



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

(11) PI 0910340-6 B1



(22) Data do Depósito: 07/04/2009

(45) Data de Concessão: 20/10/2020

(54) Título: MÉTODO E APARELHO DE DECODIFICAÇÃO DE VÍDEO

(51) Int.Cl.: H04N 19/39; H04N 19/172; H04N 19/187; H04N 19/63.

(52) CPC: H04N 19/39; H04N 19/172; H04N 19/187; H04N 19/63.

(30) Prioridade Unionista: 24/04/2008 US 12/108,680.

(73) Titular(es): GOOGLE TECHNOLOGY HOLDINGS LLC.

(72) Inventor(es): TAMER SHANABLEH; SHIH-TA HSIANG; FAISAL ISHTIAQ.

(86) Pedido PCT: PCT US2009039739 de 07/04/2009

(87) Publicação PCT: WO 2009/131830 de 29/10/2009

(85) Data do Início da Fase Nacional: 30/09/2010

(57) Resumo: MÉTODO E APARELHO PARA CODIFICAÇÃO E DECODIFICAÇÃO DE VÍDEO Um método e um aparelho para a codificação e a decodificação de vídeo realizam uma transformação de pelo menos uma porção de um quadro de vídeo de resolução alta em uma imagem de resolução baixa e uma pluralidade de conjuntos de dados de melhoramento, codifica a imagem de resolução baixa como uma imagem pictórica codificada primária em um formato de fluxo de bit e codifica cada um da pluralidade de conjuntos de dados de melhoramento como uma imagem pictórica codificada redundante diferente no formato de fluxo de bit. Para a decodificação, uma imagem de resolução baixa decodificada e uma pluralidade de conjuntos de dados de melhoramento decodificados são geradas e uma transformada inversa é realizada para a construção de uma imagem de resolução alta decodificada. A imagem pictórica codificada primária e a imagem pictórica codificada redundante podem ser formatadas de acordo com a especificação de Codificação Avançada da ITU-T H.264. A transformada pode ser uma transformada de polifase ou de sub-banda.

MÉTODO E APARELHO DE DECODIFICAÇÃO DE VÍDEO**Pedidos Relacionados**

O presente pedido está relacionado ao pedido U.S. a seguir comumente possuído com este pedido pela Motorola, Inc.: N° de Série 11/866.771, depositado em 3 de outubro de 2007, intitulado "METHOD AND APPARATUS FOR INTRA FRAME SPACTIAL SCALABLE VIDEO CODING" (protocolo legal N° CML04718EV).

Campo da Exposição

10 A presente exposição se refere geralmente a uma comunicação de vídeo e, mais particularmente, a técnicas de comunicação de vídeo provendo uma resolução de imagem escalonável e resiliência melhorada a erro.

Antecedentes

15 No padrão de codificação de vídeo H.264/AVC datado de 3/2005, cada unidade de acesso (p. 4, seção 3.1, definição de "unidade de acesso") compreende uma imagem pictórica primária codificada para a provisão da informação requerida para a representação de uma imagem pictórica em movimento
20 codificada. Uma unidade de acesso pode conter, ainda, uma ou mais imagens pictóricas codificadas redundantes, as quais "não têm efeito normativo sobre o processo de codificação". A Seção 7.4.3 ("Slice header semantics: redundant_pic_cnt") destas normas estabelece: "Não há um
25 processo de decodificação requerido para uma fatia codificada ou uma partição de dados de fatia codificada de uma imagem pictórica codificada redundante. Quando redundant_pic_cnt no cabeçalho de uma fatia de uma fatia codificada é maior do que 0, o decodificador poderá
30 descartar a fatia codificada. NOTA 6 - quando algumas das

amostras na imagem pictórica primária decodificada não podem ser decodificadas corretamente devido a erros ou a perdas na transmissão da seqüência e uma fatia redundante codificada pode ser decodificada corretamente, o

5 decodificador deve substituir as amostras da imagem pictórica primária decodificada pelas amostras correspondentes da fatia redundante decodificada. Quando mais de uma fatia redundante cobre a região relevante da imagem pictórica primária, a fatia redundante tendo o valor

10 mais baixo de `redundant_pic_cnt` deve ser usada". Assim, a presença de uma ou mais imagens pictóricas codificadas redundantes no fluxo de bit codificado é opcional e o uso pretendido (mas não um uso requerido) de uma imagem pictórica codificada redundante é a correção de erros de

15 transmissão na imagem pictórica codificada primária. A Wikipedia.com apresentou (em 15 de março de 2008) o recurso de imagem pictórica codificada redundante de H.264 conforme se segue: "Fatias redundantes (RS), um recurso de robustez a erro / perda permitindo que um codificador envie uma

20 representação extra de uma região de imagem pictórica (tipicamente a uma fidelidade mais baixa) que pode ser usada se a representação primária estiver corrompida ou for perdida".

A nova versão pré-publicada da norma de MPEG-4

25 AVC/H.264 ainda provê ferramentas de codificação de vídeo escalonáveis e permite que um decodificador recupere a imagem pictórica codificada primária ou uma forma de resolução mais alta da imagem pictórica codificada primária. Contudo, estas novas ferramentas de codificação

30 padronizadas envolvendo uma complexidade aumentada de

implementação não foram adotadas rapidamente pela indústria.

Breve Descrição das Figuras

As figuras associadas, onde números de referência
5 iguais se referem a elementos funcionalmente idênticos ou similares por todas as várias vistas separadas, em conjunto com a descrição detalhada abaixo, são incorporadas no e fazem parte do relatório descritivo, e servem para ilustração adicional das modalidades de conceitos que
10 incluem a invenção reivindicada, e explicam vários princípios e vantagens daquelas modalidades.

A FIG. 1 é um diagrama que ilustra uma porção de um quadro de vídeo colorido de acordo com certas modalidades.

As FIG. 2 e 3 são, respectivamente, um diagrama de
15 blocos funcional e um fluxograma de um aparelho e algumas etapas de um método para a codificação de pelo menos uma porção de um quadro de vídeo, de acordo com certas modalidades.

A FIG. 4 é um fluxograma que mostra uma etapa de um
20 método para a realização de uma transformação, de acordo com certas modalidades.

A FIG. 5 é um diagrama de blocos de uma função de transformada, de acordo com certas modalidades.

A FIG. 6 é um diagrama que ilustra a grade de
25 amostragem para uma porção de um quadro de vídeo, de acordo com certas modalidades.

A FIG. 7 é um fluxograma que mostra algumas etapas de um método para a realização de uma transformada, de acordo com certas modalidades.

30 As FIG. 8 e 9 são, respectivamente, um diagrama de

blocos de um decodificador e etapas de um método para a decodificação, mostrados de acordo com certas modalidades.

A FIG. 10 é um diagrama de blocos de uma função de transformada inversa para uma função de cálculo de média
5 descrita com referência à FIG. 6.

Os técnicos versados apreciarão que os elementos nas FIG. são ilustrados por simplicidade e clareza e não necessariamente foram desenhados em escala. Por exemplo, as dimensões de alguns dos elementos nas FIG. podem estar
10 exageradas em relação a outros elementos, para se ajudar no entendimento de modalidades da presente invenção.

Os componentes de aparelho e de método foram representados, onde apropriado, por símbolos convencionais nos desenhos, mostrando apenas aqueles detalhes específicos
15 que são pertinentes ao entendimento das modalidades da presente invenção, de modo a não se obscurecer a exposição com detalhes que serão prontamente evidentes para aqueles de conhecimento comum na técnica tendo o benefício da descrição aqui.

