República Federativa do Brasil

Ministério do Desenvolvimento, Indústria, Comércio e Serviços

Instituto Nacional da Propriedade Industrial

(11) BR 112015016678-4 B1

(22) Data do Depósito: 10/01/2014

(45) Data de Concessão: 18/04/2023

(54) Título: SÍNTESE DE VISUALIZAÇÃO EM VÍDEO 3D

(51) Int.Cl.: H04N 19/583; H04N 19/597; H04N 19/70; H04N 19/176; H04N 19/577.

(30) Prioridade Unionista: 09/01/2014 US 14/151,586; 10/01/2013 US 61/751,211.

(73) Titular(es): QUALCOMM INCORPORATED.

(72) Inventor(es): YE-KUI WANG; YING CHEN; LI ZHANG.

(86) Pedido PCT: PCT US2014011119 de 10/01/2014

(87) Publicação PCT: WO 2014/110426 de 17/07/2014

(85) Data do Início da Fase Nacional: 10/07/2015

(57) Resumo: SÍNTESE DE VISUALIZAÇÃO EM VÍDEO 3D. Em um exemplo, um método de decodificação de dados de vídeo inclui determinar se um índice de referência para um bloco atual corresponde a uma imagem de referência de intervisualização, e quando o índice de referência para o bloco atual corresponder à imagem de referência de intervisualização, obter, a partir de uma corrente de bits codificada, dados que indicam um modo de previsão de síntese de visualização (VSP) do bloco atual, em que o modo de VSP para o índice de referência indica se o bloco atual está previsto com a previsão de síntese de visualização a partir da imagem de referência de intervisualização.

“SÍNTESE DE VISUALIZAÇÃO EM VÍDEO 3D”

[0001] O presente pedido reivindica o benefício do Pedido Provisório nº U.S. 61/751.211, depositado em 10 de janeiro de 2013, cujo conteúdo é incorporado ao presente documento a título de referência em sua totalidade.

CAMPO DA TÉCNICA

[0002] A presente revelação refere-se à codificação de vídeo.

ANTECEDENTES

[0003] Os recursos de vídeo digital podem ser incorporados em uma ampla faixa de dispositivos, incluindo televisões digitais, sistemas de difusão direta digital, sistemas de difusão sem fio, assistentes pessoais digitais (PDAs), computadores do tipo desktop ou do tipo laptop, computadores do tipo tablet, leitores de livro eletrônico, câmeras digitais, dispositivos de gravação digital, reprodutores de mídias digitais, dispositivos de jogo eletrônico, consoles de jogo eletrônico, telefones de rádio celulares ou por satélite, então chamados de "telefones inteligentes”, dispositivos de teleconferência de vídeo, dispositivos de transmissão contínua, e semelhantes. Os dispositivos de vídeo digital implantam técnicas de compressão de vídeo, tais como, aqueles descritos nos padrões definidos como MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, Parte 10, Codificação de Vídeo Avançada (AVC), o padrão de Codificação de Vídeo de Alta Eficiência (HEVC) atualmente sob desenvolvimento, e extensões de tais padrões. Os dispositivos de vídeo podem transmitir, receber, codificar, decodificar e/ou armazenar informações de vídeo digital de maneira mais eficiente implantando-se tais técnicas de compressão de vídeo.

[0004] As técnicas de compressão de vídeo

realizam previsão espacial (intrafigura) e/ou previsão temporal (interfigura) a fim de reduzir ou de remover redundância inerente em sequências de vídeo. Para codificação de vídeo com base em bloco, a fatia de vídeo (isto é, uma figura ou uma porção de a figura) pode ser particionada em blocos de vídeo, que também podem ser denominados de blocos em árvore, unidades de codificação (CUs) e/ou nós de codificação. Os blocos de vídeo em uma fatia intracodificada (I) de a figura são codificados com o uso de previsão espacial em relação às amostras de referência em blocos vizinhos na mesma figura. Os blocos de vídeo em uma fatia intercodificada (P ou B) de a figura podem usar previsão espacial em relação às amostras de referência em blocos vizinhos na mesma figura ou na previsão temporal em relação às amostras de referência em outras figurações de referência.

