

República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial.

(21) **PI0706434-9 A2**



(22) Data de Depósito: 08/01/2007
(43) Data da Publicação: 29/03/2011
(RPI 2099)

(51) *Int.Cl.:*
H04N 7/26
H04N 7/36
H04N 13/00

(54) Título: **MÉTODO E APARELHO PARA CODIFICAÇÃO DE VÍDEO MULTIVISUALIZAÇÃO**

(30) Prioridade Unionista: 09/01/2006 US 60/757,289

(73) Titular(es): Thomson Licensing

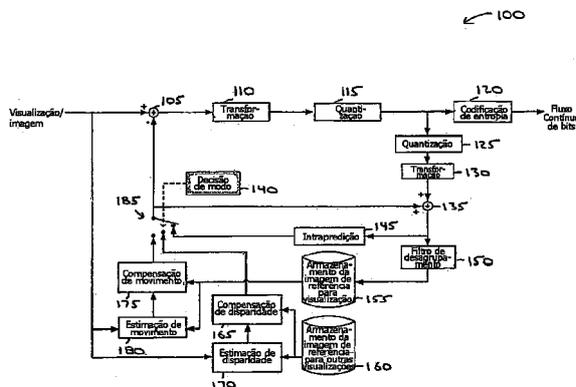
(72) Inventor(es): Cristina Gomila, Peng Yin, Yeping Su

(74) Procurador(es): Nellie Anne Daniel Shores

(86) Pedido Internacional: PCT US2007000199 de 08/01/2007

(87) Publicação Internacional: WO 2007/081756 de 19/07/2007

(57) **Resumo:** MÉTODOS E APARELHO PARA CODIFICAÇÃO DE VÍDEO MULTIVISUALIZAÇÃO São fornecidos métodos e aparelho para codificação de vídeo multivisualização. Um codificador de vídeo inclui um codificador (100) para codificar um bloco em uma imagem pela escolha entre prognóstico temporal e prognóstico de visualização cruzada para habilitar um prognóstico para o bloco. A imagem é uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos de visualização em relação a uma mesma cena ou a uma cena similar. A imagem representa um dos diferentes pontos de visualização. Uma sintaxe de alto nível é usada para indicar o uso do prognóstico de visualização cruzada para o bloco.



"MÉTODOS E APARELHO PARA CODIFICAÇÃO DE VÍDEO MULTIVISUALIZAÇÃO"

REFERÊNCIA CRUZADA A PEDIDOS RELACIONADOS

Este pedido reivindica o benefício do pedido provisório US 60/757.289, intitulado "Multi-View Video Coding System", depositado em 9 de janeiro de 2006, que é aqui incorporado pela referência. Além do mais, este pedido está relacionado ao pedido não provisório, Documento Judicial PU060118, intitulado "Methods and Apparatus for Multi-View Video Coding", do mesmo requerente, incorporado pela referência em sua íntegra e depositado simultaneamente.

CAMPO DA INVENÇÃO

A presente invenção diz respeito, no geral, a codificadores e decodificadores de vídeo e, mais particularmente, a métodos e aparelho para Codificação de Vídeo Multivisualização.

ANTECEDENTES DA INVENÇÃO

Codificação de vídeo multivisualização (MVC) é a estrutura de compressão para a codificação das seqüências multivisualização. Uma seqüência de Codificação de Vídeo Multivisualização (MVC) é um conjunto de duas ou mais seqüências de vídeo que captura a mesma cena de um diferente ponto de visualização.

Reconhece-se amplamente que a Codificação de Vídeo Multivisualização é uma tecnologia chave que serve a uma ampla variedade de aplicações, incluindo aplicações com ponto de visualização livre e aplicações de vídeo em 3D, entretenimento doméstico e vigilância. Nestas aplicações multivisu-

alização, a quantidade de dados de vídeo envolvida é enorme. Assim, existe uma necessidade de tecnologias de compressão eficientes para melhorar a eficiência de codificação das atuais soluções de codificação de vídeo que realizam a difusão simultânea de visualizações independentes.

Nos últimos anos, foram feitos muitos esforços no projeto de métodos eficientes para a compressão de vídeo estereoscópico. Métodos de compressão monoscópica convencionais podem ser aplicados independentemente às visualizações esquerda e direita de um par de imagem estéreo. Entretanto, maiores taxas de compressão podem ser alcançadas se a alta correlação entre as visualizações for explorada.

Considerando uma abordagem da tecnologia anterior na qual ambas as visualizações de um par de imagem estereoscópica são codificadas, um Perfil Multivisualização (MVP) foi definido no padrão Grupo de Especialistas em Imagens em Movimento-2 (MPEG-2) da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional (ISO/IEC) para transmitir um par de sinais de vídeo. O MVP se baseia em uma abordagem de representação de sinal multicamadas de maneira tal que uma visualização (frequentemente, a visualização esquerda) é atribuída a uma camada base, e a outra visualização é atribuída a uma camada de melhoria. A codificação monoscópica com as mesmas ferramentas do Perfil Principal (MP) é aplicada na camada base. A camada de melhoria é codificada usando ferramentas de escalabilidade temporal e um prognóstico híbrido do movimento e dos campos de disparidade.

Em métodos da tecnologia anterior relacionados ao padrão de Codificação Avançada de Vídeo (AVC) parte 10 do Grupo de Especialistas de Imagens em Movimento-4 (MPEG-4) da Organização Internacional para a Padronização / Comissão Eletrotécnica Internacional (ISO/IEC) / recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações (ITU-T) (doravante o "padrão MPEG-4 AVC"), a codificação estereoscópica de vídeo pode ser realizada de duas maneiras diferentes: (i) como um caso em particular de codificação de imagem entrelaçada, em que todos os campos de uma paridade em particular são atribuídos à visualização esquerda e todos os campos da paridade oposta são considerados a visualização direita do conteúdo de visualização estéreo; ou, alternativamente, (ii) pela alternância de quadros das visualizações esquerda e direita para criar uma única sequência monoscópica de vídeo. Uma mensagem de informação de melhoria complementar (SEI) de estereovisão fornece uma indicação ao decodificador se a sequência de vídeo codificada representa conteúdo estereoscópico ou não e qual método foi usado para codificar o conteúdo correspondente.

Estes métodos previamente conhecidos exigem mínimas modificações das técnicas de codificação monoscópica existentes. Entretanto, eles mostram uma capacidade limitada de reduzir a redundância existente entre as duas visualizações em um par estereoscópico. Em decorrência disto, a codificação da visualização estéreo resulta em um grande sobreprocessamento, se comparado com a codificação de uma única visualização monoscópica. Além do mais, não há suporte ante-

rior para a codificação de mais do que duas visualizações de câmara.

SUMÁRIO DA INVENÇÃO

Estes e outros inconvenientes e desvantagens da tecnologia anterior são abordados pela presente invenção, que é direcionada a métodos e aparelho para Codificação de Vídeo Multivisualização.

De acordo com um aspecto da presente invenção, é fornecido um codificador de vídeo. O codificador de vídeo inclui um codificador para codificar um bloco em uma imagem pela escolha entre prognóstico temporal e prognóstico de visualização cruzada para habilitar um prognóstico para o bloco. A imagem é uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos de visualização em relação a uma mesma cena ou a uma cena similar. A imagem representa um dos diferentes pontos de visualização. Uma sintaxe de alto nível é usada para indicar o uso do prognóstico de visualização cruzada para o bloco.

De acordo com um outro aspecto da presente invenção, é fornecido um codificador de vídeo. O codificador de vídeo inclui um codificador para codificar um bloco em uma imagem usando um vetor de disparidade. A imagem corresponde ao conteúdo de vídeo multivisualização e é codificada para fornecer um fluxo contínuo de bits resultante, em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação

H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma extensão desta.

De acordo com um ainda outro aspecto da presente invenção, é fornecido um codificador de vídeo. O codificador de vídeo inclui um codificador para codificar um bloco em uma imagem que corresponde ao conteúdo de vídeo multivisualização com base em um preditor do vetor de disparidade que usa pelo menos um de um vetor de disparidade temporalmente co-localizado e de um vetor de disparidade vizinho de um bloco temporalmente co-localizado.

De acordo com um ainda outro aspecto da presente invenção, é fornecido um método para codificação de vídeo. O método inclui codificar um bloco em uma imagem pela escolha entre o prognóstico temporal e o prognóstico de visualização cruzada para habilitar um prognóstico para o bloco. A imagem é uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos de visualização em relação em relação a uma mesma cena ou a uma cena similar. A imagem representa um dos diferentes pontos de visualização. Uma sintaxe de alto nível é usada para indicar o uso do prognóstico de visualização cruzada para o bloco.

De acordo com um aspecto adicional da presente invenção, é fornecido um método para codificação de vídeo. O método inclui codificar um bloco em uma imagem usando um vetor de disparidade em que a imagem corresponde ao conteúdo de vídeo multivisualização e é codificada para fornecer um fluxo contínuo de bits resultante, em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo

de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma extensão desta.

De acordo com um aspecto ainda adicional da presente invenção, é fornecido um método para codificação de vídeo. O método inclui codificar um bloco em uma imagem correspondente ao conteúdo de vídeo multivisualização com base em um preditor do vetor de disparidade que usa pelo menos um de um vetor de disparidade temporalmente co-localizado e de um vetor de disparidade vizinho de um bloco temporalmente co-localizado.

De acordo com um aspecto adicional da presente invenção, é fornecido um decodificador de vídeo. O decodificador de vídeo inclui um decodificador para decodificar um bloco em uma imagem pela escolha entre prognóstico temporal e prognóstico de visualização cruzada para habilitar um prognóstico para o bloco. A imagem é uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos de visualização em relação a uma mesma cena ou a uma cena similar. A imagem representa um dos diferentes pontos de visualização. Uma sintaxe de alto nível é lida para determinar o uso do prognóstico de visualização cruzada para o bloco.

De acordo com um aspecto ainda adicional da presente invenção, é fornecido um decodificador de vídeo. O decodificador de vídeo inclui um decodificador para decodifi-

car um bloco em uma imagem usando um vetor de disparidade. A imagem corresponde ao conteúdo de vídeo multivisualização e é decodificada a partir de um fluxo contínuo de bits em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma extensão desta.

10 De acordo com um outro aspecto da presente invenção, é fornecido um decodificador de vídeo. O decodificador de vídeo inclui um decodificador para decodificar um bloco em uma imagem correspondente ao conteúdo de vídeo multivisualização com base em um preditor do vetor de disparidade que
15 usa pelo menos um de um vetor de disparidade temporalmente co-localizado e de um vetor de disparidade vizinho de um bloco temporalmente co-localizado.