20

Descrição Detalhada

De acordo com certos aspectos das presentes modalidades, são descritas técnicas que provêm uma transformação de um quadro de vídeo de resolução alta original em uma imagem de resolução baixa e uma pluralidade
25 de conjuntos de dados de melhoramento, que então são codificados e podem ser transmitidos como fluxos de bit codificados. A imagem de resolução baixa pode ser decodificada e usada para a geração de uma versão de resolução baixa do quadro de vídeo de resolução alta
30 recuperado. A imagem de resolução baixa e a pluralidade de

conjuntos de dados de melhoramento podem ser decodificadas e transformadas inversamente para a construção de um quadro de vídeo de resolução alta decodificado. Em algumas modalidades, a maneira pela qual o vídeo é codificado e decodificado é compatível com a especificação publicada pela International Telecommunication Union, que é identificada como a especificação ITU-T H.264 com data de 03/2005, intitulada "SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS Infrastructure of audiovisual services - Coding of moving video / Advanced video coding for generic audiovisual services", a qual será referida a partir deste ponto como a "Codificação Avançada da ITU-T H.264", e a especificação equivalente publicada pelo Grupo de Especialistas em Imagens com Movimento, a qual é comumente referida como a especificação de AVC Parte 10 MPEG-4. Mais ainda, a maneira pela qual o vídeo é codificado e decodificado pode ser compatível com especificações ainda não lançadas. Em particular, as técnicas descritas aqui exploram o novo uso da ferramenta de imagem pictórica redundante, o que não tem efeito normativo sobre o processo de decodificação, em conjunto com uma codificação de sub-banda e uma codificação de descrição múltipla (MDC). Estas técnicas também podem ser compatíveis com outras técnicas de codificação de vídeo além daquelas especificadas pela Codificação Avançada da ITU-T H.264.

Uma transformada de sub-banda / ondulação decompõe uma imagem de entrada em uma representação de sinal de resolução múltipla. Foi demonstrado ser um dos métodos mais eficientes para compressão de imagem na literatura, e é utilizado na norma internacional JPEG 2000 para aplicações

de codificação de imagem e vídeo (no formato de JPEG 2000 em Movimento) na indústria. Graças à alta compactação de energia de uma transformada de sub-banda / ondulação, os codificadores de sub-banda / ondulação são capazes de
5 obterem excelente performance de compressão sem artefatos de bloco tradicionais associados à transformada de bloco. De forma mais importante, eles podem acomodar facilmente a funcionalidade de codificação escalonável espacial desejável quase sem penalidade na eficiência de compressão,
10 porque a transformada de sub-banda de ondulação é de resolução escalonável por natureza.

Certas formas de redução de amostragem de polifase têm sido usadas em uma codificação de descrição múltipla. Uma codificação de vídeo de descrição múltipla (MD) é uma
15 abordagem para transmissão de vídeo em difusão simultânea e multidifusão por redes propensas a erros. Uma codificação de vídeo de descrição múltipla é conceitualmente diferente da codificação de vídeo escalonável de camada múltipla tradicional disponível em codificadores - decodificadores
20 híbridos, tais como MPEG-2, MPEG-4 e na família H.26x de especificações. Em uma codificação de vídeo de MD, vários descritores de vídeo igualmente importantes são gerados e, talvez, entregues através de canais ou percursos de transmissão diferentes. Cada descritor de vídeo é
25 individualmente decodificável e capaz de reproduzir a resolução de vídeo original, embora possivelmente a uma qualidade mais baixa. Portanto, o vídeo codificado pode ser recuperado e consumido por um cliente remoto, desde que pelo menos uma descrição tenha sido recebida de forma bem
30 sucedida. A noção de camadas de vídeo hierárquicas não se

aplica; daí, quanto mais descritores de vídeo disponíveis para o receptor, mais alto é o vídeo decodificado, independentemente da ordenação dos descritores recebidos. Em outras palavras, uma codificação de vídeo de descrição múltipla remove as dependências entre camadas, e, portanto, a priorização de camada de vídeo introduzida pela codificação de camada múltipla tradicional não é mais necessária.

A FIG. 1 é um diagrama que mostra uma grade de amostragem para uma porção 100 de um quadro de vídeo colorido, de acordo com certas modalidades. É similar à figura 6-1 da Codificação Avançada da ITU-T H.264, a qual é intitulada "Localizações verticais e horizontais nominais de amostras de luma e croma a 4:2:0 em um quadro". Na FIG. 1, as amostras de luma 105 do quadro de vídeo são mostradas como Xs e um de dois conjuntos associados de amostras de croma 110 são mostrados como Os. As amostras de croma no outro conjunto de croma ocorreriam nas mesmas posições mostradas para amostras de croma 110. A quantidade de amostras de luma 105 está relacionada à resolução e ao tamanho da imagem de vídeo no quadro. A quantidade de amostras de croma está relacionada à quantidade de amostras de luma. A FIG. 1 ilustra um quadro de vídeo especificado em um formato 4:2:0 no qual a quantidade de amostras de croma é igual a $\frac{1}{4}$ da quantidade de amostras de luma. A norma de Codificação Avançada da ITU-T H.264 provê outras relações, tais como 1 para 1 e $\frac{1}{2}$ para 1. Obviamente, uma imagem monocromática não tem amostras de croma. Quando um quadro de vídeo, ou uma porção de um quadro de vídeo, tal como uma divisão, um macrobloco ou uma região de interesse

é codificada, por exemplo, de acordo com a norma de Codificação Avançada da ITU-T H.264, as operações realizadas nas amostras de luma também são realizadas nas amostras de croma, escalonadas para a quantidade de amostras de croma. O mesmo é verdadeiro para as operações 5 únicas definidas aqui. Para simplificar a descrição das operações, elas são descritas em termos de pixels, onde um pixel pode ser tomado como uma amostra de luma ou uma amostra de croma. Para simplificar mais os termos usados aqui, quando o termo "porção" é usado, ele significa "pelo 10 menos uma porção", o que pode incluir uma parte, partes ou o todo. Quando o quadro de vídeo é um quadro colorido, as operações descritas aqui são realizadas nos conjuntos de luma e de croma de amostras para a porção do quadro de vídeo, escalonada para os tamanhos dos conjuntos de 15 amostras.

As FIG. 2 e 3 são, respectivamente, um diagrama de blocos funcional 200 e um fluxograma 300 de um aparelho e algumas etapas de um método para a codificação de pelo 20 menos uma porção de um quadro de vídeo, de acordo com certas modalidades. Com referência à FIG. 2, uma informação a partir de uma seqüência de quadros de vídeo é acoplada a uma entrada 210 de uma função de transformada 215. Tipicamente, a informação é operada sobre um quadro de vídeo em um momento, ou pelo menos uma porção (tal como uma 25 fatia) de um quadro de vídeo. Neste exemplo, o quadro de vídeo 206 de uma seqüência de quadro de vídeo 205 está sendo operado pela função de transformada 215. O quadro de vídeo 206 é um quadro de vídeo de resolução alta. A função de transformada 215 transforma (305, FIG. 3) pelo menos uma 30

porção do quadro de vídeo 206 (FIG. 2) em uma imagem de resolução baixa 220 e uma pluralidade de conjuntos de dados de melhoramento 225, 230, 235. No exemplo mostrado na FIG. 2, há três conjuntos de dados de melhoramento 225, 230, 235. A pluralidade de conjuntos de dados de melhoramento poderia ser uma outra quantidade além de três em algumas modalidades. A transformação usada pela função de transformada 215 tem várias modalidades, conforme descrito em maiores detalhes abaixo. Os tipos de transformações que são usados pela função de transformada 215 em certas modalidades são tais que uma realização da inversa da transformação da imagem de resolução baixa e da pluralidade de conjuntos de dados melhorados reconstrua o quadro de vídeo de resolução alta (ou uma porção do mesmo).