[0005] A previsão espacial ou temporal resulta em um bloco previsível para um bloco a ser codificado. Os dados residuais representam diferenças de pixel entre o bloco original a ser codificado e o bloco previsível. Um bloco intercodificado é codificado de acordo com um vetor de moção que aponta para um bloco de amostras de referência que forma o bloco previsível e aos dados residuais que indicam a diferença entre o bloco codificado e o bloco previsível. Um bloco intracodificado é codificado de acordo com um modo de intracodificação e com os dados residuais. Para uma compressão adicional, os dados residuais podem ser transformados a partir do domínio espacial para um domínio de transformada, resultando em coeficientes de transformada residuais que podem ser, em seguida, quantizados. Os coeficientes de transformada quantizados, dispostos inicialmente em um arranjo bidimensional, podem ser varridos a fim de produzir um vetor unidimensional de

coeficientes de transformada, e a codificação de entropia pode ser aplicada para alcançar ainda mais compressão.

SUMÁRIO

[0006] Em geral, a presente revelação descreve técnicas para codificar informações relacionadas à previsão de síntese de visualização. A previsão de síntese de visualização se refere, em geral, a uma técnica de codificação de vídeo para codificação de vídeo de multivisualização na qual um componente de visualização sintetizado é como uma figura de referência durante a codificação de uma figura de uma visualização com o uso de previsão de intervisualização. As técnicas desta revelação incluem determinar caso deve-se aplicar previsão de síntese de visualização a um particular bloco de dados de vídeo com base em uma indicação fornecida no fluxo de bits para figurações de referência de intervisualização.

[0007] Em um exemplo, um método para decodificar dados de vídeo inclui determinar se um indicie de referência para um bloco atual corresponde a uma figura de referência de intervisualização e, quando o indicie de referência para o bloco atual corresponde à figura de referência de intervisualização, obtendo, a partir de um fluxo de bits codificado, os dados que indicam um modo de previsão de síntese de visualização (VSP) do bloco atual, em que o modo de VSP para o indicie de referência indica se o bloco atual é previsto para previsão de síntese de visualização a partir da figura de referência de intervisualização.

[0008] Em outro exemplo, um aparelho para decodificar dados de vídeo inclui uma memória que armazena dados de vídeo que incluem um bloco atual e um ou mais processadores configurados para determinar se um indicie de referência para o bloco atual corresponde a uma figura de

referência de intervisualização e quando o indicie de referência para o bloco atual corresponde à figura de referência de intervisualização, obter, a partir de um fluxo de bits codificado, os dados que indicam um modo de previsão de síntese de visualização (VSP) do bloco atual, em que o modo de VSP para o indicie de referência indica se o bloco atual é previsto com uma previsão de síntese de visualização a partir da figura de referência de intervisualização.

[0009] Em outro exemplo, um método para codificar os dados de vídeo inclui determinar se um indicie de referência para um bloco atual corresponde a uma figura de referência de intervisualização, quando o indicie de referência para o bloco atual corresponde à figura de referência de intervisualização, gerar os dados que indicam um modo de previsão de síntese de visualização (VSP) do bloco atual, em que o modo de VSP para o indicie de referência indica se o bloco atual é previsto com previsão de síntese de visualização a partir da figura de referência de intervisualização, e codificar os dados que indicam o modo de VSP em um fluxo de bits.

[0010] Em outro exemplo, um aparelho para codificar dados de vídeo inclui uma memória que armazena dados de vídeo que incluem um bloco atual e um ou mais processadores configurados para determinar se um indicie de referência para um bloco atual corresponde a uma figura de referência de intervisualização, quando o indicie de referência para o bloco atual corresponde à figura de referência de intervisualização, para gerar dados que indicam um modo de previsão de síntese de visualização (VSP) do bloco atual, em que o modo de VSP para o indicie de referência indica se o bloco atual é previsto com uma previsão de síntese de visualização a partir da figura de

referência de intervisualização, e para codificar os dados que indicam o modo de VSP in a fluxo de bits.

[0011] Em outro exemplo, uma mídia legível por computador não transitória tem instruções armazenadas na mesma que, quando executadas, fazem com que um ou mais processadores determinem se um indicie de referência para um bloco atual corresponde a uma figura de referência de intervisualização e, quando o indicie de referência para o bloco atual corresponde à figura de referência de intervisualização, codifiquem os dados que indicam um modo de previsão de síntese de visualização (VSP) do bloco atual, em que o modo de VSP para o indicie de referência indica se o bloco atual é previsto com uma previsão de síntese de visualização a partir da figura de referência de intervisualização.