De acordo com um ainda outro aspecto da presente invenção, é fornecido um método para decodificação de vídeo.
20 O método inclui decodificar um bloco em uma imagem pela escolha entre prognóstico temporal e prognóstico de visualização cruzada para habilitar um prognóstico para o bloco. A imagem é uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos
25 de visualização em relação a uma mesma cena ou a uma cena similar. A imagem representa um dos diferentes pontos de visualização. Uma sintaxe de alto nível é lida para determinar o uso do prognóstico de visualização cruzada para o bloco.

De acordo com um ainda outro aspecto da presente invenção, é fornecido um método para decodificação de vídeo. O método inclui decodificar um bloco em uma imagem usando um vetor de disparidade. A imagem corresponde ao conteúdo de vídeo multivisualização e é decodificada a partir de um fluxo contínuo de bits em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma extensão desta (400).

De acordo com um aspecto adicional da presente invenção, é fornecido um método para decodificação de vídeo. O método inclui decodificar um bloco em uma imagem correspondente ao conteúdo de vídeo multivisualização com base em um preditor do vetor de disparidade que usa pelo menos um de um vetor de disparidade temporalmente co-localizado e de um vetor de disparidade vizinho de um bloco temporalmente co-localizado.

Estes e outros aspectos, recursos e vantagens da presente invenção ficaram aparentes a partir da seguinte descrição detalhada das modalidades exemplares, que deve ser lida em conjunto com os desenhos anexos.

25 DESCRIÇÃO RESUMIDA DOS DESENHOS

A presente invenção pode ser mais bem entendida de acordo com as seguintes figuras exemplares, nas quais:

a figura 1 é um diagrama de blocos para um codifi-

cador de Codificação de Vídeo Multivisualização (MVC) no qual os presentes princípios podem ser aplicados, de acordo com uma modalidade dos presentes princípios;

5 a figura 2 é um diagrama de blocos para um decodificador de Codificação de Vídeo Multivisualização (MVC) no qual os presentes princípios podem ser aplicados, de acordo com uma modalidade dos presentes princípios;

10 a figura 3 é um fluxograma para um método exemplar para a construção da lista de referência para conteúdo de vídeo multivisualização, de acordo com uma modalidade dos presentes princípios;

15 a figura 4 é um fluxograma para um método exemplar para realizar uma decisão de modo de visualização temporal / cruzada, de acordo com uma modalidade dos presentes princípios;

20 a figura 5 é um fluxograma para um método exemplar para processar vetores de movimento e de disparidade para a mesma fatia correspondente ao conteúdo de vídeo multivisualização, de acordo com uma modalidade dos presentes princípios; e

a figura 6 é um fluxograma para um outro método para processar vetores de movimento e de disparidade para conteúdo de vídeo multivisualização, de acordo com uma modalidade dos presentes princípios.

25 DESCRIÇÃO DETALHADA

A presente invenção é direcionada a métodos e aparelho para Codificação de Vídeo Multivisualização.

A presente descrição ilustra os princípios da pre-

sente invenção. Assim, percebe-se que versados na técnica podem conceber vários arranjos que, embora não explicitamente descritos ou mostrados aqui, incorporam os princípios da invenção e estão incluídos no seu espírito e escopo.

5 Pretende-se que todos os exemplos e linguagem condicional aqui citados tenham propósitos pedagógicos para auxiliar o leitor no entendimento dos princípios da invenção e dos conceitos contribuídos pelo inventor para desenvolver a tecnologia, e devem ser interpretados sem limitação a tais
10 exemplos e condições especificamente citados.

Além do mais, pretende-se que todas as declarações que aqui citam princípios, aspectos e modalidades da invenção, bem como seus exemplos específicos, abranjam tanto seus equivalentes estruturais quanto funcionais. Adicionalmente,
15 pretende-se que tais equivalentes incluam tanto equivalentes atualmente conhecidos quanto equivalentes desenvolvidos no futuro, isto é, qualquer elemento desenvolvido que realize a mesma função, independente da estrutura.

Assim, por exemplo, versados na técnica percebem
20 que os diagramas de blocos aqui apresentados representam visualizações conceituais do sistema de circuitos ilustrativo que incorpora os princípios da invenção. Similarmente, percebe-se que todos os fluxogramas, diagramas de fluxo, diagramas de transição de estado, pseudocódigos e congêneres
25 representam vários processos que podem ser substancialmente representados em mídia legível por computador e, portanto, executados por um computador ou processador, caso tal computador ou processador sejam explicitamente mostrados ou não.

As funções dos vários elementos mostrados nas figuras podem ser fornecidas por meio do uso de hardware dedicado, bem como de hardware que pode executar software em conjunto com software apropriado. Quando fornecidas por um processador, as funções podem ser fornecidas por um único processador dedicado, por um único processador compartilhado ou por uma pluralidade de processadores individuais, alguns dos quais podem ser compartilhados. Além do mais, o uso explícito do termo "processador" ou "controlador" não deve ser interpretado dizendo respeito exclusivamente a hardware que pode executar software, e pode incluir implicitamente, mas sem limitações, hardware processador de sinal digital ("DSP"), memória exclusiva de leitura ("ROM") para armazenar software, memória de acesso aleatório ("RAM") e armazenamento não volátil.

Outro hardware, convencional e/ou personalizado, também pode ser incluído. Similarmente, todas as chaves mostradas nas figuras são somente conceituais. Sua função pode ser realizada por meio da operação de lógica de programa, por meio de lógica dedicada, por meio da interação do controle do programa e da lógica dedicada, ou mesmo manualmente, a técnica em particular sendo selecionada pelo implementador, como é mais especificamente entendido a partir do contexto.

Nas reivindicações anexas, pretende-se que todos os elementos expressos como um dispositivo para realizar uma função específica abrangam qualquer maneira de realizar aquela função, incluindo, por exemplo, a) uma combinação de

elementos de circuito que realiza aquela função ou b) software em qualquer forma, incluindo, portanto, software embarcado, microcódigo ou congêneres, combinados com sistema de circuitos apropriado para executar aquele software para realizar a função. A invenção definida por tais reivindicações fica residente no fato de que as funcionalidades fornecidas pelos vários dispositivos citados são combinadas e articuladas da maneira que as reivindicações exigem. Assim, considera-se que todos os dispositivos que podem fornecer aquelas funcionalidades são equivalentes àqueles aqui mostrados.

Nesta especificação, a referência a "uma modalidade" ou "a modalidade" dos presentes princípios significa que um recurso, estrutura, característica e assim por diante em particular descrito em conjunto com a modalidade está incluído em pelo menos uma modalidade dos presentes princípios. Assim, nem todas as aparências da frase "em uma modalidade" ou "na modalidade" que aparecem em vários locais por toda a especificação, se referem, necessariamente à mesma modalidade.

Voltando para a figura 1, um codificador de Codificação de Vídeo Multivisualização (MVC) exemplar é indicado, no geral, pelo número de referência 100. O codificador 100 inclui um combinador 105 com uma saída conectada em comunicação de sinal com uma entrada de um transformador 110. Uma saída do transformador 110 é conectada em comunicação de sinal com uma entrada do quantizador 115. Uma saída do quantizador 115 é conectada em comunicação de sinal com uma entrada de um codificador de entropia 120 e com uma entrada de

um quantizador invertido 125. Uma saída do quantizador invertido 125 é conectada em comunicação de sinal com uma entrada de um transformador invertido 130. Uma saída do transformador invertido 130 é conectada em comunicação de sinal com uma primeira entrada não invertida de um combinador 135. Uma saída do combinador 135 é conectada em comunicação de sinal com uma entrada de um intrapreditor 145 e com uma entrada de um filtro de desagrupamento 150. Uma saída do filtro de desagrupamento 150 é conectada em comunicação de sinal com uma entrada de um armazenamento da imagem de referência 155 (para a visualização i). Uma saída do armazenamento da imagem de referência 155 é conectada em comunicação de sinal com uma primeira entrada de um compensador de movimento 175 e com uma primeira entrada de um estimador de movimento 180. Uma saída do estimador de movimento 180 é conectada em comunicação de sinal com uma segunda entrada do compensador de movimento 175.

Uma saída de um armazenamento da imagem de referência 160 (para outras visualizações) é conectada em comunicação de sinal com uma primeira entrada de um estimador de disparidade / iluminação 170 e com uma primeira entrada de um compensador de disparidade / iluminação 165. Uma saída do estimador de disparidade / iluminação 170 é conectada em comunicação de sinal com uma segunda entrada do compensador de disparidade / iluminação 165.

Uma saída do decodificador de entropia 120 fica disponível como uma saída do codificador 100. Uma entrada não invertida do combinador 105 fica disponível como uma en-

trada do codificador 100 e é conectada em comunicação de sinal com uma segunda entrada do estimador de disparidade / iluminação 170 e com uma segunda entrada do estimador de movimento 180. Uma saída de uma chave 185 é conectada em comunicação de sinal com uma segunda entrada não invertida do combinador 135 e com uma entrada invertida do combinador 105. A chave 185 inclui uma primeira entrada conectada em comunicação de sinal com uma saída do compensador de movimento 175, uma segunda entrada conectada em comunicação de sinal com uma saída do compensador de disparidade / iluminação 165, e uma terceira entrada conectada em comunicação de sinal com uma saída do intrapreditor 145.

Um módulo de decisão de modo 140 tem uma saída conectada na chave 185 para controlar qual entrada é selecionada pela chave 185.

Voltando para a figura 2, um decodificador de Codificação de Vídeo Multivisualização é indicado, no geral, pelo número de referência 200. O decodificador 200 inclui um decodificador de entropia 205 com uma saída conectada em comunicação de sinal com uma entrada de um quantizador invertido 210. Uma saída do quantizador invertido é conectada em comunicação de sinal com uma entrada de um transformador invertido 215. Uma saída do transformador invertido 215 é conectada em comunicação de sinal com uma primeira entrada não invertida de um combinador 220. Uma saída do combinador 220 é conectada em comunicação de sinal com uma entrada de um filtro de desagrupamento 225 e com uma entrada de um intrapreditor 230. Uma saída do filtro de desagrupamento 225 é

conectada em comunicação de sinal com uma entrada de um armazenamento da imagem de referência 240 (para visualização i). Uma saída do armazenamento da imagem de referência 240 é conectada em comunicação de sinal com uma primeira entrada
5 de um compensador de movimento 235.

Uma saída do armazenamento da imagem de referência 245 (para outras visualizações) é conectada em comunicação de sinal com uma primeira entrada de um compensador de disparidade / iluminação 250.

10 Uma entrada de um codificador de entropia 205 fica disponível como uma entrada para o decodificador 200 para receber o fluxo contínuo de bits residual. Além do mais, uma entrada de um módulo de modo 260 também fica disponível como uma entrada para o decodificador 200 para receber a sintaxe
15 de controle para controlar qual entrada é selecionada pela chave 255. Adicionalmente, uma segunda entrada do compensador de movimento 235 fica disponível como uma entrada do decodificador 200 para receber os vetores de movimento. Também, uma segunda entrada do compensador de disparidade / i-
20 luminação 250 fica disponível como uma entrada para o decodificador 200 para receber vetores de disparidade e sintaxe de compensação de iluminação.

