fluxo de bits de um microbloco é gravado em três pacotes: um pacote de microbloco DC, um pacote de microbloco passa-baixa e um pacote de microbloco passa-alta. Além disso, pode haver outras camadas de dados.

[00070] Sendo assim, para o sistema representativo, uma imagem é organizada nas seguintes "dimensões":

- Dimensão espacial : Quadro → Microbloco → Macrobloco;
- Dimensão de frequência : DC | Passa-baixa | Passa-alta; e
- Dimensão de canal: Luminância | Crominância_0 | Crominância_1 ... (por exemplo, como Y | U | V).

[00071] As setas acima indicam uma hierarquia, ao passo que as barras verticais indicam uma divisão.

[00072] Embora o sistema representativo organize os dados de mídia digital comprimidos em dimensões espacial, de frequência e canal, a abordagem de quantização flexível descrita na presente invenção pode ser aplicada em sistemas de codificador/decodificador alternativos que organizam seus dados ao longo de menos dimensões, dimensões adicionais ou dimensões diferentes. Por exemplo, a abordagem de quantização flexível pode ser aplicada à codificação usando um

número maior de bandas de frequência, outro formato de canais de cor (por exemplo, YIQ, RGB, etc.), canais de imagem adicionais (por exemplo, para "stereo vision" ou múltiplos arranjos de câmeras).

[00073] Tendo descrito e ilustrado os princípios da invenção com referência às várias concretizações descritas, reconhecer-se-á que as várias concretizações descritas podem ser modificadas, tanto na configuração como nos detalhes, sem divergir de tais princípios. Deve-se entender que os programas, processos ou métodos descritos neste documento não estão relacionados ou limitados a qualquer tipo específico de ambiente de computação, salvo indicação ao contrário. Vários tipos de ambientes de computação de uso geral ou específico podem ser usados com, ou realizar, operações de acordo com os ensinamentos descritos no presente documento. Os elementos das concretizações ilustradas em software podem ser implementados em hardware, e vice versa.

[00074] Em vista das diversas concretizações possíveis às quais os princípios da invenção revelada podem ser aplicados, deve-se reconhecer que as concretizações ilustradas são apenas exemplos preferidos da invenção e não devem ser interpretados como limitações ao âmbito da invenção. Em vez disso, o âmbito da invenção é definido pelas reivindicações a seguir. Portanto, reivindicamos como nossa invenção tudo o que se enquadrar no âmbito e espírito dessas reivindicações.

REIVINDICAÇÕES

1. Método, **caracterizado pelo fato de que** compreende as etapas de:

receber dados de imagem digital codificados (620) para uma imagem fixa, em que a imagem fixa é codificada usando uma transformada sobreposta reversível, os dados de imagem digital codificados (620) compreendendo informações de crominância subamostradas;

decodificar as informações de posição de crominância para as informações de crominância subamostradas, em que as informações de posição de crominância são sinalizadas no cabeçalho da imagem, em que as informações de posição de crominância são determinadas conforme uma regra de posição de crominância que mapeia as informações de posição de crominância para as informações de fase em uma dimensão X e para as informações de fase em uma dimensão Y, em que a regra de posição de crominância mapeia as informações de posição de crominância para informações de fase que incluem posições de fase -0,5, -0,25, 0, 0,25, 0,5, 0,75, 1, 1,25 e 1,5, onde as posições de fase estão em unidades de pixel, em que a decodificação das informações de posição de crominância compreendem a decodificação de um primeiro elemento de sintaxe que representa a informação de fase na dimensão X e um segundo elemento de sintaxe que representa a informação de fase na dimensão Y;

executar uma operação de rotação ou inversão para os dados de imagem digital codificados (620) enquanto a imagem fixa está no domínio compactado, incluindo a execução da operação de inversão através da modificação dos coeficientes da transformada em blocos da transformada gerados usando a transformada sobreposta reversível ou a execução da operação de rotação através da modificação e transposição dos coeficientes da transformada nos blocos da

transformada, em que a modificação dos coeficientes da transformada compreende negar o sinal de coeficientes da transformada de simetria ímpar nos blocos da transformada, e modificar as informações de posição de cromaticidade para refletir a operação de rotação ou inversão;

decodificar a imagem fixa a partir dos dados de imagem digital codificados (620) usando um inverso de uma transformada sobreposta reversível; e

gerar uma imagem fixa decodificada (790).

2. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** as informações de cromaticidade subamostradas são subamostradas por um fator de dois em uma ou mais dimensões.

3. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** as informações de fase na dimensão X compreendem informações de fase de meia unidade de pixel.

4. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** as informações de fase na dimensão Y ou na dimensão X compreendem informações de fase de um quarto de unidade de pixel.

5. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** ainda compreende usar as informações de posição de cromaticidade para sobreamostrar as informações de cromaticidade subamostradas.

6. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** ainda compreende usar as informações de posição de cromaticidade para executar a operação de rotação ou inversão para a imagem codificada (620).

7. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** realizar a operação de inversão ainda compreende varrer os blocos da transformada modificados em uma sequência invertida lateralmente.

8. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** realizar a operação de rotação ainda compreende a transposição dos blocos modificados e a varredura dos blocos da transformada modificados em uma sequência invertida lateralmente.

9. Meio de armazenamento **caracterizado pelo fato de que** contém o método conforme definido em qualquer uma das reivindicações 1 a 8.

10. Sistema de processamento de mídia digital, **caracterizado pelo fato de que** está configurado para executar o método conforme definido em qualquer uma das reivindicações 1 a 8.

Figura 1

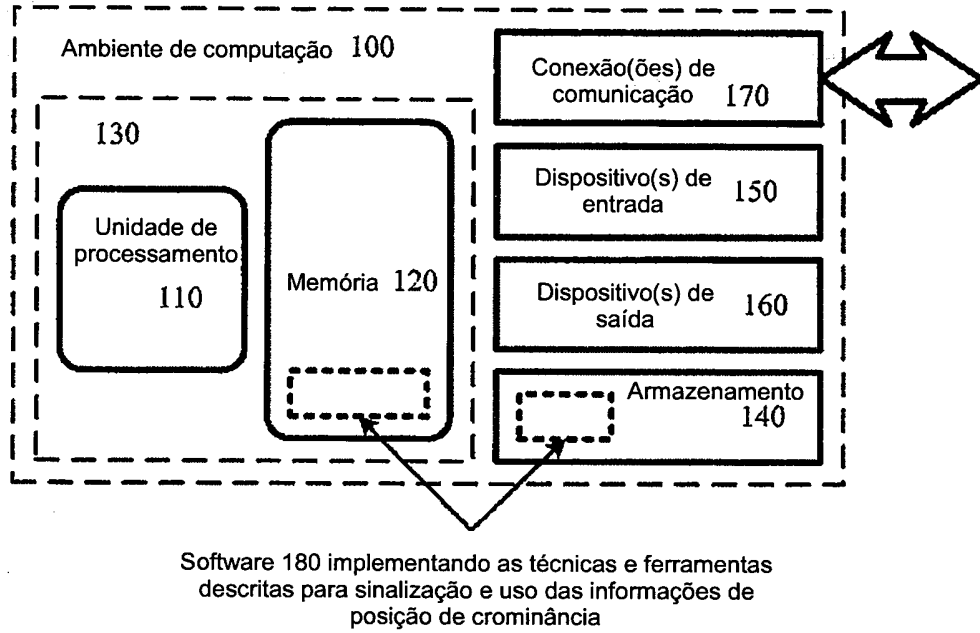
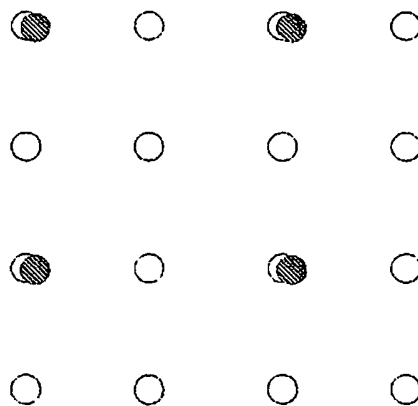


Figura 2



○ amostra de luminância

● amostra de crominância (ilustrada com ligeiro desvio por clareza)