[0012] Em outro exemplo, um aparelho para codificar dados de vídeo inclui um meio para determinar se um indicie de referência para um bloco atual corresponde a uma figura de referência de intervisualização e, quando o indicie de referência para o bloco atual corresponde à figura de referência de intervisualização, um meio para codificar dados que indicam um modo de previsão de síntese de visualização (VSP) do bloco atual, em que o modo de VSP para o indicie de referência indica se o bloco atual é previsto com previsão de síntese de visualização a partir da figura de referência de intervisualização.

[0013] Os detalhes de um ou mais exemplos são apresentados nos desenhos anexos e na descrição abaixo. Outros recursos, objetivos e vantagens estarão aparentes a partir da descrição e dos desenhos e das reivindicações.

BREVE DESCRIÇÃO DOS DESENHOS

[0014] A Figura 1 é um diagrama de blocos que ilustra um sistema de codificação e decodificação de vídeo

exemplificativo que pode utilizar as técnicas descritas na presente revelação.

[0015] A Figura 2 é um diagrama de blocos que ilustra um codificador de vídeo exemplificativo que pode implantar as técnicas descritas na presente revelação.

[0016] A Figura 3 é um diagrama de blocos que ilustra um decodificador de vídeo exemplificativo que pode implantar as técnicas descritas na presente revelação.

[0017] A Figura 4 é um diagrama conceitual que ilustra uma codificação de uma sequência de Codificação de Vídeo de Múltiplas Visualizações (MVC).

[0018] A Figura 5 é um diagrama conceitual que ilustra um modelo de previsão de MVC.

[0019] A Figura 6 é um fluxograma que ilustra um método exemplificativo para codificar informações relacionadas à previsão de síntese de visualização.

[0020] A Figura 7 é um fluxograma que ilustra outro método exemplificativo para codificar informações relacionadas à previsão de síntese de visualização.

DESCRIÇÃO DETALHADA

[0021] Em geral, a presente revelação descreve técnicas para codificar e processar dados de vídeo de multivisualização, por exemplo, os dados de vídeo usados para produzir um efeito tridimensional (3D). Desse modo, os aspectos da presente revelação pode estar relacionados a codificação de vídeo 3D com base em codecs avançados, incluindo a codificação de duas ou mais visualizações de uma figura com mapas de profundidade. Em particular, os aspectos da presente revelação podem estar relacionados à sinalização de suporte de previsão de síntese de visualização em um processo de codificação de vídeo 3D.

[0022] As técnicas da presente revelação podem ser aplicadas a uma variedade diferentes padrões de

codificação de vídeo. Por exemplo, os padrões de codificação de vídeo incluem ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 ou ISO/IEC MPEG-2 Visual, ITU-T H.263, ISO/IEC MPEG-4 Visual e ITU-T H.264 (também conhecido como ISO/IEC MPEG-4 AVC ou H.264/AVC), incluindo as extensões de Codificação de Vídeo Escalonável (SVC) e Codificação de Vídeo de Múltiplas Visualizações (MVC) do mesmo. Um recente rascunho conjunto de MVC é descrito em “Advanced Video Coding for generic audiovisual services”, Recomendação de ITU-T H.264, março de 2010.

[0023] Além disso, há um novo padrão de codificação de vídeo, a saber HEVC, que foi desenvolvido pela Equipe de Colaboração Conjunta em Codificação de Vídeo (JCT-VC) do Grupo de Especialistas de Codificação de Vídeo (VCEG) ITU-T e do Grupo de especialistas de figura de moção (MPEG) ISO/IEC. Um Rascunho de Trabalho Recente (WD) de HEVC, é denominado de WD4 de HEVC doravante, está disponível junto a http://phenix.int-evry.fr/jct/doc_end_user/documents/6_Torino/wgl l/JCTVC-F803-v8.zip. Outro WD de HEVC, e denominado de HEVC WD9 doravante, está disponível junto a http://phenix.int-evry .fr/j ct/ doc_end_user/ documents/ 11 _Shanghai/wg 11 /JCTVC-K1003 -v 10.zip .