Uma saída de uma chave 255 é conectada em comunicação de sinal com uma segunda entrada não invertida do combinador 220. Uma primeira entrada da chave 255 é conectada
25 em comunicação de sinal com uma saída do compensador de disparidade / iluminação 250. Uma segunda entrada da chave 255 é conectada em comunicação de sinal com uma saída do compen-

sador de movimento 235. Uma terceira entrada da chave 255 é conectada em comunicação de sinal com uma saída do intrapreditor 230. Uma saída do módulo de modo 260 é conectada em comunicação de sinal com a chave 255 para controlar qual entrada é selecionada pela chave 255. Uma saída do filtro de desagrupamento 225 fica disponível como uma saída do decodificador.

A codificação de vídeo multivisualização (MVC) é a estrutura de compressão para a codificação das seqüências multivisualização. Uma seqüência de Codificação de Vídeo Multivisualização (MVC) é um conjunto de duas ou mais seqüências de vídeo que captura a mesma cena a partir de um ponto de visualização diferente.

Já que uma fonte de vídeo multivisualização inclui múltiplas visualizações da mesma cena, existe um alto grau de correlação entre as imagens de múltiplas visualizações. Portanto, a redundância de visualização pode ser explorada além da redundância temporal, e é alcançada pela realização do prognóstico de visualização através das diferentes visualizações. Dessa maneira, modalidades dos presentes princípios aqui descritos podem envolver prognóstico tanto temporal quanto de visualização cruzada.

Com propósitos ilustrativos, modalidades dos presentes princípios são aqui descritas em relação ao padrão MPEG-4 AVC. Entretanto, percebe-se que a presente invenção não é limitada ao padrão MPEG-4 AVC e, dados os preceitos dos presentes princípios aqui fornecidos, versados na técnica percebem que este e outros padrões de codificação de ví-

deo podem fazer a Codificação de Vídeo Multivisualização à qual os presentes princípios podem ser aplicados, mantendo o escopo dos presentes princípios. Modalidades dos presentes princípios aqui descritos que dizem respeito ao padrão MPEG-4 AVC podem envolver, por exemplo, mudanças no filtro de desagrupamento e/ou na codificação de entropia das sintaxes.

Em uma modalidade, no nível da fatia, as listas do prognóstico de visualização cruzada são introduzidas para habilitar prognóstico de disparidade, e uma sintaxe tipo codificação de visualização cruzada é adicionada para indicar o tipo de codificação do prognóstico de disparidade. No nível do macrobloco (MB), uma sintaxe de indicador é introduzida para indicar se a compensação de movimento ou a compensação de disparidade é usada para cada bloco de sinal. Além do mais, outras mudanças que podem ser utilizadas em modalidades direcionadas para o padrão MPEG-4 AVC incluem, por exemplo, um filtro de desagrupamento, contextos de Codificação Aritmética Binária Adaptativa ao Contexto (CABAC) para as novas sintaxes, e sintaxes adicionais no nível do conjunto de parâmetro e no nível do cabeçalho da fatia.

Agora, uma descrição será dada considerando listas do tipo codificação de visualização cruzada e listas de referência de visualização cruzada de acordo com uma modalidade dos presentes princípios.

O padrão MPEG-4 AVC realiza o prognóstico interquadros pela formação de duas listas de prognóstico, List0 e List1. Portanto, um bloco de imagem no quadro atual pode ser compensado tanto pelo uso de somente uma imagem de referên-

cia na List0 quanto pelo uso de duas imagens de referência, uma de cada lista. No cabeçalho da fatia, uma sintaxe slice_type é assinalada para indicar o tipo de codificação temporal para cada fatia. Quando slice_type = P_SLICE, somente List0 será usada na compensação de movimento. Quando slice_type = B_SLICE, possivelmente, tanto a List0 quanto a List1 podem ser usadas na compensação de movimento.

Para habilitar o prognóstico de visualização cruzada entre as diferentes visualizações, uma modalidade dos presentes princípios envolve o uso de duas novas listas de prognóstico: ViewList0 e ViewList1. Imagens na ViewList0 / ViewList1 são imagens de referência das visualizações da câmera diferentes da visualização atual. Uma nova sintaxe view_slice_type no cabeçalho da fatia é usada para indicar o tipo de codificação para o prognóstico de visualização cruzada. Por exemplo, se uma fatia específica tiver slice_type = B_SLICE e view_slice_type = P_SLICE, então, um macrobloco (MB) naquela fatia pode ser tanto temporalmente codificado como um tipo de codificação B_SLICE quanto codificado em visualização cruzada como um tipo de codificação P_SLICE.

Uma maneira alternativa de habilitar os prognósticos de visualização cruzada na estrutura do padrão MPEG-4 AVC envolve inserir imagens de referência de outras visualizações nas listas List0 / List1 sem introduzir novas listas de prognóstico de visualização e tipo de codificação de visualização cruzada. Entretanto, as vantagens da primeira abordagem são como segue. Uma vantagem da primeira abordagem é que, já que as imagens de referência em ViewList0 / Vie-

wList1 somente incluem referências de visualização cruzada, sinalizar o ref_Idx gastará menos bits do que ter tanto as referências de mesma visualização quanto as referências de visualização cruzada na mesma lista. Uma outra vantagem da primeira abordagem é que ter duas novas listas ViewList0 / ViewList1 fornece uma maneira separada de tratar os prognósticos temporal e de visualização cruzada. Isto diz respeito ao caso em que as List0 / List1 incluem tanto referências temporais quanto referências de visualização cruzada, para que o padrão MPEG-4 AVC que reordena o processo para a reconstrução das listas da imagem de referência tenha que ser modificado, e, necessariamente, será mais complexo.

Em uma modalidade, as listas de referência de visualização cruzada para cada fatia podem ser formadas de acordo com as seguintes regras. Em relação à primeira regra, no cabeçalho da fatia, o número de imagens de referência de visualização cruzada e de seus view_id's é assinalado tanto para ViewList0 quanto para ViewList1. Os view_id's são distintos em cada uma das duas listas de prognóstico de visualização cruzada. Em relação à segunda regra, as imagens de referência na lista de prognóstico de visualização cruzada são ordenadas na mesma seqüência que elas aparecem no cabeçalho da fatia. Para cada visualização referida, a imagem de referência com o número de Contagem de Ordem de Imagem (POC) mais próximo (em relação ao POC da fatia atual) é usada na lista de prognóstico de visualização cruzada da fatia atual.

Sintaxes de reordenamento de referência adicionais podem ser incluídas para permitir tratamento mais flexível

das imagens de referência de visualização cruzada.

Voltando para a figura 3, um método exemplar para a construção da lista de referência para o conteúdo de vídeo multivisualização é indicado, no geral, pelo número de referência 300. O método 300 inclui um bloco de início 305 que
5 passa o controle para um bloco de decisão 310. O bloco de decisão 310 determina se um tipo de fatia atual é fatia P ou fatia B ou não. Então, se for o caso, o controle é passado para um bloco de função 315. Caso contrário, o controle é
10 passado para um bloco de decisão 330.

O bloco de função 315 constrói a List0 usando referências temporais, e passa o controle para um bloco de decisão 320. O bloco de decisão 320 determina se o tipo de fatia atual é fatia B ou não. Então, se for o caso, o controle
15 é passado para o bloco de função 325. Caso contrário, o controle é passado para o bloco de decisão 330.

O bloco de função 325 constrói a List1 usando referências temporais e passa o controle para o bloco de decisão 330.

20 O bloco de decisão 330 determina se o tipo de fatia de visualização atual é fatia P ou fatia B ou não. Então, se for o caso, o controle é passado para um bloco de função 335. Caso contrário, o controle é passado para um bloco de limite de laço 350.

25 O bloco de função 335 constrói a ViewList0 usando referências de visualização cruzada e passa o controle para um bloco de decisão 340. O bloco de decisão 340 determina se o tipo de fatia de visualização atual é fatia B ou não. En-

tão, se for o caso, o controle é passado para um bloco de função 345. Caso contrário, o controle é passado para o bloco de limite de laço 350.

5 O bloco de função 345 constrói ViewList0 usando referências de visualização cruzada e passa o controle para o bloco de limite de laço 350.

10 O bloco de limite de laço 350 começa um laço em cada macrobloco incluindo o ajuste de uma faixa para o laço usando uma variável $mb = 0$ até $MacroBlocksinPic-1$ e passa o controle para um bloco de função 355. O bloco de função 355 codifica um macrobloco atual usando List0 / List1 e passa o controle para um bloco de decisão 360. O bloco de decisão 360 determina se o tipo de fatia de visualização atual é igual à fatia P ou fatia B ou não. Se for o caso, o controle é passado para um bloco de função 365. Caso contrário, o controle é passado para um bloco de função 370.

O bloco de função 365 codifica o macrobloco atual usando ViewList0 / ViewList1 e passa o controle para o bloco de função 370.

20 O bloco de função 370 seleciona o melhor modo, ajusta o `mvc_prediction_flag` e passa o controle para um bloco de função 375. O bloco de função 375 realiza processamento de armazenamento temporário do vetor de movimento / disparidade e passa o controle para um bloco de limite de laço 380.

25 O bloco de limite de laço termina o laço e passa o controle para um bloco de função 385. O bloco de função 385 salva a imagem codificada em armazenamento temporário de imagens decodificadas (dqb) e passa o controle para um bloco de fim

390.

Já que o prognóstico de visualização cruzada de cada fatia é completamente configurável usando listas do tipo de codificação de visualização cruzada e do prognóstico de visualização, o codec de Codificação de Vídeo Multivisualização (MVC) pode suportar ordem de codificação de visualização e escalabilidade de visualização arbitrárias.

Em uma modalidade, no nível MB, uma nova sintaxe chamada `mvc_pred_flag` indica se prognóstico temporal ou prognóstico de visualização cruzada é usado para codificar cada bloco de sinal. No caso de `mvc_pred_flag = 0`, `List0 / List1` será utilizada para a compensação de movimento dependendo do `slice_type`. Quando `mvc_pred_flag = 1`, então, `ViewList0 / ViewList1` serão utilizadas dependendo do `view_slice_type`.

Voltando para a figura 4, um método exemplar para realizar uma decisão de modo temporal / visualização cruzada é indicado, no geral, pelo número de referência 400. O método 400 inclui um bloco de início 405 que passa o controle para um bloco de decisão 410. O bloco de decisão 410 determina se o tipo de fatia atual é fatia P ou fatia B ou não. Então, se for o caso, o controle é passado para um bloco de função 415. Caso contrário, o controle é passado para um bloco de decisão 430.

O bloco de função 415 constrói `List0` usando referências temporais e passa o controle para um bloco de decisão 420. O bloco de decisão 420 determina se o tipo de fatia atual é fatia B ou não. Se for o caso, o controle é passado

para um bloco de função 425. Caso contrário, o controle é passado para o bloco de decisão 430.