Figura 3

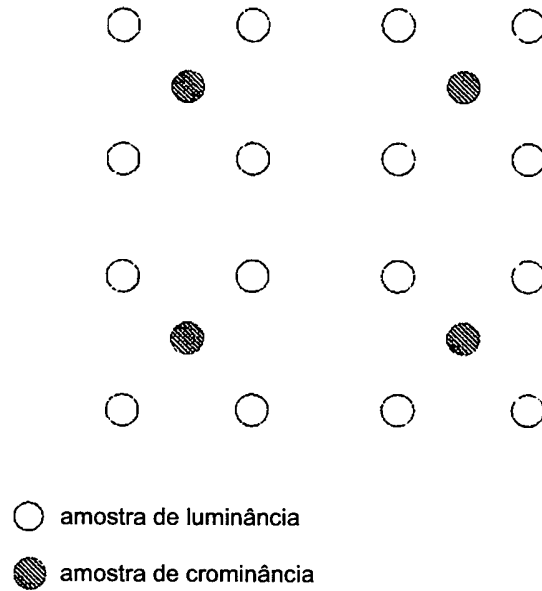


Figura 4

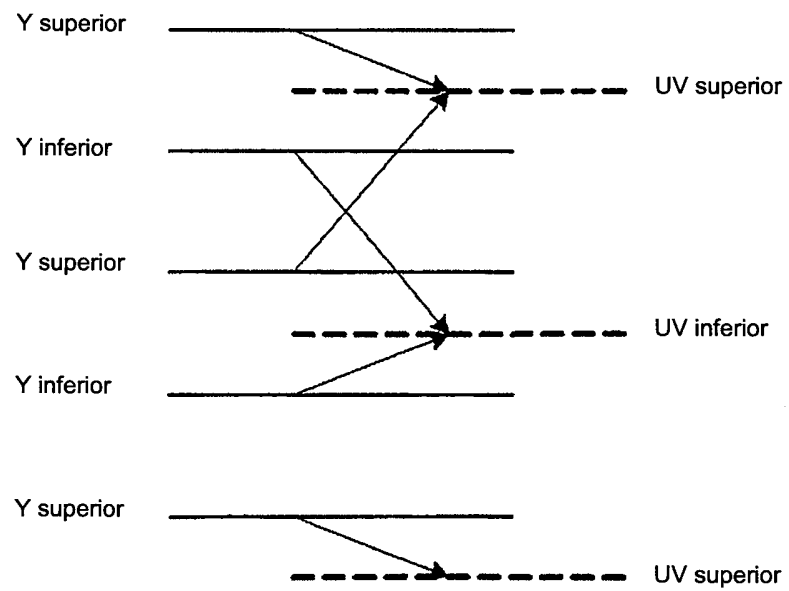


Figura 5