A imagem de resolução baixa 220 é acoplada a um codificador de imagem pictórica primária (codificador de PP) 240, o qual codifica a imagem de resolução baixa usando técnicas de codificação para uma codificação de imagem pictórica primária, gerando uma imagem pictórica codificada primária 250. Cada um da pluralidade de conjuntos de dados de melhoramento é acoplado a um codificador de imagem pictórica redundante (codificador de RP) 245, o qual codifica o conjunto de dados de melhoramento correspondente usando técnicas de codificação para a codificação de imagem pictórica redundante, gerando uma imagem pictórica codificada redundante 255, 260, 265. De acordo com certas modalidades, o codificador de imagem pictórica primária 240 codifica (310, FIG. 3) a imagem de resolução baixa como uma imagem pictórica primária em um formato que é especificado pela Codificação Avançada da ITU-T H.264, e cada

codificador de imagem pictórica redundante 245 codifica (315, FIG. 3) um dos conjuntos de dados de melhoramento como uma imagem pictórica codificada redundante diferente no formato que é especificado pela Codificação Avançada da ITU-T H.264. Quando outras técnicas de codificação além 5 daquelas especificadas pela Codificação Avançada da ITU-T H.264 são usadas, a codificação descrita acima como uma codificação de imagem pictórica primária pode ser realizada por qualquer método de codificação da técnica que é 10 compatível com uma codificação de uma imagem de resolução baixa que normalmente é decodificada por um decodificador para aquela técnica, e a codificação descrita acima como uma codificação de imagem pictórica redundante pode ser realizada por qualquer método de codificação da técnica que 15 é compatível com uma codificação de um conjunto de dados de melhoramento ou suplementar - isto é, um que seja identificado para um decodificador para aquela técnica como estando associado ao quadro de vídeo de resolução alta, e para o qual uma decodificação pode ser ótima. Os formatos 20 de imagem e de codificação usados para outras técnicas podem ser identificados por outros nomes além de uma imagem pictórica codificada primária e uma "imagem pictórica codificada redundante". Quando uma imagem de resolução baixa em uma codificação de conjunto de dados de 25 melhoramento é completada, a imagem pictórica codificada primária 250 e a pluralidade de imagens pictóricas codificadas redundantes 255, 260, 265 são acopladas a uma função de formação de fluxo de bit 270, a qual os forma e um fluxo de bit 275, o qual então pode ser transmitido 30 (320, FIG. 3).

Com referência à FIG. 4, um fluxograma mostra uma etapa de um método para a realização de uma transformação, de acordo com certas modalidades. Nestas modalidades, a função de transformada 215 realiza em 405 uma transformação de sub-banda da porção do quadro de vídeo de resolução alta par a geração de uma representação de resolução baixa e conjuntos de dados de melhoramento comumente referidos como sub-bandas LL, HL, LH e HH do quadro de vídeo (ou de uma porção do mesmo).

10 Com referência à FIG. 5, um diagrama de blocos da função de transformada de sub-banda 215 é mostrado, de acordo com certas modalidades. A função de transformada 215 compreende bancos de filtro de análise de sub-banda 505, 515, os quais são bancos de filtro de sub-banda 15 convencionais e, em algumas modalidades, bancos de filtro de ondulação, que são bem conhecidos na técnica. O quadro de vídeo de entrada (ou uma porção do mesmo) 206 é processado, respectivamente, em primeiro lugar, por um filtro de passa baixa 506 e um filtro de passa alta 507, 20 seguidos por uma operação de redução de amostragem ao longo da direção horizontal, gerando as quatro sub-bandas LL 220, HL 225, LH 230 e HH 235 que são submetidas a uma codificação pelo codificador de PP 240 e pelos codificadores de PR 245 na FIG. 2, respectivamente. Este 25 processo é comumente referido como decomposição de sub-banda. Os bancos de filtro de síntese de sub-banda correspondentes que são usados na função de transformada inversa 875 de decodificador 800 (FIG. 8) podem construir a 30 decodificada do quadro de vídeo de resolução alta original

206 que foi decomposto pelos bancos de filtro de análise de sub-banda 505, 515 (FIG. 5). Os filtros usados nos bancos de filtro de análise / sub-banda 505, 515 podem pertencer a uma família de filtros de ondulação ou a uma família de bancos de filtro de sub-banda que podem ser de uma estrutura diferente daquela mostrada, que é conhecida na técnica.

Com referência à FIG. 6, um diagrama ilustra uma porção de um quadro de vídeo 600, de acordo com certas modalidades. Os pixels da porção do quadro de vídeo mostrado na FIG. 6 tiveram uma redução de amostragem em fases da porção do quadro de vídeo. Neste exemplo, a redução de amostragem foi realizada com um fator de dois em cada dimensão espacial. Assim, quatro fases de conjuntos unicamente diferentes de pixels foram formadas, cada uma compreendendo aproximadamente um quarto do número de pixels na porção do quadro de vídeo. A quantidade de um quarto é estabelecida como sendo aproximada, devido a efeitos de borda. Os pixels das quatro fases diferentes são descritos como triângulos, xs, quadrados e círculos. Neste exemplo, uma vizinhança de pixel pode ser definida, que é válida para qualquer pixel no quadro de vídeo. Neste exemplo, uma vizinhança de pixel é definida para incluir um pixel selecionado, tal como o pixel 615, e os outros pixels incluídos na fronteira de vizinhança 610. Em geral, uma vizinhança de pixel pode ter quaisquer formato e tamanho. Um grupo de fase de pixel pode ser definido para qualquer pixel no quadro de vídeo. Um grupo de fase de pixel inclui um pixel a partir de cada fase localizado na mesma posição na respectiva grade de imagem de amostragem reduzida. No

exemplo da FIG. 6, o pixel 615 foi selecionado, e identifica o grupo de fase de pixel na fronteira 605. O tamanho de um grupo de fase de pixel é determinado pelo fator de redução de amostragem. Para as finalidades deste documento, o pixel selecionado 615 é designado como o pixel (0, 0). O pixel identificado pelo x no grupo de fase de pixel é designado como o pixel (1, 0). O pixel identificado por um quadrado no grupo de fase de pixel é designado como o pixel (0, 1). O pixel identificado pelo círculo no grupo de fase de pixel é designado como o pixel (1, 1). Embora um grupo de fase de pixel deva ter uma relação topográfica consistente com um pixel selecionado de uma fase em particular para uma modalidade em particular, a relação topográfica pode ser diferente em outras modalidades. Por exemplo, o grupo de fase de pixel mostrado na FIG. 6 usa pixels que são os vizinhos mais próximos do pixel selecionado e os quais estão à direita e abaixo do pixel selecionado, mas um outro conjunto de quatro vizinhos mais próximos únicos poderia ser usado para um grupo de fase.

Com referência à FIG. 7, um fluxograma 700 mostra algumas etapas de um método 700 para a realização de uma transformação, de acordo com certas modalidades. Na etapa 705, a função de transformada 215 forma uma pluralidade de fases da porção do quadro de vídeo de resolução alta 206. No exemplo mostrado na FIG. 6, a pluralidade equivale a quatro. Na etapa 710, a função de transformada 215 forma a imagem de resolução baixa para compreender pixels de resolução baixa nos quais um valor de cada pixel de resolução baixa é determinado usando-se uma função de cálculo de média de valores de pixels em uma vizinhança de

pixel de cada pixel que está em uma primeira fase da pluralidade de fases. Na etapa 715, a função de transformada 215 determina cada conjunto de dados de melhoramento da pluralidade de conjuntos de dados de melhoramento como um conjunto único que compreende todos os pixels de uma outra fase além da primeira fase. Por exemplo, os conjuntos de dados de melhoramento para a porção transformada de quadro de vídeo 600 mostrada na FIG. 6 compreenderiam um conjunto de todos os pixels na porção do quadro de vídeo designada por quadrados, um conjunto de todos os pixels na porção do quadro de vídeo 600 designados por xs e um conjunto de todos os pixels na porção do quadro de vídeo 600 designada por círculos. Deve ser notado que a primeira fase pode ser qualquer fase em particular da pluralidade de fases que seja útil para o método.