O bloco de função 425 constrói a List1 usando referências temporais e passa o controle para o bloco de decisão 430.

O bloco de decisão 430 determina se o tipo de fatia de visualização atual é fatia P ou fatia B ou não. Então, se for o caso, o controle é passado para um bloco de função 435. Caso contrário, o controle é passado para um bloco de limite de laço 450.

O bloco de função 435 constrói ViewList0 usando referências de visualização cruzada e passa o controle para um bloco de decisão 440. O bloco de decisão 440 determina se o tipo de fatia de visualização atual é fatia B ou não. Então, se for o caso, o controle é passado para um bloco de função 445. Caso contrário, o controle é passado para o bloco de limite de laço 450.

O bloco de função 445 constrói a ViewList0 usando referências de visualização cruzada e passa o controle para o bloco de limite de laço 450.

O bloco de limite de laço 450 começa um laço em cada macrobloco incluindo o ajuste de uma faixa do laço usando uma variável $mb = 0$ até $MacroBlocksinPic-1$ e passa o controle para um bloco de decisão 455. O bloco de decisão 455 determina se $mvc_prediction_flag$ é igual a 1 ou não. Então, se for o caso, o controle é passado para um bloco de função 460. Caso contrário, o controle é passado para um bloco de função 465.

O bloco de função 460 decodifica um macrobloco usando ViewList0 / ViewList1 e passa o controle para um bloco de função 470.

5 O bloco de função 465 decodifica o macrobloco usando List0 / List1 e passa o controle para um bloco de função 470.

O bloco de função 470 realiza o processamento do armazenamento temporário do vetor de movimento / disparidade e passa o controle para um bloco de limite de laço 475. O
10 bloco de limite de laço 475 termina o laço e passa o controle para um bloco de função 480. O bloco de função 480 salva a imagem decodificada em armazenamento temporário de imagens decodificadas (dqb) e passa o controle para um bloco de fim
485.

15 Três novos contextos CABAC são adicionados para codificar a sintaxe mvc_pred_dir. A modelagem de contexto é a mesma da sintaxe transform_size_8x8_flag.

Na extensão multivisualização do padrão MPEG-4 AVC, o armazenamento temporário da imagem decodificada (dqb)
20 precisa ser capaz de tratar as imagens decodificadas a partir de múltiplas visualizações. Considerando que há N visualizações de entrada, uma modalidade dos presentes princípios pode envolver N dpb's separados. Cada dpb armazena imagens decodificadas de uma visualização específica.

25 Uma maneira alternativa de gerenciar dpb é colocar todas as imagens de visualização em um único dpb. Entretanto, a primeira abordagem tem as seguintes vantagens. Uma vantagem da primeira abordagem é que cada visualização tem

seu próprio dpb, com o mesmo processo de marcação de referência decodificada que o padrão MPEG-4 AVC. Esta abordagem mais simples reduz as complicações de gerenciar diferentes imagens de visualização no mesmo dpb. Uma outra vantagem da primeira abordagem diz respeito à indesejabilidade de reduzir o número de quadros de referência temporal disponíveis, já que a correlação temporal é, no geral, mais forte do que a correlação de visualização cruzada. Com cada visualização gerenciando suas próprias imagens de referência no seu dpb, o prognóstico temporal terá a mesma capacidade de prognóstico de quadro de referência múltipla que na difusão simultânea.

Um traço característico da MVC, comparando com a codificação de vídeo convencional, é a coexistência tanto de movimento quanto de disparidade. Os blocos que são temporalmente prognosticados precisarão assinalar vetores de movimento (MV) em relação a vetores de disparidade (DV) para prognóstico de visualização cruzada.

Dois métodos exemplares são aqui descritos para lidar tanto com os vetores de movimento quanto com os vetores de disparidade para a mesma fatia. Entretanto, percebe-se que dados os preceitos da presente invenção aqui fornecidos, versados na técnica contemplam estes e outros métodos para a mesma, mantendo o escopo da presente invenção.

No primeiro método, para cada bloco, sinalizar e armazenar um vetor de movimento ou um vetor de disparidade, mas não ambos. Se um vetor de movimento ou um vetor de disparidade será assinalado e armazenado depende da sintaxe

`mvc_pred_flag`. Isto exigirá menos armazenamento de memória, mas o campo do vetor combinado não será consistente.

No segundo método, para cada bloco, armazenar tanto um vetor de movimento quanto um vetor de disparidade. Isto pode ser alcançado tanto pela sinalização de ambos os vetores quanto pela sinalização de somente um deles e pelo enchimento do outro usando a interpolação do campo do vetor. Esta abordagem tomará mais armazenamento de memória, mas a consistência de ambos os campos de movimento e de disparidade pode ser mais bem preservada.

Uma modalidade exemplar do primeiro método é mostrada e descrita em relação à figura 5. Uma modalidade exemplar do segundo método é mostrada e descrita em relação à figura 6.

Voltando para a figura 5, um método exemplar para processar vetores de movimento e de disparidade para a mesma fatia correspondente ao conteúdo de vídeo multivisualização é indicado, no geral, pelo número de referência 500. O método 500 inclui um bloco de início 505 que passa o controle para um bloco de decisão 510. O bloco de decisão 510 determina se `mvc_pred_flag` é igual a 0 ou não. Então, se for o caso, o controle é passado para um bloco de função 515. Caso contrário, o controle é passado para um bloco de função 520. O bloco de função 515 forma o preditor do vetor de disparidade, processa o vetor de disparidade DV, armazena o vetor de disparidade DV em `VectorBuffer` e passa o controle para um bloco de fim 525.

O bloco de função 520 forma o preditor do vetor de

movimento, processa o vetor de movimento MV, armazena o vetor de movimento MV em VectorBuffer e passa o controle para o bloco de fim 525.

Voltando para a figura 6, um outro método para processar vetores de movimento e de disparidade para conteúdo de vídeo multivisualização é indicado, no geral, pelo número de referência 600. O método 600 inclui um bloco de início 605 que passa o controle para um bloco de função 610. O bloco de função 610 forma o preditor do vetor de disparidade, processa o vetor de disparidade DV, armazena o vetor de disparidade DV em VectorBuffer1 e passa o controle para um bloco de função 615. O bloco de função 615 forma o preditor do vetor de movimento, processa o vetor de movimento MV, armazena o vetor de movimento MV em VectorBuffer2 e passa o controle para um bloco de fim 620.

A implicação de ambos os vetores de movimento e de disparidade na codificação da mesma fatia surge nos seguintes aspectos: (1) codificação preditiva dos vetores de movimento / disparidade; e (2) modos Direto e Ignorar.

No padrão MPEG-4 AVC, os componentes do vetor de movimento são diferencialmente codificados usando prognóstico de mediana ou direcional a partir dos blocos vizinhos. Na Codificação de Vídeo Multivisualização, os blocos vizinhos podem ter uma(s) direção(s) de prognóstico diferente(s) em relação ao bloco atual. A fim de economizar bits na codificação dos vetores de movimento / disparidade, é preferível usar a informação mais correlacionada para formar um preditor. Dependendo se há tanto os vetores de movimento quanto

os vetores de disparidade disponíveis para os blocos vizinhos, para o primeiro método, use somente aqueles blocos vizinhos que têm a mesma direção de prognóstico. Para o segundo método, use somente os vetores de movimento dos blocos vizinhos na formação do preditor do vetor de movimento, e use somente os vetores de disparidade dos blocos vizinhos na formação do preditor de disparidade.

Aparte dos blocos espaciais vizinhos, blocos temporalmente co-localizados também podem ser usados para melhorar o prognóstico de disparidade em virtude de os campos de disparidade ser usualmente estacionários na dimensão temporal.

Os modos Direto e Ignorar no padrão MPEG-4 AVC são ferramentas de codificação efetivas que exploram melhor a correlação espaço-temporal que existe entre os macroblocos adjacentes, em virtude de eles poderem representar movimento sem ter que transmitir os vetores de movimento. Na Codificação de Vídeo Multivisualização, aqueles modos devem ser adaptados a fim de considerar a correlação de visualização cruzada adicional.

Para modos P_Skip, o sinal reconstruído é obtido similar ao sinal de prognóstico de um macrobloco tipo P_16x16 que referencia a imagem que está localizada no índice 0 de List0. O vetor de movimento usado para reconstruir o macrobloco P_Skip é similar ao preditor do vetor de movimento para o bloco 16x16. Na MVC, a adaptação supradescrita do preditor do vetor de movimento / disparidade ajudará a tornar o modo P_Skip mais usado.

Para a codificação de B_SLICE, os modos B_Skip / B_Direct_16x16 / B_Direct_8x8 devem ser adaptados para considerar a mistura de movimento e de disparidade. Há dois diferentes modos Direto suportados no padrão MPEG-4 AVC, a saber, Direto temporal e Direto espacial.

Para o modo Direto temporal, os vetores de movimento são derivados da posição co-localizada na primeira referência da List1. Quando a primeira referência de List1 for disparidade prognosticada, o sistema pode tanto procurar vetores de movimento na posição co-localizada em outras referências da List1 ($ref_idx > 0$) quanto usar o preditor do vetor de movimento espacial.

Para o modo Direto espacial, os vetores de movimento são derivados de uma maneira similar à empregada pelo P_SKIP, mas com ambas List0 / List1 consideradas. A mesma adaptação feita em P_SKIP também pode ser estendida na List1.

As Tabelas 1-4 ilustram várias sintaxes para Codificação de Vídeo Multivisualização incluindo aquelas de acordo com várias modalidades dos presentes princípios. A Tabela 1 ilustra a sintaxe RBSP ajustada do parâmetro seqüência para Codificação de Vídeo Multivisualização. A Tabela 2 ilustra a sintaxe RBSP ajustada do parâmetro imagem para Codificação de Vídeo Multivisualização. A Tabela 3 ilustra a sintaxe do cabeçalho de fatia para Codificação de Vídeo Multivisualização. A Tabela 4 ilustra a sintaxe de camada de macrobloco para Codificação de Vídeo Multivisualização.