Técnica Anterior

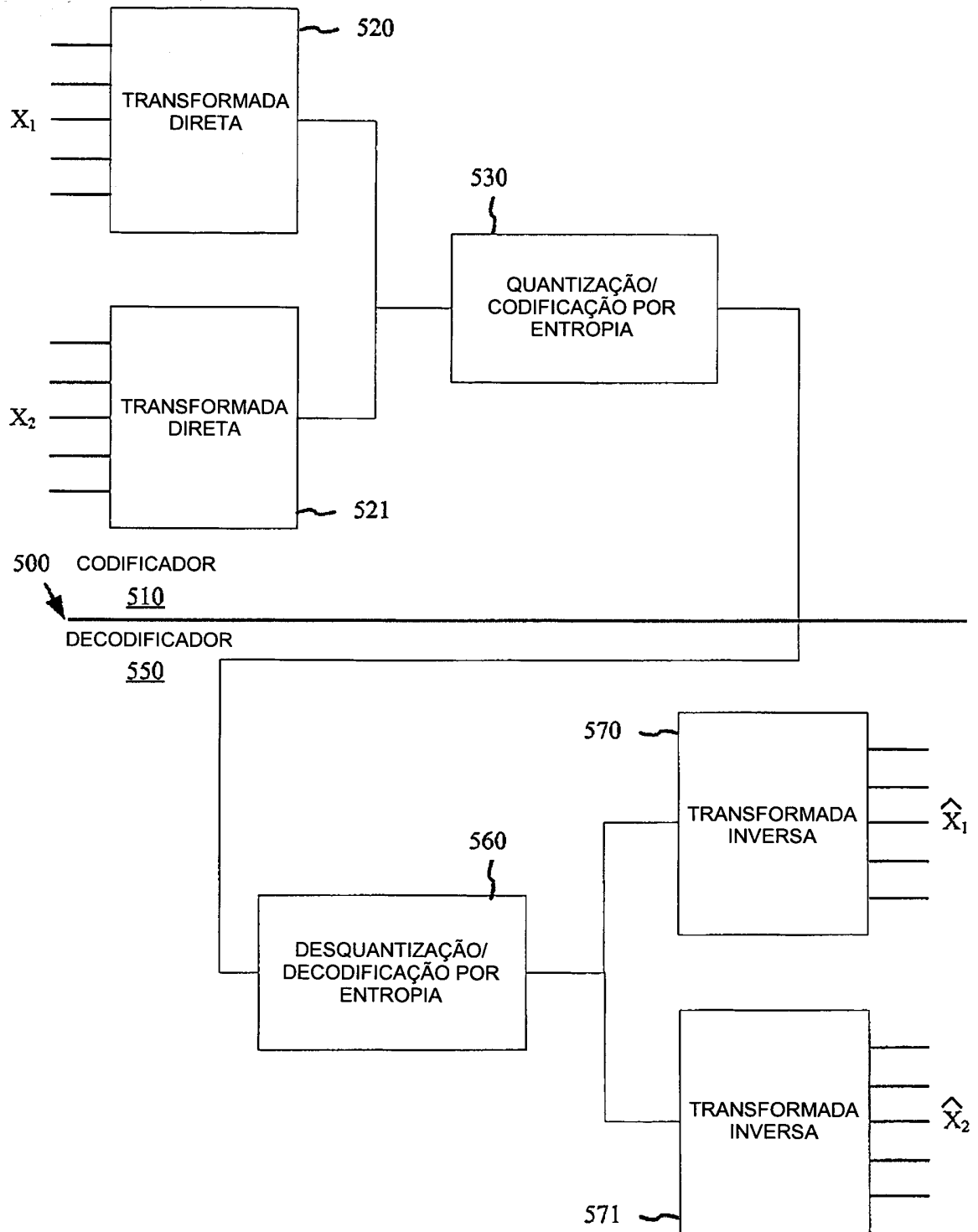


Figura 6

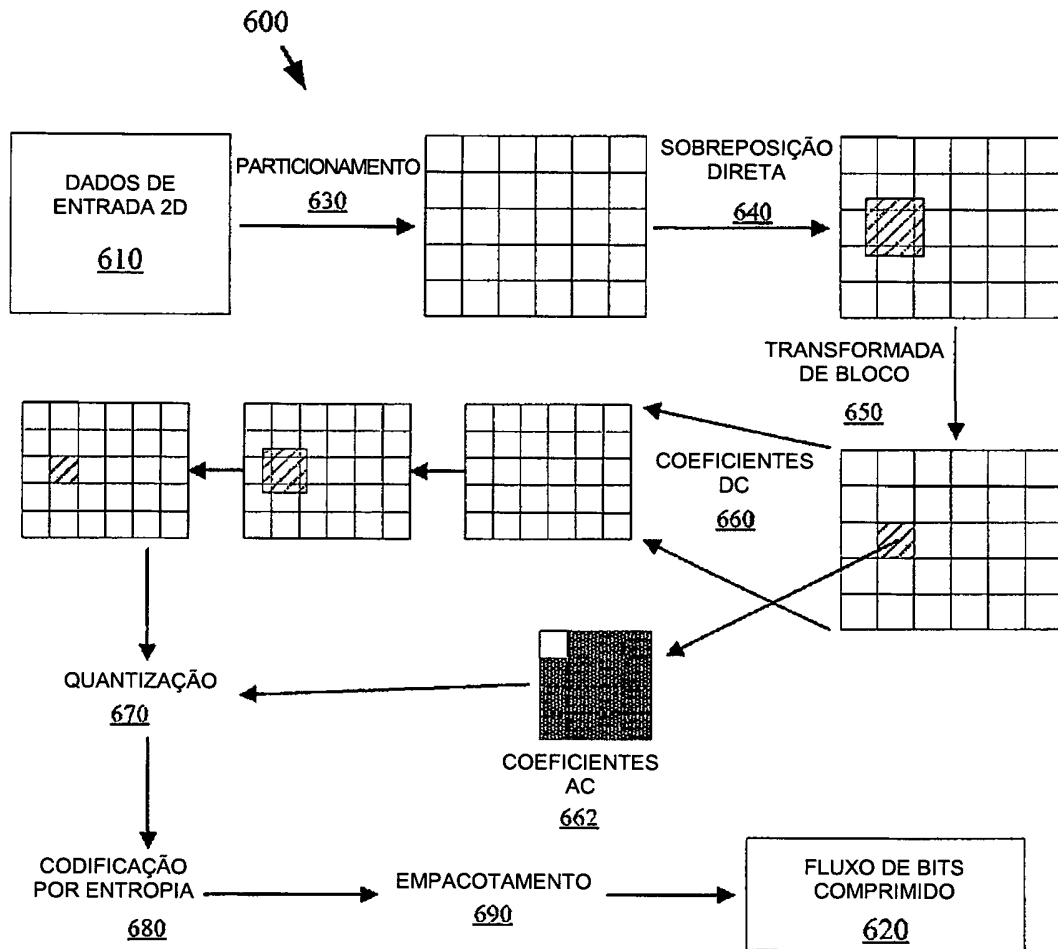


Figura 7