De acordo com certas modalidades, a função de média é uma função de impulso, a qual determina o valor de cada pixel de resolução baixa como sendo igual ao valor do pixel correspondente da primeira fase. Em outras palavras, o cálculo da média é uma média de um número. No exemplo mostrado na FIG. 6, a imagem de resolução baixa simplesmente compreende todos os pixels designados pelos triângulos.

De acordo com certas modalidades, a função de cálculo de média determina o valor de cada pixel de resolução baixa como sendo igual à média dos valores de pixels em uma vizinhança de pixel de um pixel da primeira fase. No exemplo mostrado na FIG. 6, o valor do pixel da imagem de resolução baixa que está relacionada ao pixel 615 de uma fase em particular é determinado pelo cálculo da média dos

valores dos nove pixels na fronteira 610. De acordo com certas modalidades, a função de cálculo de média determina o valor de cada pixel de resolução baixa como sendo igual à média dos valores do pixel no grupo de fase de pixel de um pixel da primeira fase. No exemplo mostrado na FIG. 6, o valor do pixel da imagem de resolução baixa que está relacionado ao pixel 615 da primeira fase pode ser determinado pelo cálculo da média dos valores dos quatro pixels na fronteira 605.

De acordo com certas modalidades, o número de fases é quatro. Os quatro pixels identificados com referência à FIG. 6 pelo triângulo, pelo x, pelo quadrado e pelo círculo em cada grupo de fase de pixel são alternativamente identificados, respectivamente, como os pixels (0, 0), (0, 1), (1, 0) e (1, 1). Os pixels (0, 0) e (1, 1) são posicionados de forma diagonal. A função de cálculo da média determina o valor de cada pixel de resolução baixa como sendo igual à média do valor de um pixel da primeira fase (pixel 615 na FIG. 6) e um preditor, y. O preditor, y, é determinado a partir dos pixels no mesmo grupo de fase de pixel que o pixel da primeira fase, conforme se segue:

$$y = \max((0, 1), (1, 0)), \text{ quando } (1, 1) \text{ for } \geq \max((0, 1), (1, 0)) \quad (1)$$

$$y = \min((0, 1), (1, 0)), \text{ quando } (1, 1) \text{ for } \leq \min((0, 1), (1, 0)) \quad (2)$$

$$y = (0, 1) + (1, 0) - (1, 1) \text{ caso contrário} \quad (3)$$

Com referência às FIG. 8 e 9, um diagrama de blocos de um decodificador 800 e as etapas de um método 900 para decodificação são respectivamente mostrados de acordo com certas modalidades. Um fluxo de bit 805 (FIG. 8) que foi codificado de acordo com uma modalidade descrita com referência às FIG. 1 a 7 é acoplado a um separador de fluxo

de bit 810 (FIG. 8). Uma imagem pictórica codificada primária 815 e uma pluralidade de imagens pictóricas codificadas redundantes 820, 825, 830 são recebidas (905, FIG. 9) no fluxo de bit e separadas pelo separador de fluxo de bit 810. Um decodificador de imagem pictórica primária 835 (FIG. 2) gera (910, FIG. 9) uma imagem de resolução baixa 855 pela decodificação da imagem pictórica codificada primária 815. Na etapa 915 (FIG. 9), quando uma seleção tiver sido feita na etapa 930 (FIG. 9) para se usar uma imagem de resolução baixa, uma função de transformada inversa 875 (FIG. 8) pode simplesmente escalonar a informação na imagem de resolução baixa e gerar uma imagem de resolução baixa decodificada 885 na saída 880, e nenhuma etapa adicional do método 900 pode ser realizada. Na etapa 915 (FIG. 9), quando uma seleção tiver sido feita para se usar uma imagem de resolução alta, uma pluralidade de decodificadores de imagem pictórica redundante 840 (FIG. 8) gera (920, FIG. 9) uma pluralidade de conjuntos de dados de melhoramento 860, 865, 870 pela decodificação da pluralidade correspondente de imagens pictóricas codificadas redundantes 820, 825, 830. A função de transformada inversa 875 (FIG. 8) realiza (925, FIG. 9) uma função de transformada inversa da imagem de resolução baixa 855 (FIG. 2), e cada uma da pluralidade de conjuntos de dados de melhoramento 860, 865, 870 (FIG. 2), construindo a versão decodificada da posição do quadro de vídeo de resolução alta 885 a partir da qual a imagem pictórica codificada primária e uma pluralidade de conjuntos de dados de melhoramento foram geradas na saída 880, como uma parte de uma seqüência de quadro de vídeo 890. A etapa 925 pode

ser declarada de forma ligeiramente diferente como pelo menos substancialmente reconstruindo a porção do quadro de vídeo de resolução alta 206.

Quando da reconstrução da porção do quadro de vídeo de resolução alta 885, a função de transformada inversa 875 usa uma inversa da transformada que foi empregada para a codificação do quadro de vídeo de resolução alta 206. Para uma versão codificada de um quadro de vídeo de resolução alta que foi gerada usando-se uma transformação de sub-
10 banda, conforme descrito na FIG. 4, a função de transformada inversa 875 é uma operação de síntese de sub-banda que sintetiza o quadro de vídeo de resolução alta decodificado usando-se os bancos de filtro de síntese duplos, conforme é conhecido por aqueles de conhecimento
15 comum na técnica. Para um quadro de vídeo de resolução alta que foi gerado usando-se uma transformada de polifase, conforme descrito na FIG. 7, a função de transformada inversa 875 compreende uma função de cálculo de média inversa. Com referência à FIG. 10, um diagrama de blocos de
20 uma função de transformada inversa 1000 para a função de cálculo de média descrita acima com referência à FIG. 6 é mostrado, de acordo com certas modalidades. A função de transformada inversa 1000 compreende uma memória 1005, uma função de cálculo de média inversa 1010, e uma função de
25 reconstrução 1035. A pluralidade de conjuntos de dados de melhoramento 860, 865, 870 é acoplada à memória 1005, a qual armazena os conjuntos de dados de melhoramento como fases 2, 3 e 4 da imagem de resolução alta. A imagem de resolução baixa 855 é acoplada à função de cálculo de média
30 inversa 1010. Quando uma seleção é feita para uma imagem de

resolução baixa no método 900 (FIG. 9), a função de cálculo de média inversa 1010 gera a imagem de resolução baixa 855 na saída 880. A imagem de resolução baixa 855 é recebida pela função de reconstrução 1035 e é passada através dali para a saída 880. A função de cálculo de média inversa 1010 ou a função de reconstrução 1035, ou ambas, pode aplicar um escalonamento aos valores da imagem de resolução baixa 855. Quando uma seleção é feita para uma imagem de resolução alta no método 900 (FIG. 9), a função de cálculo de média inversa 1010 determina a primeira fase da porção do quadro de vídeo de resolução alta a partir da imagem de resolução baixa 855 e as outras fases da porção do quadro de vídeo de resolução alta 1020, 1025, 1030 usando métodos bem conhecidos por alguém de conhecimento comum na técnica. A primeira fase da porção do quadro de vídeo de resolução alta é acoplada a partir da saída 1015 à função de reconstrução 1035, juntamente com as outras três fases da porção do quadro de vídeo de resolução alta 1020, 1025, 1030, onde elas são combinadas para a reconstrução da porção do quadro de vídeo de resolução alta 885 na saída 880.