TABELA 1

| seq_parameter_set_rbsp(){ | C | Descriptor |
|--|---|-------------------------------|
| log2 max view num minus1 | 0 | ue(v) |
| num_views_sps | 0 | u(log2_max_view_num_minus1+1) |
| view_id_sps | 0 | u(log2_max_view_num_minus1+1) |
| profile_idc | 0 | u(8) |
| constraint_set0_flag | 0 | u(1) |
| constraint_set1_flag | 0 | u(1) |
| constraint_set2_flag | 0 | u(1) |
| constraint_set3_flag | 0 | u(1) |
| reserved_zero_4bits/* equal to 0 */ | 0 | u(4) |
| ... | | |

TABELA 2

| pic_parameter_set_rbsp(){ | C | Descriptor |
|-------------------------------|---|-------------------------------|
| view id pps | 0 | u(log2 max view num minus1+1) |
| pic_parameter_set_id | 1 | ue(v) |
| seq parameter set id | 1 | ue(v) |
| en- tropy coding mode flag | 1 | u(1) |
| plc order present flag | 1 | u(1) |
| num slice groups minus1 | 1 | ue(v) |

| | | |
|--|---|------|
| Reserved_zero_4bits/* equal to 0 */ | 0 | u(4) |
| ... | | |

TABELA 3

| | | |
|--|---|-------------------------------|
| slice header(){ | C | Descriptor |
| first mb in slice | 2 | ue(v) |
| view id | 2 | u(log2 max view num minus1+1) |
| view slice type | 2 | ue(v) |
| se(view_slice_type == VL SLICE){ | | |
| num_ref_idx_ll_active_minus1 | 2 | ue(v) |
| para(i=0; i<=num_ref_idx_ll_active_minus1; i++){ | | |
| left_ref_view_id[i] | 2 | ue(v) |
| } | 0 | u(4) |
| } | | |
| se(view_slice_type == VR SLICE){ | | |
| num_ref_idx_lr_active_minus1 | 2 | ue(v) |
| para(i=0; i<= num_ref_idx_lr_active_minus1; i++){ | | |
| right_ref_view_id[i] | 2 | ue(v) |
| } | | |
| } | | |

| | | |
|---|---|-------|
| se(view_slice_type == VB SLICE) { | | |
| num_ref_idx_ll_active_minus1 | 2 | ue(v) |
| para(i=0; i<= num_ref_idx_ll_active_minus1; i++){ | | |
| left_ref_view_id[i] | 2 | ue(v) |
| } | | |
| num_ref_idx_lr_active_minus1 | 2 | ue(v) |
| para(i=0; i<= num_ref_idx_lr_active_minus1; i++){ | | |
| right_ref_view_id[i] | 2 | ue(v) |
| } | | |
| } | | |
| ... | | |
| slice_type | 2 | ue(v) |
| pic_parameter_set_id | 2 | ue(v) |
| frame_num | 2 | U(v) |
| ... | | |

TABELA 4

| macroblock_layer() { | C | Descriptor |
|---------------------------|---|-------------|
| mvc_pred_flag | 2 | u(i)lac(v) |
| mb_type | 2 | ue(v)lac(v) |
| se(mb_type == I_PCM) { | | |
| enquanto(!byte_aligned()) | | |

| | | |
|------------------------|---|------|
| pcm_alignment_zero_bit | 2 | f(1) |
| para(i=0; i<256; i++) | | |
| pcm_sample_luma[i] | 2 | U(v) |
| ... | | |

Agora, será dada uma descrição de algumas das muitas vantagens / recursos presentes da presente invenção, algumas das quais fora mencionadas anteriormente. Por exemplo, uma vantagem / recurso é um codificador de vídeo que inclui

5 um codificador para codificar um bloco em uma imagem pela escolha entre prognóstico temporal e prognóstico de visualização cruzada para habilitar um prognóstico para o bloco. A imagem é uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos

10 de visualização em relação à mesma cena ou a cenas similares. A imagem representa um dos diferentes pontos de visualização. Uma sintaxe de alto nível é usada para indicar o uso do prognóstico de visualização cruzada para o bloco.

Uma outra vantagem / recurso é o codificador de

15 vídeo supradescrito, em que o codificador codifica a imagem para fornecer um fluxo contínuo de bits resultante, em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização

20 / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma extensão desta.

Uma ainda outra vantagem / recurso é o codificador

de vídeo supradescrito, em que a sintaxe de alto nível inclui uma sintaxe em nível de fatia.

Além do mais, uma outra vantagem / recurso é o codificador de vídeo supradescrito, em que uma sintaxe em nível de fatia é usada para indicar se compensação de movimento ou compensação de disparidade é aplicada no bloco.

Adicionalmente, uma outra vantagem / recurso é um codificador de vídeo que inclui um codificador para codificar um bloco em uma imagem usando um vetor de disparidade. A imagem corresponde ao conteúdo de vídeo multivisualização e é codificada para fornecer um fluxo contínuo de bits resultante, em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma extensão desta.

Também, uma outra vantagem / recurso é o codificador de vídeo supradescrito, em que o codificador mistura um vetor de movimento e o vetor de disparidade, ambos correspondendo ao bloco em um único campo do vetor.

Adicionalmente, uma outra vantagem / recurso é o codificador de vídeo supradescrito, em que o codificador codifica o bloco usando pelo menos um de um preditor do vetor de movimento e de um preditor do vetor de disparidade derivado somente dos blocos vizinhos que têm pelo menos um de um mesmo indicador do prognóstico temporal e de um mesmo indicador do prognóstico de visualização cruzada como o bloco.

Além do mais, uma outra vantagem / recurso é o codificador de vídeo supradescrito, em que o codificador codifica o bloco usando campos do vetor separados para um vetor de movimento e um vetor de disparidade correspondentes ao
5 bloco.

Também, uma outra vantagem / recurso é o codificador de vídeo supradescrito, em que o codificador codifica o bloco usando pelo menos um de um preditor do vetor de movimento, que diz respeito somente a um campo do vetor de movimento correspondente, e de um preditor do vetor de disparidade, que diz respeito somente a um campo do vetor de disparidade correspondente, o campo do vetor de movimento e o campo do vetor de disparidade incluídos nos campos do vetor separados.
10

Adicionalmente, uma outra vantagem / recurso é um codificador de vídeo que inclui um codificador para codificar um bloco em uma imagem correspondente ao conteúdo de vídeo multivisualização com base em um preditor do vetor de disparidade que usa pelo menos um de um vetor de disparidade temporalmente co-localizado e de um vetor de disparidade vizinho de um bloco temporalmente co-localizado.
15
20

Estes e outros recursos e vantagens da presente invenção podem ser prontamente apurados pelos versados na técnica com base nos preceitos aqui descritos. Entende-se
25 que os preceitos da presente invenção podem ser implementados em várias formas de hardware, software, software embarcado, processadores de uso pessoal ou combinações destes.

Mais preferivelmente, os preceitos da presente in-

venção são implementados como uma combinação de hardware e software. Além do mais, o software pode ser implementado como um programa de aplicação tangivelmente incorporado em uma unidade de armazenamento de programa. O programa de aplicação pode ser carregado em uma máquina que compreende qualquer arquitetura e executado por ela. Preferivelmente, a máquina é implementada em uma plataforma de computador com hardware tal como uma ou mais unidades centrais de processamento ("CPU"), uma memória de acesso aleatório ("RAM") e interfaces de entrada / saída ("I/O"). A plataforma de computador também pode incluir um sistema operacional e código de microinstrução. Os vários processos e funções aqui descritos podem ser tanto parte do código de microinstrução quanto parte do programa de aplicação, ou qualquer combinação destes, que podem ser executados por uma CPU. Além do mais, várias outras unidades periféricas podem ser conectadas na plataforma do computador, tais como uma unidade de armazenamento de dados adicional e uma unidade de impressão.

Entende-se adicionalmente que, em virtude de alguns dos componentes e métodos do sistema constituinte representados nos desenhos anexos ser preferivelmente implementados em software, as conexões reais entre os componentes do sistema ou os blocos de função do processo podem diferir dependendo da maneira na qual a presente invenção é programada. Dados os preceitos aqui descritos, versados na técnica podem contemplar estas e similares implementações e configurações da presente invenção.

Embora as modalidades ilustrativas tenham sido a-

qui descritas em relação aos desenhos anexos, entende-se que a presente invenção não é limitada a estas precisas modalidades, e que várias mudanças e modificações podem ser efetuadas pelos versados na técnica sem fugir do escopo ou do espírito da presente invenção. Pretende-se que todas tais mudanças e modificações sejam incluídas no escopo da presente invenção apresentado nas reivindicações anexas.

REIVINDICAÇÕES

1. Codificador de vídeo, **CARACTERIZADO** pelo fato de que compreende:

um codificador (100) para codificar um bloco em
5 uma imagem pela escolha entre prognóstico temporal e prognóstico de visualização cruzada para habilitar um prognóstico para o bloco, a imagem sendo uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos de visualização em relação a uma mesma
10 cena ou a uma cena similar, a imagem representando um de diferentes pontos de visualização,

em que uma sintaxe de alto nível é usada para indicar o uso do prognóstico de visualização cruzada para o bloco.

15 2. Codificador de vídeo, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o dito codificador (100) codifica a imagem para fornecer um fluxo contínuo de bits resultante, em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas
20 de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma extensão desta.

25 3. Codificador de vídeo, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que a sintaxe de alto nível compreende uma sintaxe em nível de fatia.

4. Codificador de vídeo, de acordo com a reivindi-

cação 1, **CARACTERIZADO** pelo fato de que uma sintaxe em nível de bloco é usada para indicar se a compensação de movimento ou a compensação de disparidade é aplicada no bloco.

5. Codificador de vídeo, **CARACTERIZADO** pelo fato
5 de que compreende:

um codificador (100) para codificar um bloco em uma imagem usando um vetor de disparidade em que a imagem corresponde ao conteúdo de vídeo multivisualização e é codificada para fornecer um fluxo contínuo de bits resultante,
10 em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União In-
15 ternacional de Telecomunicações e uma extensão desta.

6. Codificador de vídeo, de acordo com a reivindicação 5, **CARACTERIZADO** pelo fato de que o dito codificador (100) mistura um vetor de movimento e o vetor de disparidade, ambos correspondendo ao bloco em um único campo do ve-
20 tor.

7. Codificador de vídeo, de acordo com a reivindicação 6, **CARACTERIZADO** pelo fato de que o dito codificador (100) codifica o bloco usando pelo menos um de um preditor do vetor de movimento e de um preditor do vetor de disparidade derivados somente de blocos vizinhos que têm pelo menos um de um mesmo indicador do prognóstico temporal e de um mesmo indicador do prognóstico de visualização cruzada que o bloco.

8. Codificador de vídeo, de acordo com a reivindicação 5, **CARACTERIZADO** pelo fato de que o dito codificador (100) codifica o bloco usando campos do vetor separados para um vetor de movimento e para um vetor de disparidade correspondentes ao bloco.

9. Codificador de vídeo, de acordo com a reivindicação 8, **CARACTERIZADO** pelo fato de que o dito codificador (100) codifica o bloco usando pelo menos um de um preditor do vetor de movimento, que diz respeito somente a um campo do vetor de movimento correspondente, e de um preditor do vetor de disparidade, que diz respeito somente a um campo do vetor de disparidade correspondente, o campo do vetor de movimento e o campo do vetor de disparidade incluídos nos campos do vetor separados.

10. Codificador de vídeo, **CARACTERIZADO** pelo fato de que compreende:

um codificador (100) para codificar um bloco em uma imagem correspondente ao conteúdo de vídeo multivisualização com base em um preditor do vetor de disparidade que usa pelo menos um de um vetor de disparidade temporalmente co-localizado e de um vetor de disparidade vizinho de um bloco temporalmente co-localizado.

11. Método para codificação de vídeo, **CARACTERIZADO** pelo fato de que compreende:

codificar (370) um bloco em uma imagem pela escolha entre prognóstico temporal e prognóstico de visualização cruzada para habilitar um prognóstico para o bloco, a imagem sendo uma de um conjunto de imagens correspondente ao conte-

údo de vídeo multivisualização e com diferentes pontos de visualização em relação a uma mesma cena ou a uma cena similar, a imagem representando um de diferentes pontos de visualização,

5 em que uma sintaxe de alto nível é usada para indicar o uso de prognóstico de visualização cruzada para o bloco.

12. Método, de acordo com a reivindicação 11, **CARACTERIZADO** pelo fato de que a dita etapa de codificação
10 codifica a imagem para fornecer um fluxo contínuo de bits resultante, em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional /
15 ou da recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma extensão desta (300).

13. Método, de acordo com a reivindicação 11, **CARACTERIZADO** pelo fato de que a sintaxe de alto nível compreende uma sintaxe em nível de fatia (330).
20

14. Método, de acordo com a reivindicação 11, **CARACTERIZADO** pelo fato de que uma sintaxe em nível de bloco é usada para indicar se a compensação de movimento ou a compensação de disparidade é aplicada no bloco (370).

25 15. Método para codificação de vídeo, **CARACTERIZADO** pelo fato de que compreende:

codificar um bloco em uma imagem usando um vetor de disparidade em que a imagem corresponde ao conteúdo de

vídeo multivisualização e é codificada para fornecer um fluxo contínuo de bits resultante, em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma extensão desta (300).

16. Método, de acordo com a reivindicação 15,
10 **CARACTERIZADO** pelo fato de que a dita etapa de codificação mistura um vetor de movimento e o vetor de disparidade, ambos correspondentes ao bloco em um único campo do vetor (500).

17. Método, de acordo com a reivindicação 16,
15 **CARACTERIZADO** pelo fato de que a dita etapa de codificação codifica o bloco usando pelo menos um de um preditor do vetor de movimento e de um preditor do vetor de disparidade derivados somente dos blocos vizinhos que têm pelo menos um de um mesmo indicador do prognóstico temporal e de um mesmo
20 indicador do prognóstico de visualização cruzada que o bloco (515, 520).

18. Método, de acordo com a reivindicação 15,
CARACTERIZADO pelo fato de que a dita etapa de codificação codifica o bloco usando campos do vetor separados para um
25 vetor de movimento e para um vetor de disparidade correspondentes ao bloco (600).

19. Método, de acordo com a reivindicação 18,
CARACTERIZADO pelo fato de que a dita etapa de codificação

codifica o bloco usando pelo menos um de um preditor do vetor de movimento, que diz respeito somente a um campo do vetor de movimento correspondente, e de um preditor do vetor de disparidade, que diz respeito somente a um campo do vetor de disparidade correspondente, o campo do vetor de movimento e o campo do vetor de disparidade incluídos nos campos do vetor separados (610, 615).

20. Método para codificação de vídeo, **CARACTERIZADO** pelo fato de que compreende:

10 codificar (375) um bloco em uma imagem correspondente ao conteúdo de vídeo multivisualização com base em um preditor do vetor de disparidade que usa pelo menos um de um vetor de disparidade temporalmente co-localizado e de um vetor de disparidade vizinho de um bloco temporalmente co-
15 localizado.

21. Decodificador de vídeo, **CARACTERIZADO** pelo fato de que compreende:

um decodificador (200) para decodificar um bloco em uma imagem pela escolha entre prognóstico temporal e
20 prognóstico de visualização cruzada para habilitar um prognóstico para o bloco, a imagem sendo uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos de visualização em relação a uma mesma cena ou a uma cena similar, a imagem representando um
25 dos diferentes pontos de visualização,

em que uma sintaxe de alto nível é lida para determinar o uso do prognóstico de visualização cruzada para o bloco.

22. Decodificador de vídeo, de acordo com a reivindicação 21, **CARACTERIZADO** pelo fato de que o dito decodificador (200) decodifica a imagem a partir de um fluxo contínuo de bits em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma extensão desta.

10 23. Decodificador de vídeo, de acordo com a reivindicação 21, **CARACTERIZADO** pelo fato de que a sintaxe de alto nível compreende uma sintaxe em nível de fatia.

24. Decodificador de vídeo, de acordo com a reivindicação 21, **CARACTERIZADO** pelo fato de que uma sintaxe em nível de bloco é usada para indicar se a compensação de movimento ou a compensação de disparidade é aplicada no bloco.

25. Decodificador de vídeo, **CARACTERIZADO** pelo fato de que compreende:

um decodificador (200) para decodificar um bloco em uma imagem usando um vetor de disparidade em que a imagem corresponde ao conteúdo de vídeo multivisualização e é decodificada a partir de um fluxo contínuo de bits, em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma extensão desta.

26. Decodificador de vídeo, de acordo com a reivindicação 25, **CARACTERIZADO** pelo fato de que o dito decodificador (200) usa um vetor de movimento e o vetor de disparidade, ambos correspondentes ao bloco, misturados em um único campo do vetor.

27. Decodificador de vídeo, de acordo com a reivindicação 26, **CARACTERIZADO** pelo fato de que o dito decodificador (200) decodifica o bloco usando pelo menos um de um preditor do vetor de movimento e de um preditor do vetor de disparidade derivados somente de blocos vizinhos que têm pelo menos um de um mesmo indicador do prognóstico temporal e de um mesmo indicador do prognóstico de visualização cruzada que o bloco.

28. Decodificador de vídeo, de acordo com a reivindicação 25, **CARACTERIZADO** pelo fato de que o dito decodificador (200) decodifica o bloco usando campos do vetor separados para um vetor de movimento e para um vetor de disparidade correspondentes ao bloco.

29. Decodificador de vídeo, de acordo com a reivindicação 28, **CARACTERIZADO** pelo fato de que o dito decodificador (200) decodifica o bloco usando pelo menos um de um preditor do vetor de movimento, que diz respeito somente a um campo de vetor de movimento correspondente, e a um preditor do vetor de disparidade, que diz respeito somente a um campo do vetor de disparidade correspondente, o campo do vetor de movimento e o campo do vetor de disparidade incluídos nos campos do vetor separados.

30. Decodificador de vídeo, **CARACTERIZADO** pelo fa-

to de que compreende:

um decodificador (200) para decodificar um bloco em uma imagem correspondente ao conteúdo de vídeo multivisualização com base em um preditor do vetor de disparidade que usa pelo menos um de um vetor de disparidade temporalmente co-localizado e de um vetor de disparidade vizinho de um bloco temporalmente co-localizado.

31. Método para decodificação de vídeo, **CARACTERIZADO** pelo fato de que compreende:

10 decodificar um bloco em uma imagem pela escolha entre prognóstico temporal e prognóstico de visualização cruzada para habilitar um prognóstico para o bloco, a imagem sendo uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos de visualização em relação a uma mesma cena ou a uma cena similar, a imagem representando um de diferentes pontos de visualização,

em que uma sintaxe de alto nível é lida para determinar o uso de prognóstico de visualização cruzada para o bloco (455).

25 32. Método, de acordo com a reivindicação 31, **CARACTERIZADO** pelo fato de que a dita etapa de decodificação decodifica a imagem a partir de um fluxo contínuo de bits em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União Inter-

nacional de Telecomunicações e uma extensão desta.

33. Método, de acordo com a reivindicação 31, **CARACTERIZADO** pelo fato de que a sintaxe de alto nível compreende uma sintaxe em nível de fatia (430).

5 34. Método, de acordo com a reivindicação 31, **CARACTERIZADO** pelo fato de que uma sintaxe em nível de bloco é usada para indicar se compensação de movimento ou compensação de disparidade é aplicada no bloco (455).

10 35. Método para decodificação de vídeo, **CARACTERIZADO** pelo fato de que compreende:

decodificar um bloco em uma imagem usando um vetor de disparidade em que a imagem corresponde ao conteúdo de vídeo multivisualização e é decodificada a partir de um fluxo contínuo de bits, em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma
15
20 extensão desta (400).

36. Método, de acordo com a reivindicação 35, **CARACTERIZADO** pelo fato de que a dita etapa de decodificação usa um vetor de movimento e o vetor de disparidade, ambos correspondendo ao bloco, misturados em um único campo de vetor
25 tor (500).

37. Método, de acordo com a reivindicação 36, **CARACTERIZADO** pelo fato de que a dita etapa de decodificação decodifica o bloco usando pelo menos um de um preditor do

vetor de movimento e de um preditor do vetor de disparidade derivados somente de blocos vizinhos que têm pelo menos um de um mesmo indicador do prognóstico temporal e de um mesmo indicador do prognóstico de visualização cruzada que o bloco
5 (515, 520).

38. Método, de acordo com a reivindicação 35, **CARACTERIZADO** pelo fato de que a dita etapa de decodificação decodifica o bloco usando campos de vetor separados para um vetor de movimento e um vetor de disparidade correspondente
10 ao bloco (600).

39. Método, de acordo com a reivindicação 38, **CARACTERIZADO** pelo fato de que a dita etapa de decodificação decodifica o bloco usando pelo menos um de um preditor do vetor de movimento, que diz respeito somente a um campo do vetor de movimento correspondente, e um preditor do vetor de
15 disparidade que diz respeito somente a um campo do vetor de disparidade correspondente, o campo do vetor de movimento e o campo do vetor de disparidade incluídos nos campos de vetor separados (610, 615).

20 40. Método para decodificação de vídeo, **CARACTERIZADO** pelo fato de que compreende:

decodificar (470) um bloco em uma imagem correspondente ao conteúdo de vídeo multivisualização com base em um preditor do vetor de disparidade que usa pelo menos um de
25 um vetor de disparidade temporalmente co-localizado e de um vetor de disparidade vizinho de um bloco temporalmente co-localizado.

41. Estrutura de sinal de vídeo para codificação

de vídeo, **CARACTERIZADA** pelo fato de que compreende:

um bloco em uma imagem codificado pela escolha entre prognóstico temporal e prognóstico de visualização cruzada para habilitar um prognóstico para o bloco, a imagem sendo uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos de visualização em relação a uma mesma cena ou a uma cena similar, a imagem representando um dos diferentes pontos de visualização,

em que uma sintaxe no nível da fatia é usada para indicar o uso do prognóstico de visualização cruzada para o bloco.

42. Mídia de armazenamento com dados de sinal de vídeo codificados, **CARACTERIZADA** pelo fato de que compreende:

um bloco em uma imagem codificado pela escolha entre prognóstico temporal e prognóstico de visualização temporal para habilitar um prognóstico para o bloco, a imagem sendo uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos de visualização em relação a uma mesma cena ou a uma cena similar, a imagem representando um dos diferentes pontos de visualização,

em que uma sintaxe em nível de fatia é usada para indicar o uso do prognóstico de visualização cruzada para o bloco.