Será apreciado que certas etapas do método descritas aqui podem não precisar ser realizadas na ordem descrita. Apenas como um exemplo, a etapa 920 do método 900 na FIG. 9 poderia ser realizada antes da etapa 915; os resultados da etapa 920 então podem não ser usados quando a imagem de resolução baixa for escolhida.

As modalidades descritas aqui provêm novas abordagens para uma codificação de imagem / vídeo escalonável e uma codificação de descrição múltipla requerendo mudanças

substanciais em codificadores e decodificadores existentes não escalonáveis. Em particular, as técnicas podem utilizar adicionalmente a ferramenta de imagem pictórica redundante e ser facilmente aplicáveis aos sistemas de codificação de vídeo tradicionais que são projetados apenas para serem compatíveis com os perfis da ITU-T H.264 não escalonáveis.

Será apreciado que algumas modalidades podem compreender um ou mais processadores genéricos ou especializados (ou "dispositivo de processamento"), tais como microprocessadores, processadores de sinal digital, processadores personalizados e arranjos de porta programáveis de campo (FPGAs) e instruções de programa armazenadas únicas (incluindo software e firmware) que controlam um ou mais processadores a implementar, em conjunto com certos circuitos não de processador, alguns, a maioria, ou todas as funções dos métodos e/ou aparelhos descritos aqui. Alternativamente, algumas, a maioria ou todas estas funções poderiam ser implementadas por uma máquina de estado que não tem instruções de programa armazenadas, ou em um ou mais circuitos integrados específicos de aplicação (ASICs), em que cada função ou algumas combinações de certas funções são implementadas como uma lógica personalizada. Obviamente, uma combinação das duas abordagens poderia ser usada.

Mais ainda, uma modalidade pode ser implementada como um meio de armazenamento que pode ser lido em computador tendo um código que pode ser lido em computador armazenado nele para a programação de um computador (por exemplo, compreendendo um processador) para a realização de um método conforme descrito e reivindicado aqui. Os exemplos

desses meios de armazenamento que podem ser lidos em computador incluem, mas não estão limitados a um disco rígido, um CD-ROM, um dispositivo de armazenamento ótico, um dispositivo de armazenamento magnético, uma ROM (memória de acesso randômico), uma PROM (memória de acesso randômico programável), uma EPROM (memória de acesso randômico programável apagável), uma EEPROM (memória de acesso randômico programável apagável eletricamente) e uma memória flash. Ainda, é esperado que alguém de conhecimento comum, não obstante possivelmente com esforço significativo e motivado por muitas escolhas de projeto, por exemplo, tempo disponível, tecnologia atual e considerações econômicas, quando guiado pelos conceitos e princípios mostrados aqui, seja prontamente capaz de gerar essas instruções de software e esses programas e ICs com uma experimentação mínima.

No relatório descritivo precedente, as modalidades específicas foram descritas. Contudo, alguém de conhecimento comum na técnica aprecia que várias modificações e mudanças podem ser feitas, sem que se desvie do escopo da invenção, conforme estabelecido nas reivindicações abaixo. Assim sendo, o relatório descritivo e as FIG. devem ser considerados em um sentido ilustrativo, ao invés de restritivo, e se pretende que todas essas modificações estejam incluídas no escopo dos presentes ensinamentos. Os benefícios, as vantagens e soluções dos problemas, e quaisquer elementos que possam fazer com que qualquer benefício, vantagem ou solução ocorra ou se torne mais pronunciado não devem ser construídos como recursos ou elementos críticos, requeridos ou essenciais de todas ou

quaisquer reivindicações. A invenção é definida unicamente pelas reivindicações em apenso incluindo quaisquer emendas feitas durante a pendência deste pedido e todos os equivalentes daquelas reivindicações, conforme emitidas.

5 Mais ainda, neste documento, termos relacionais tais como primeiro e segundo, topo e fundo, e similares podem ser usados unicamente para a distinção de uma entidade ou ação de outra entidade ou ação sem necessariamente requerer ou implicar em qualquer relação real como essa ou ordem

10 entre essas entidades ou ações. Os termos "compreende", "compreendendo", "tem", "tendo", "inclui", "incluindo", "contém", "contendo" ou qualquer outra variação dos mesmos são pretendidos para cobrirem uma inclusão não exclusiva, de modo que um processo, método, artigo ou aparelho que

15 compreenda, tenha, inclua, contenha uma lista de elementos não inclua apenas aqueles elementos, mas possa incluir outros elementos não expressamente listados ou inerentes para esse processo, método, artigo ou aparelho. Um elemento seguido por "compreende... um", "tem... um", "inclui... um",

20 "contém... um" não impede, sem maiores restrições, a existência de elementos idênticos adicionais no processo, método, artigo ou aparelho que compreende, tem, inclui, contém o elemento. Os termos "um" e "uma" são definidos como um ou mais, a menos que explicitamente declarado de

25 outra forma aqui. Os termos "substancialmente", "essencialmente", "aproximadamente", "em torno de" ou qualquer outra versão dos mesmos são definidos como sendo próximos de, conforme entendido por alguém de conhecimento comum na técnica, e em uma modalidade não limitativa o

30 termo é definido como estando em 10%, em uma outra

modalidade em 5%, em uma outra modalidade em 1% e em uma outra modalidade em 0,5%. O termo "acoplado" conforme usado aqui é definido como conectado, embora não necessariamente de forma direta e não necessariamente de forma mecânica. Um dispositivo ou uma estrutura que é "configurada" de certa forma é configurado pelo menos daquela forma, mas também pode ser configurado de formas que não são listadas.

O Resumo da Exposição é provido para se permitir que o leitor rapidamente avalie a natureza da exposição técnica.

10 É submetido com o entendimento que não será usado para interpretar ou limitar o escopo ou o significado das reivindicações. Além disso, na Descrição Detalhada precedente, pode ser visto que vários recursos são agrupados em conjunto em várias modalidades para fins de
15 otimização da exposição. Este método de exposição não é para ser interpretado como refletindo uma intenção que as modalidades reivindicadas requerem mais recursos do que expressamente recitado em cada reivindicação. Ao invés disso, conforme as reivindicações a seguir refletem, o
20 assunto inventivo está em menos do que todos os recursos de uma única modalidade mostrada. Assim, as reivindicações a seguir desse modo são incorporadas na Descrição detalhada, com cada reivindicação sendo auto-suficiente como um assunto reivindicado separadamente.

REIVINDICAÇÕES

1. Método de decodificação de vídeo, caracterizado por compreender:

a geração de uma imagem de resolução baixa
5 decodificada pela decodificação de uma imagem pictórica codificada primária codificada de acordo com a Codificação Avançada da ITU-T H.264 Especificação 03/2005;

a geração de uma pluralidade de conjuntos de dados de
10 melhoramento decodificados pela decodificação de uma pluralidade correspondente de imagens pictóricas codificadas redundantes codificadas de acordo com a Codificação Avançada da ITU-T H.264 Especificação 03/2005;
e

a realização de uma transformação inversa da imagem de
15 resolução baixa decodificada e da pluralidade de conjuntos de dados de melhoramento decodificados para a construção de uma versão decodificada de uma porção de um quadro de vídeo de resolução alta;

em que a imagem de resolução baixa compreender pixels
20 de resolução baixa nos quais um valor de cada pixel de resolução baixa é determinado usando-se uma função de cálculo de média de valores de pixels definidos em uma vizinhança de pixel de cada pixel em uma fase em particular de uma pluralidade de fases da porção do quadro de vídeo de
25 resolução alta, e cada um da pluralidade de conjuntos de dados de melhoramento ser um conjunto único que compreende todos os pixels de uma outra fase da porção do quadro de vídeo de resolução alta além da fase em particular, onde a transformação inversa compreende:

30 a regulagem de cada outra fase do quadro de vídeo

de resolução alta além da fase em particular para um conjunto correspondente da pluralidade de conjuntos de dados de melhoramento;

5 a determinação da fase em particular da porção do vídeo de resolução alta a partir da imagem de resolução baixa e das outras fases da porção do vídeo de resolução alta; e

10 a combinação da fase em particular e das outras fases para a construção da versão decodificada do quadro de vídeo de resolução alta; e

em que um grupo de fase de pixel para um pixel inclui um pixel de cada grupo de fase localizado na mesma posição em relação ao pixel na imagem de resolução baixa, em que o número de fases é quatro, em que os quatro pixels em cada grupo de fase de pixel são identificados como (0, 0), (0, 1), (1, 0) e (1, 1), onde os pixels (0, 0) e (1, 1) são posicionados de forma diagonal, e onde a função de cálculo da média determina o valor de cada pixel de resolução baixa como sendo igual à média do valor de um pixel da fase em particular e um preditor, y , em que o pixel da fase em particular é um pixel da fase (0, 0), onde o preditor, y , é determinado a partir dos pixels no mesmo grupo de fase de pixel que o pixel da fase em particular, conforme a seguir:

25 $y = \max ((0, 1), (1, 0)),$ quando (1, 1) for $\geq \max ((0, 1), (1, 0))$

$y = \min ((0, 1), (1, 0)),$ quando (1, 1) for $\leq \min ((0, 1), (1, 0))$

$y = (0, 1) + (1, 0) - (1, 1)$ caso contrário.

30 2. Método de decodificação de vídeo, de acordo com a reivindicação 1, caracterizado pelo fato de a vizinhança de

um pixel ser um grupo de fase de pixel.

3. Aparelho de decodificação, caracterizado por compreender:

um decodificador que recupera uma imagem de resolução
5 baixa a partir de uma imagem pictórica codificada primária
em um fluxo de bit e recupera uma pluralidade de conjuntos
de dados de melhoramento a partir de uma pluralidade
correspondente de imagens pictóricas codificadas
redundantes no fluxo de bit, a imagem pictórica codificada
10 primária e a pluralidade correspondente de imagens
pictóricas codificadas redundantes codificadas de acordo
com a Codificação Avançada da ITU-T H.264 Especificação
03/2005; e

um pós-processador que realiza uma transformada
15 inversa na imagem de resolução baixa e na pluralidade de
conjuntos de dados de melhoramento, construindo uma versão
decodificada de uma porção do quadro de vídeo de resolução
alta;

em que a imagem de resolução baixa compreender pixels
20 de resolução baixa nos quais um valor de cada pixel de
resolução baixa é determinado usando-se uma função de
cálculo de média de valores de pixels definidos em uma
vizinhança de pixel de cada pixel em uma fase em particular
de uma pluralidade de fases da porção do quadro de vídeo de
25 resolução alta, e cada um da pluralidade de conjuntos de
dados de melhoramento ser um conjunto único que compreende
todos os pixels de uma outra fase da porção do quadro de
vídeo de resolução alta além da fase em particular, onde o
dito pós-processador é configurado para executar a função
30 de transformada inversa compreende:

uma memória para a regulagem de cada outra fase do quadro de vídeo de resolução alta além da fase em particular para um conjunto correspondente da pluralidade de conjuntos de dados de melhoramento;

5 uma função de cálculo de média inversa que determina a fase em particular da porção do vídeo de resolução alta a partir da imagem de resolução baixa e das outras fases da porção do quadro de vídeo de resolução alta; e

10 um reconstrutor que combina a fase em particular e as outras fases na versão decodificada do quadro de vídeo de resolução alta; e

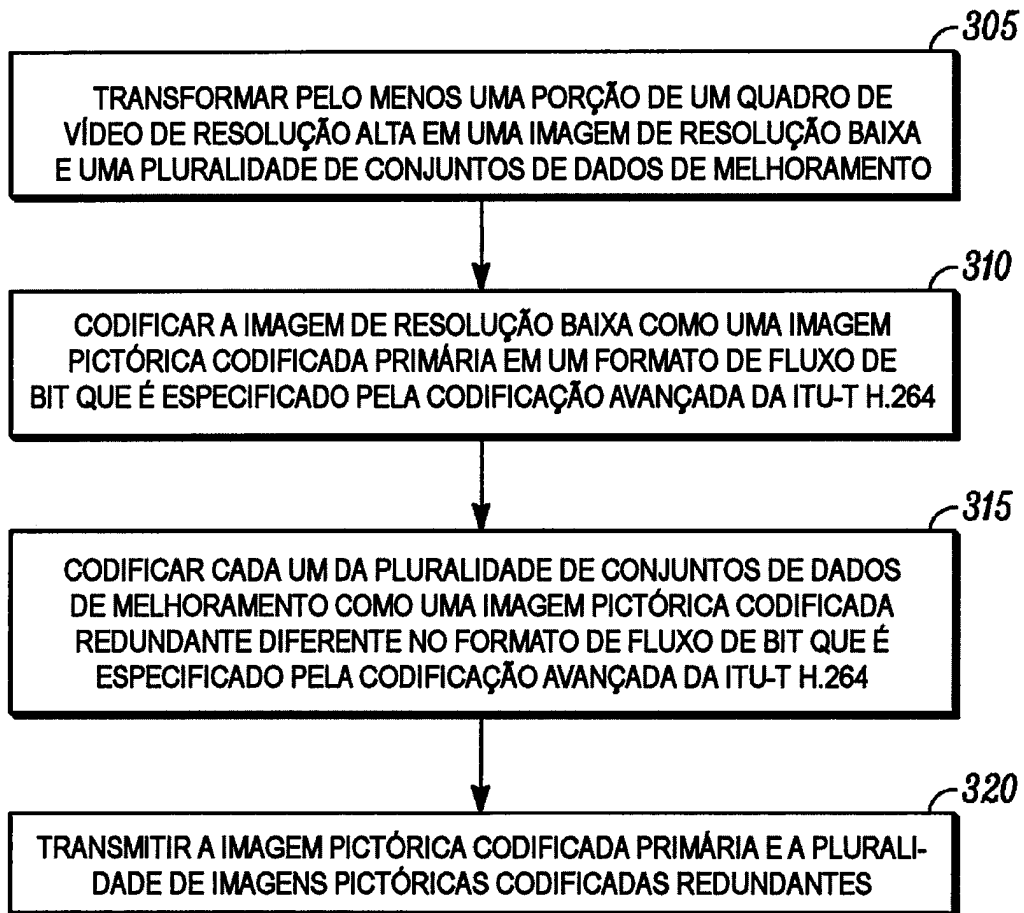
em que um grupo de fase de pixel para um pixel inclui um pixel de cada grupo de fase localizado na mesma posição em relação ao pixel na imagem de resolução baixa, em que o número de fases é quatro, em que os quatro pixels em cada grupo de fase de pixel são identificados como (0, 0), (0, 1), (1, 0) e (1, 1), onde os pixels (0, 0) e (1, 1) são posicionados de forma diagonal, e onde a função de cálculo da média determina o valor de cada pixel de resolução baixa como sendo igual à média do valor de um pixel da fase em particular e um preditor, y , em que o pixel da fase em particular é um pixel da fase (0, 0), onde o preditor, y , é determinado a partir dos pixels no mesmo grupo de fase de pixel que o pixel da fase em particular, conforme a seguir:

$y = \max ((0, 1), (1, 0))$, quando (1, 1) for $\geq \max ((0, 1), (1, 0))$

$y = \min ((0, 1), (1, 0))$, quando (1, 1) for $\leq \min ((0, 1), (1, 0))$

30 $y = (0, 1) + (1, 0) - (1, 1)$ caso contrário.