43. Estrutura de sinal de vídeo para codificação de vídeo, **CARACTERIZADA** pelo fato de que compreende:

um bloco em uma imagem codificado pela escolha entre prognóstico temporal e prognóstico de visualização cruzada para habilitar um prognóstico para o bloco, a imagem sendo uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos de visualização em relação a uma mesma cena ou a uma cena similar, a imagem representando um dos diferentes pontos de visualização,

em que uma sintaxe em nível de bloco é usada para indicar se compensação de movimento ou compensação de disparidade é aplicada no bloco.

44. Mídia de armazenamento com dados de sinal de vídeo codificados, **CARACTERIZADA** pelo fato de que compreende:

um bloco em uma imagem codificado pela escolha entre prognóstico temporal e prognóstico de visualização cruzada para habilitar um prognóstico para o bloco, a imagem sendo uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos de visualização em relação a uma mesma cena ou a uma cena similar, a imagem representando um dos diferentes pontos de visualização,

em que uma sintaxe no nível do bloco é usada para indicar se compensação de movimento ou compensação de disparidade é aplicada no bloco.

45. Estrutura de sinal de vídeo para codificação de vídeo, **CARACTERIZADA** pelo fato de que compreende:

um bloco em uma imagem codificado usando um vetor

de disparidade em que a imagem corresponde ao conteúdo de vídeo multivisualização e é codificada para fornecer um fluxo contínuo de bits resultante, em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma extensão desta.

10 46. Mídia de armazenamento com dados de sinal de vídeo codificados, **CARACTERIZADA** pelo fato de que compreende:

 um bloco em uma imagem codificado usando um vetor de disparidade em que a imagem corresponde ao conteúdo de vídeo multivisualização e é codificada para fornecer um fluxo contínuo de bits resultante, em conformidade com pelo menos um do padrão de Codificação Avançada de Vídeo do Grupo de Especialistas de Imagens em Movimento-4 Parte 10 da Organização Internacional para Padronização / Comissão Eletrotécnica Internacional / ou da recomendação H.264 do Setor de Telecomunicações da União Internacional de Telecomunicações e uma extensão desta.

 47. Estrutura de sinal de vídeo para codificação de vídeo, **CARACTERIZADA** pelo fato de que compreende:

25 um bloco em uma imagem correspondente ao conteúdo de vídeo multivisualização, o bloco codificado com base em um preditor do vetor de disparidade que usa pelo menos um de um vetor de disparidade temporalmente co-localizado e de um

vetor de disparidade vizinho de um bloco temporalmente co-localizado.

48. Mídia de armazenamento com dados de sinal de vídeo codificados, **CARACTERIZADA** pelo fato de que compreende:

um bloco em uma imagem correspondente ao conteúdo de vídeo multivisualização, o bloco codificado com base em um preditor do vetor de disparidade que usa pelo menos um de um vetor de disparidade temporalmente co-localizado e de um
10 vetor de disparidade vizinho de um bloco temporalmente co-localizado.

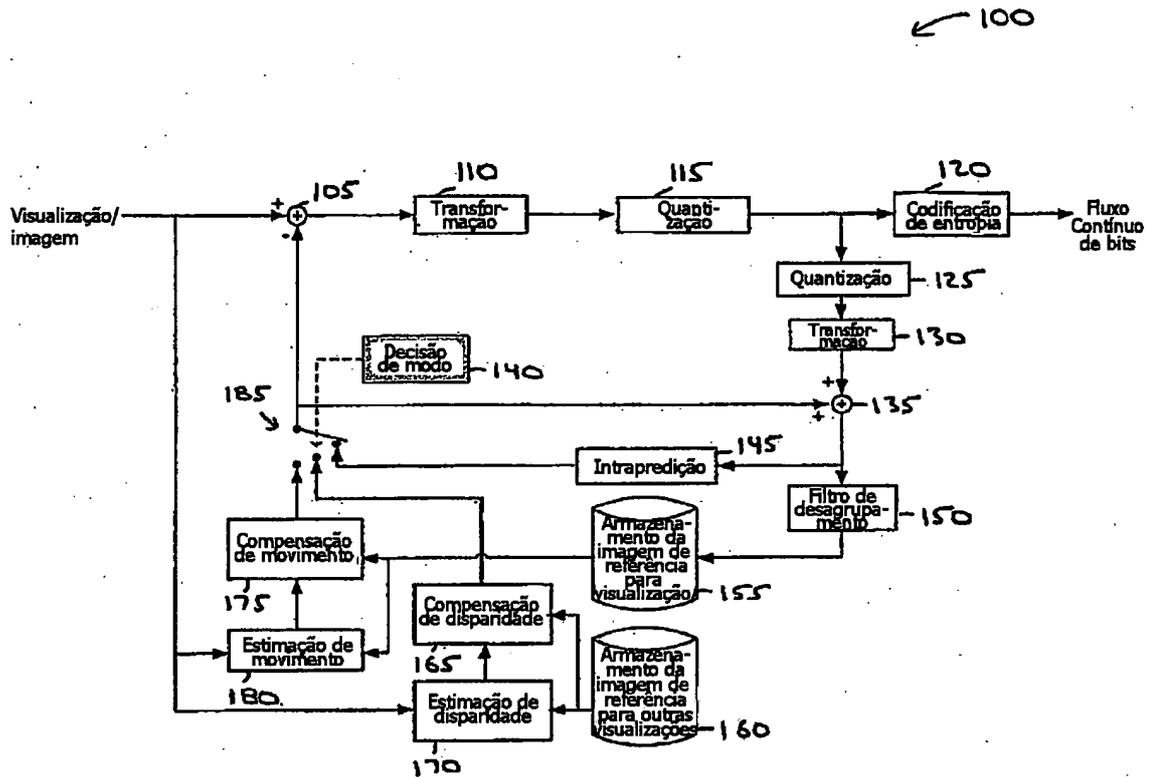


FIG. 1

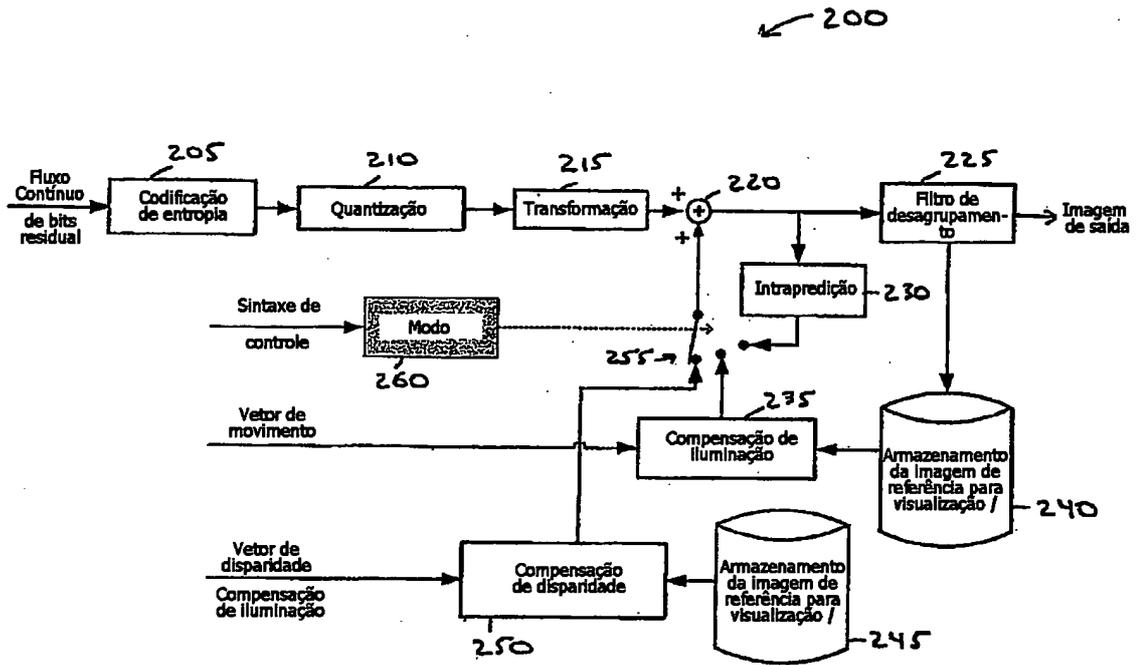


FIG. 2

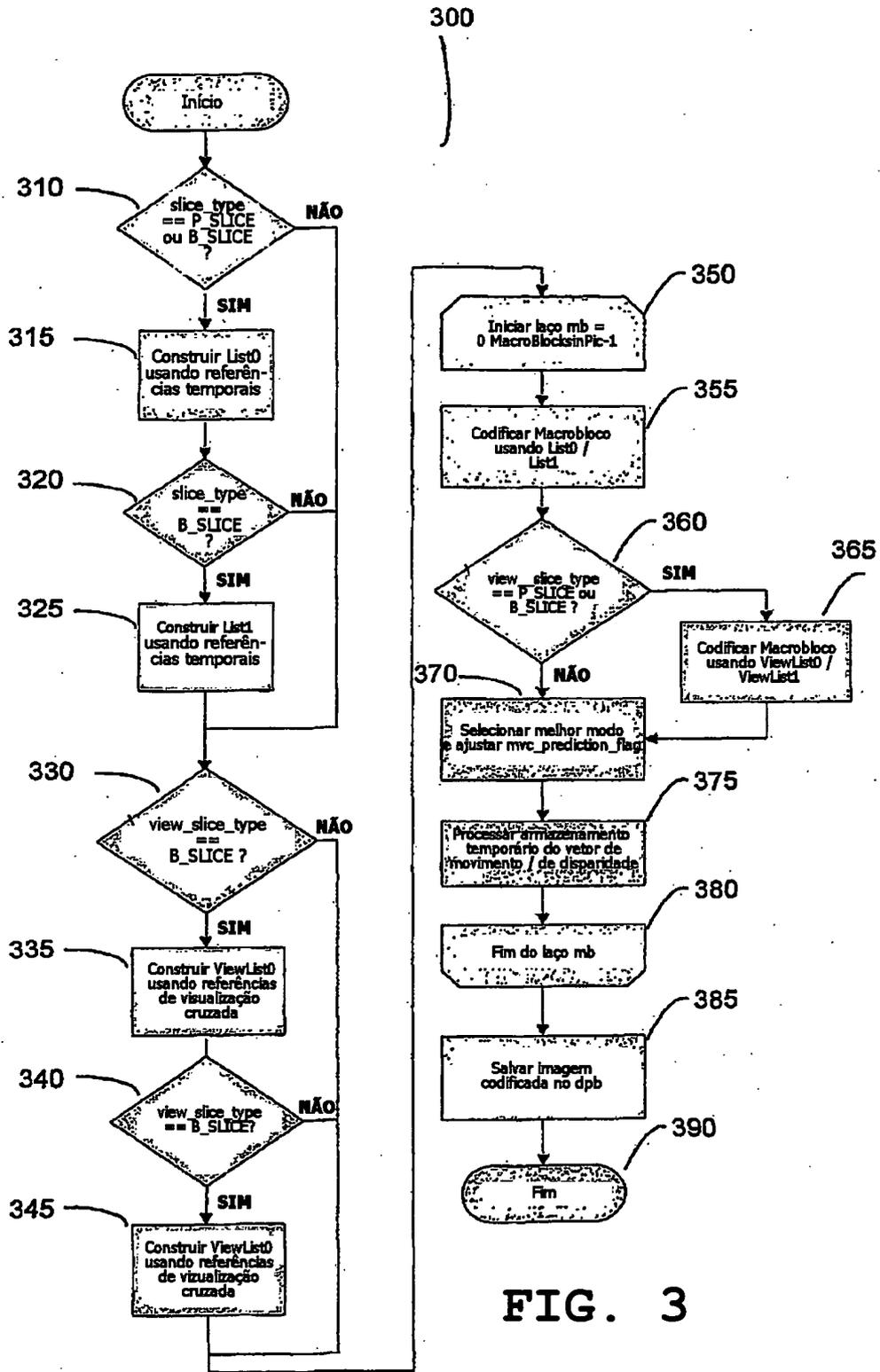


FIG. 3

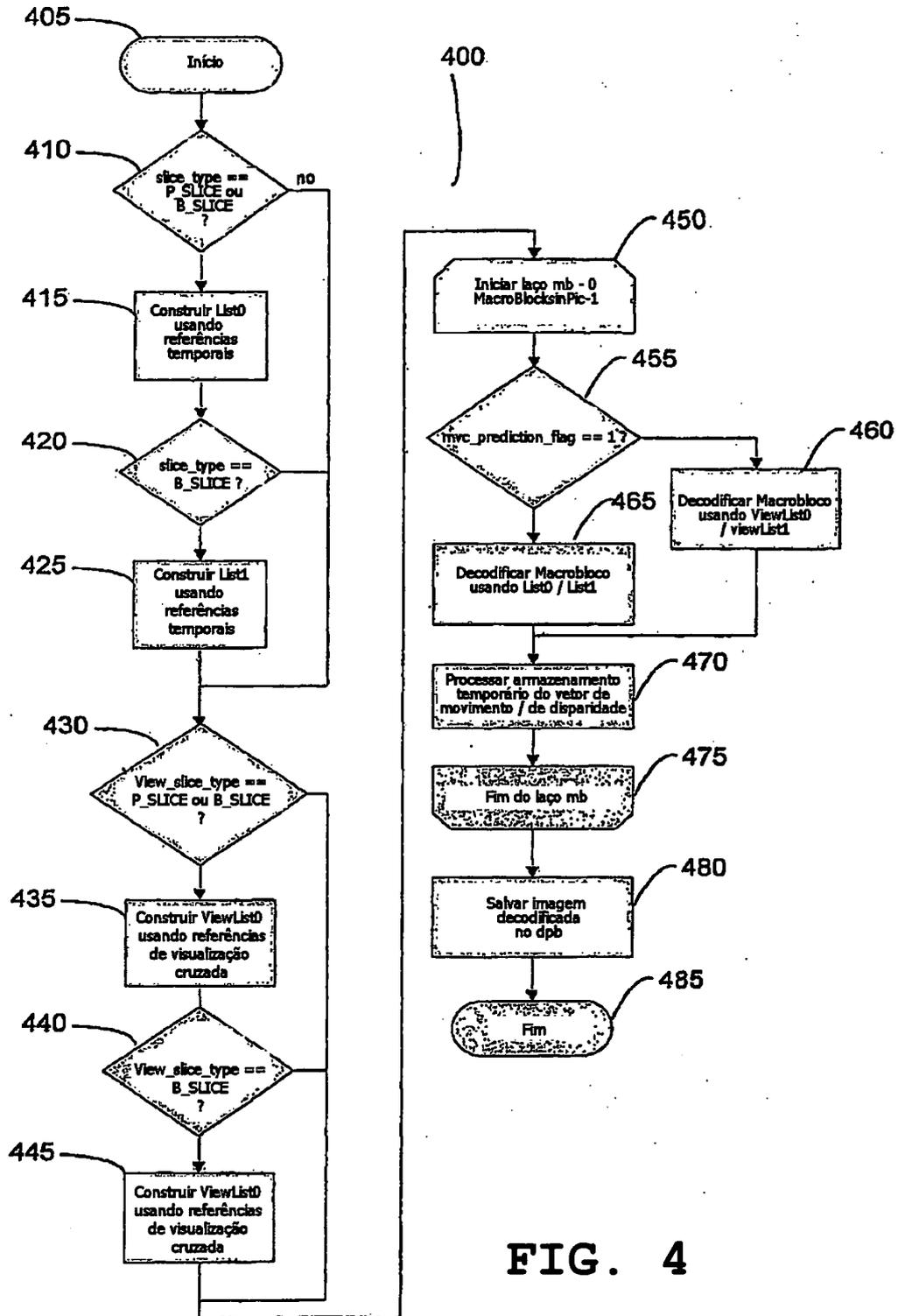


FIG. 4

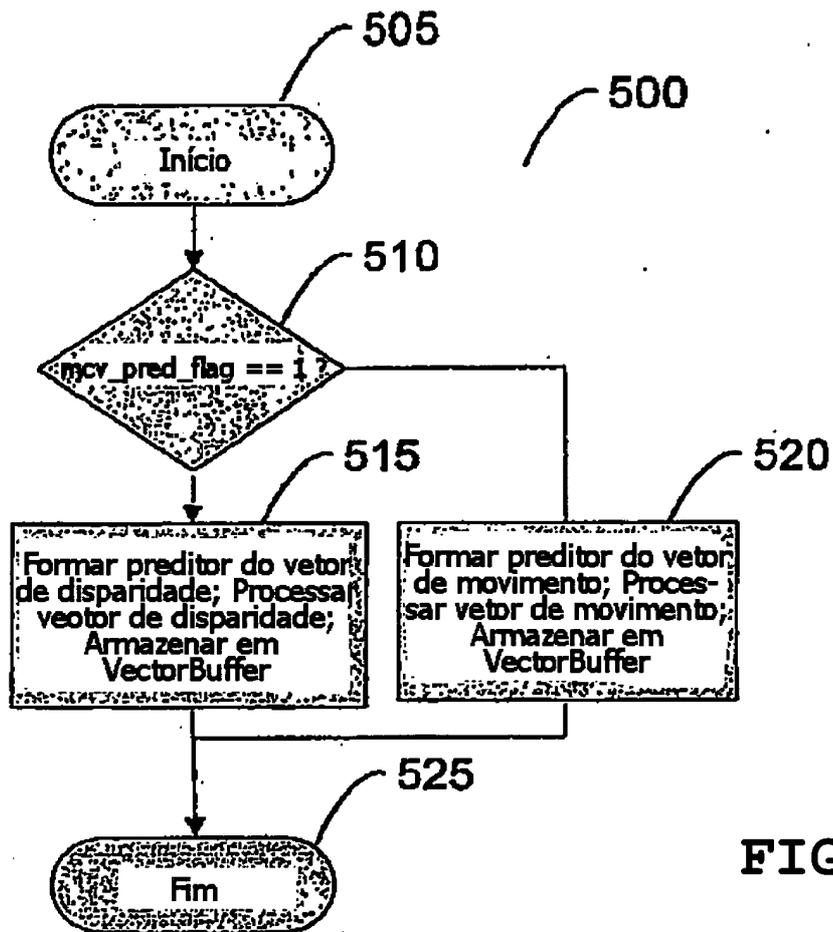


FIG. 5

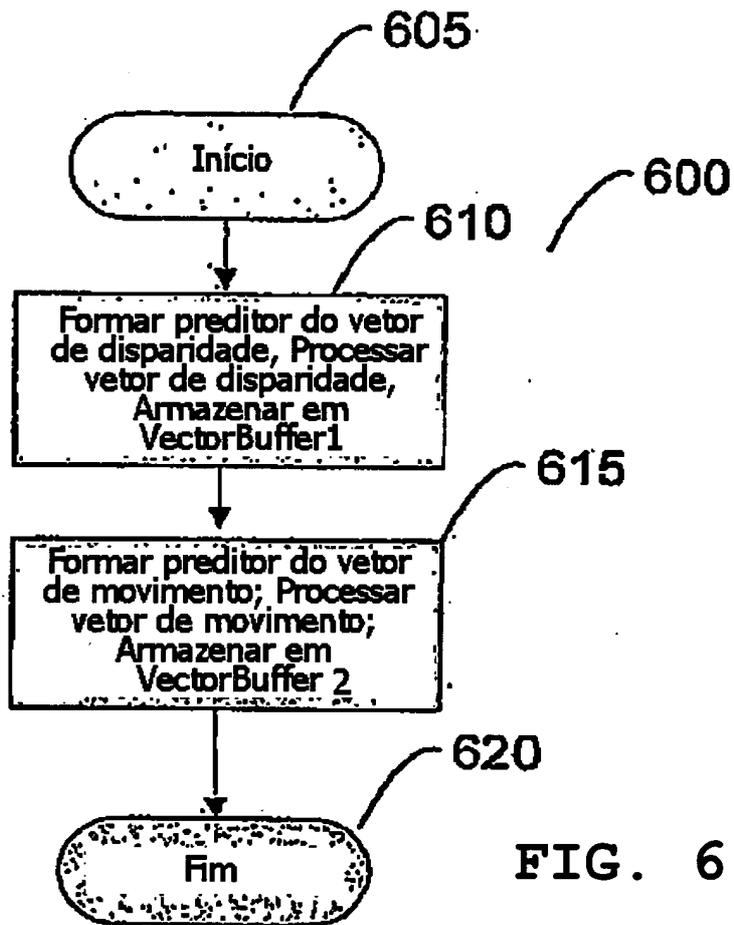


FIG. 6

RESUMO

"MÉTODOS E APARELHO PARA CODIFICAÇÃO DE VÍDEO MULTIVISUALIZAÇÃO"

São fornecidos métodos e aparelho para codificação de vídeo multivisualização. Um codificador de vídeo inclui um codificador (100) para codificar um bloco em uma imagem pela escolha entre prognóstico temporal e prognóstico de visualização cruzada para habilitar um prognóstico para o bloco. A imagem é uma de um conjunto de imagens correspondente ao conteúdo de vídeo multivisualização e com diferentes pontos de visualização em relação a uma mesma cena ou a uma cena similar. A imagem representa um dos diferentes pontos de visualização. Uma sintaxe de alto nível é usada para indicar o uso do prognóstico de visualização cruzada para o bloco.