4. Aparelho de decodificação, de acordo com a reivindicação 3, caracterizado pelo fato de a vizinhança de um pixel ser um grupo de fase de pixel.

300*FIG. 3*

405

REALIZAR UMA TRANSFORMAÇÃO DE SUB-BANDA DA PORÇÃO DO QUADRO DE VÍDEO DE RESOLUÇÃO ALTA PARA A GERAÇÃO DAS SUB-BANDAS LL, HL, LH E HH DA PORÇÃO DO QUADRO DE VÍDEO, ONDE A SUB-BANDA LL É A IMAGEM DE RESOLUÇÃO BAIXA E AS SUB-BANDAS HL, LH E HH SÃO A PLURALIDADE DE CONJUNTOS DE DADOS DE MELHORAMENTO

FIG. 4

600

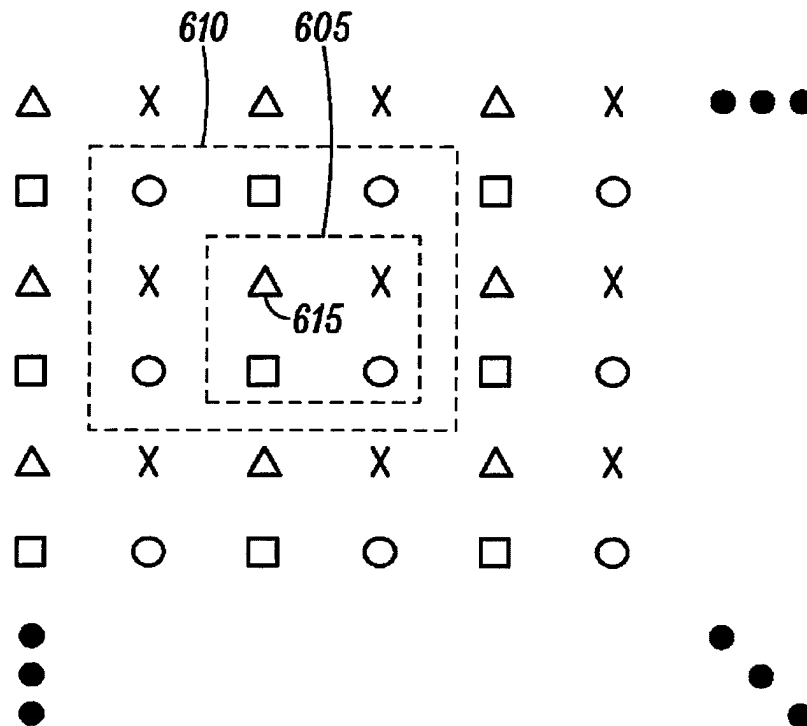


FIG. 6

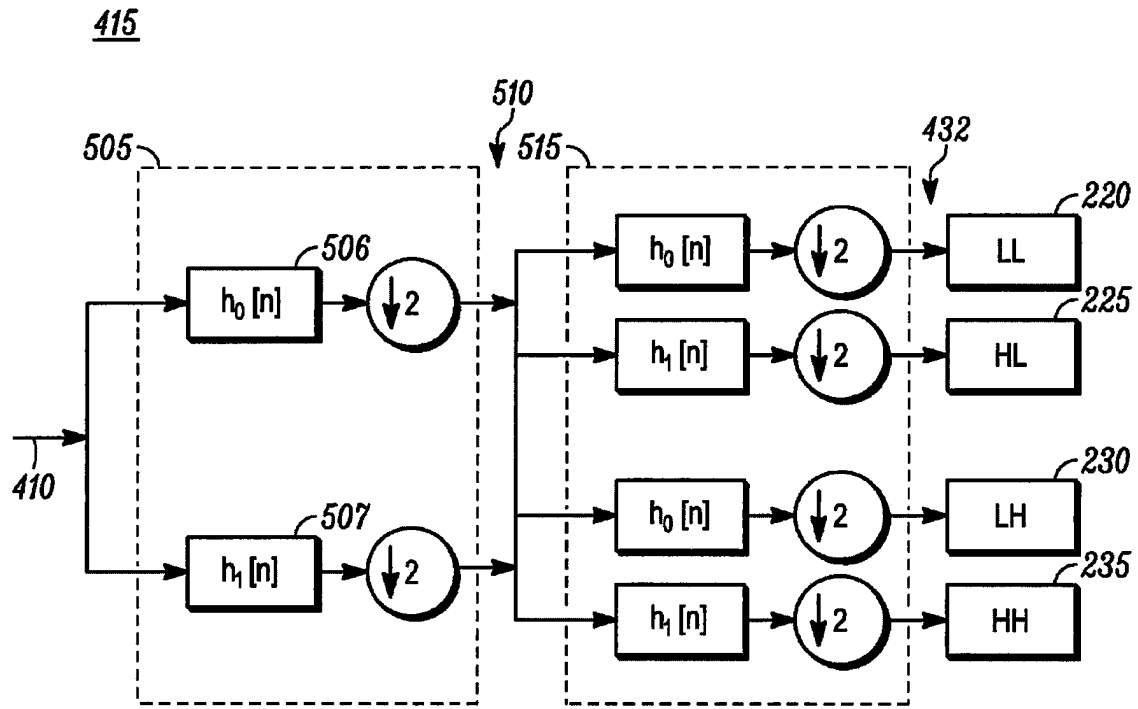


FIG. 5

700

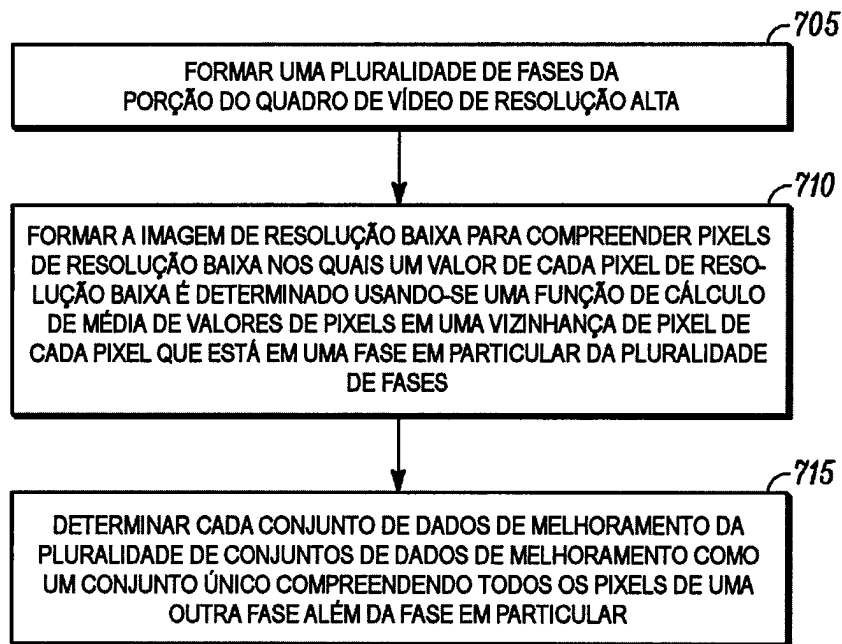


FIG. 7

800

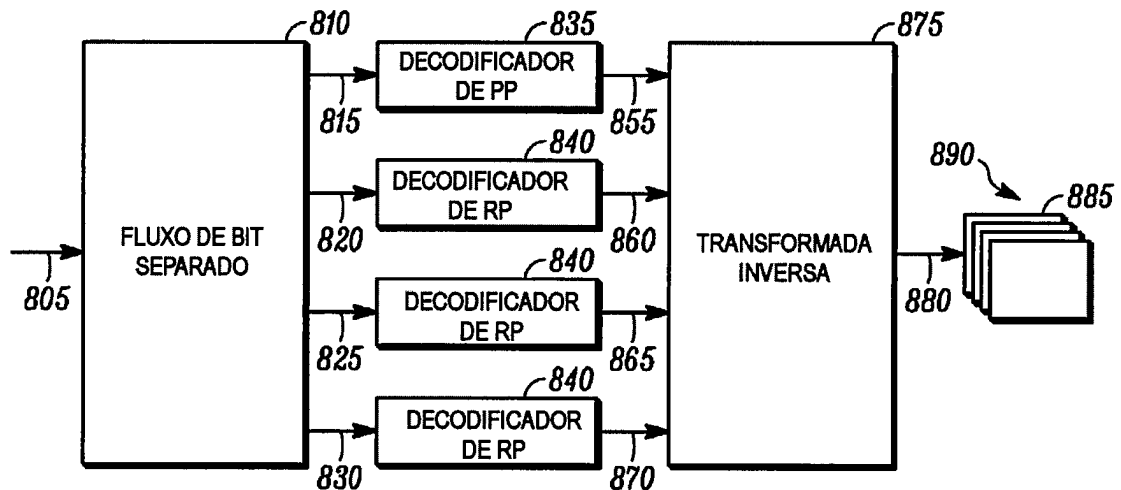


FIG. 8

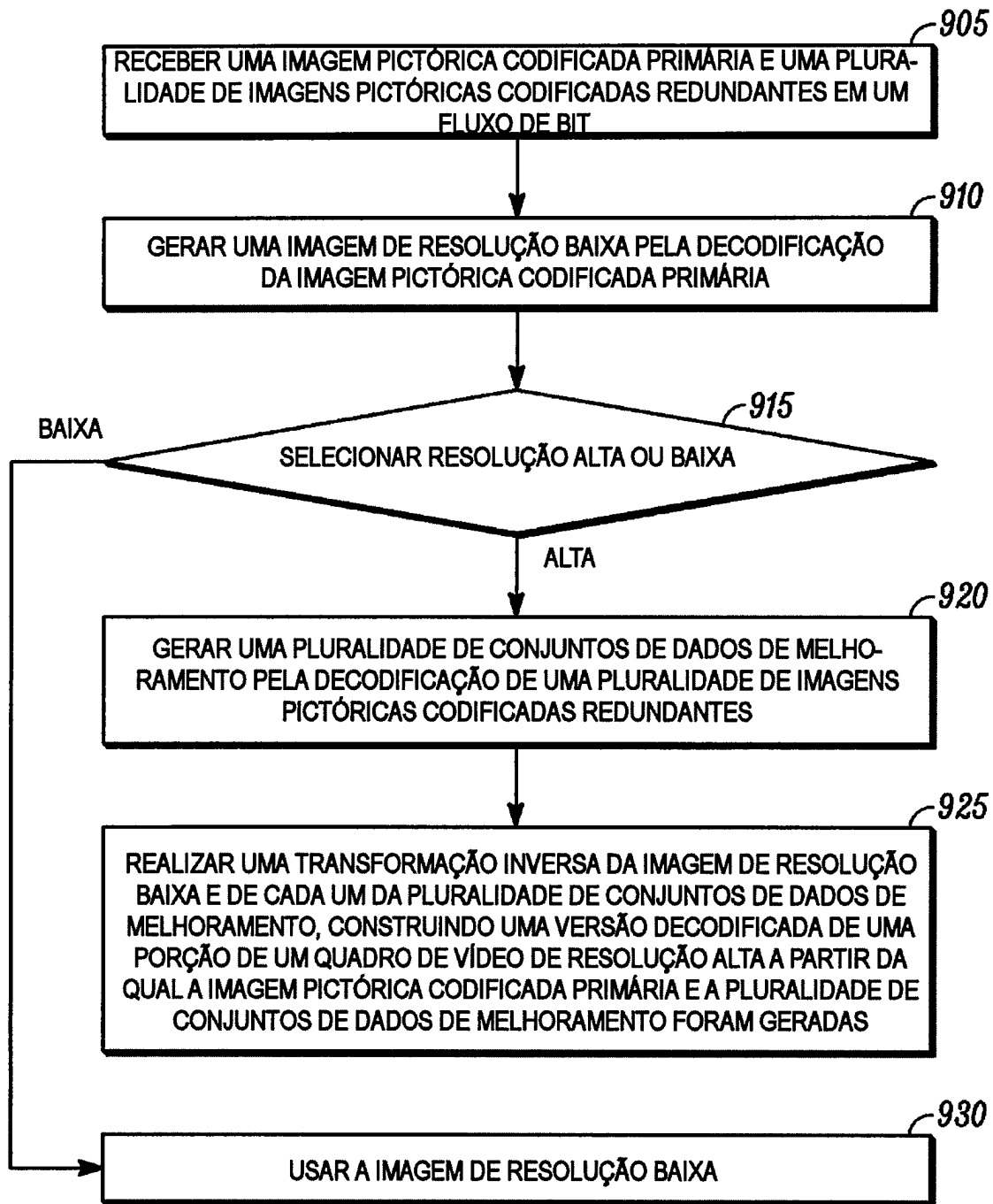
900

FIG. 9

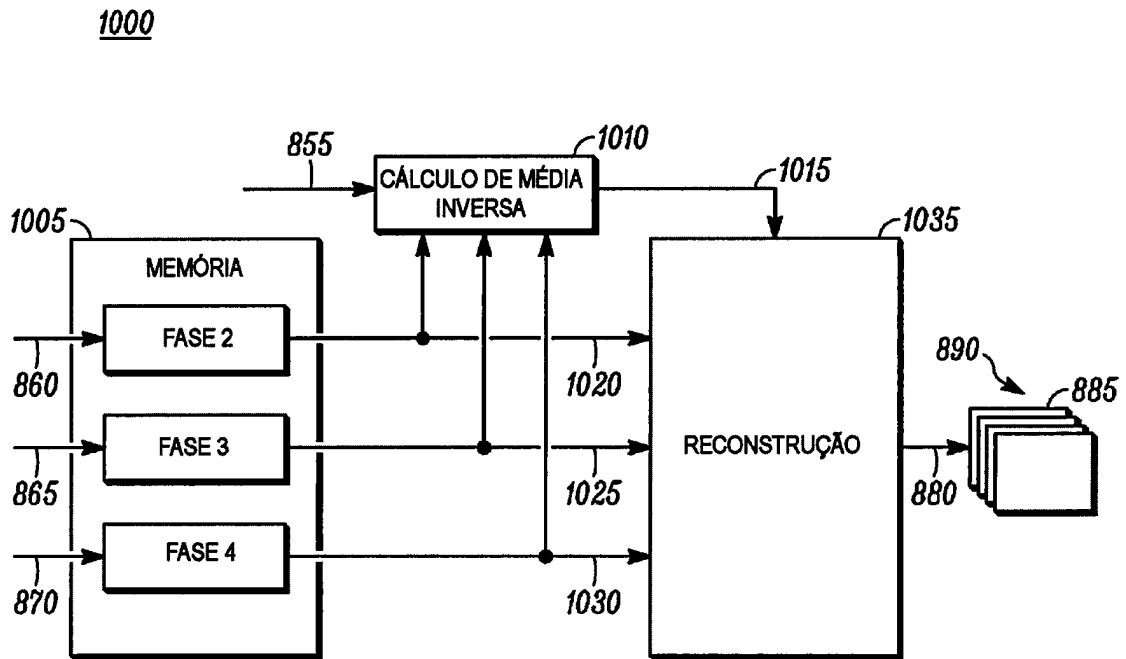


FIG. 10