[0024] As recentes atividades de padronização em MPEG incluíram a Exigência de Proposta (CfP) para tecnologias de vídeo 3D. Dentre as respostas, as técnicas são selecionadas para ser incluídas no atual software de referência 3D. Um padrão potencial é uma codificação de vídeo 3D com base em H.264/AVC, que pode codificar tanto a textura quanto a profundidade para cada visualização. As ferramentas de codificação principais para a Codificação de Vídeo 3D (AVC 3D) com base em AVC (denominada no presente documento de WD4 de 3DV) são descritas no documento JCT3V-

B1002 ("3D-AVC Draft Text 4”, Equipe Colaborativa Conjunta em Desenvolvimento de Extensão de Codificação de Vídeo 3D de ITU-T SG 16 WP 3 e ISO/IEC JTC 1/SC 29/WG 11, 2ª reunião: Xangai, CN, 13 a 19 outubro de 2012, disponível publicamente junto a http://phenix.it-sudparis.eu/jct2/doc_end_user/current_document.php?id=456). A última descrição de software de referência é “Test Model for AVC based VIDEO 3D coding”, Miska M. Hannuksela, ISO/IEC JTC1/SC29/WG11 MPEG2011/N12558, San Jose, E.U.A., fevereiro de 2012. O ultimo software de referência pode ser denominado presente documento de ATM ou de 3D-ATM e está disponível publicamente em http://mpeg3dv.research.nokia.com/svn/mpeg3dv/trunk/.

[0025] Em alguns exemplos, as técnicas da presente revelação podem ser aplicadas a um padrão de codificação de VÍDEO 3D (3DV) que se baseia em H.264/ AVC. Consequentemente, para propósitos de ilustração, as técnicas da presente revelação são descritas primariamente em relação a uma extensão de 3DV de ITU-T H.264/ AVC. No entanto, deve-se entender que as técnicas podem ser aplicadas a outros padrões para codificar dados de vídeo usados para produzir um efeito tridimensional, tal como, uma extensão de VÍDEO 3D (3DV) do padrão de HEVC ou de outro padrão de codificação.

[0026] De qualquer modo, de acordo com determinados sistemas de codificação de vídeo, a estimativa de moção e a compensação de moção podem ser usadas para remover a redundância temporal em uma sequência de vídeo, de modo que a alcançar a compressão de dados. Nesse caso, pode ser gerado um vetor de moção que identifica um bloco previsível de dados de vídeo, por exemplo, um bloco de outra figura ou fatia de vídeo, que pode ser usado para prever os valores do atual bloco de vídeo que é codificado.

Os valores do bloco de vídeo previsível são subtraídos do valor do atual bloco de vídeo para produzir um bloco de dados residuais. As informações de moção (por exemplo, um vetor de moção, índices de vetor de moção, direções de previsão, ou outras informações) são comunicadas de um codificador de vídeo a um decodificador de vídeo, junto os dados residuais. O decodificador pode localizar o mesmo bloco previsível (com base no vetor de moção) e reconstruir o bloco de vídeo terminado combinando-se os dados residuais com os dados do bloco previsível.

[0027] Em relação à codificação de vídeo 3D, a fim de produzir um efeito tridimensional no vídeo, duas visualizações de uma cena, por exemplo, uma visualização de olho esquerdo e uma visualização de olho direito, podem ser mostradas simultaneamente ou quase simultaneamente. Duas figurações da mesma cena, correspondentes à visualização de olho esquerdo e à visualização de olho direito da cena, podem ser capturadas (ou geradas, por exemplo, como um gráfico gerado por computador) a partir porções horizontais levemente diferentes, representando a disparidade horizontal entre os olhos esquerdo e direito de um visualizador. Mediante a exibição dessas duas figurações simultaneamente ou quase simultaneamente, de modo que a visualização de olho esquerdo figura seja percebida pelo olho esquerdo de visualizador e a visualização de olho direito figura seja percebida pelo olho direito do visualizador, o visualizador pode experimentar um efeito de vídeo tridimensional.

[0028] As técnicas de codificação de vídeos relacionadas à codificação de multivisualização (ou codificação de visualização estérea) podem incluir, em geral, informações sobre previsão de intervisualização e/ou profundidade codificação. A previsão de intervisualização

inclui, em geral, codificar figurações de uma visualização de base com o uso de intraprevisão ou interprevisão padrão, em seguida, prever figurações de outras visualizações em relação à visualização de base ou a outras visualizações. Dessa maneira, algumas figurações de uma visualização de não base podem ser previstas em relação às figurações da visualização de base ou de outra visualização de não base.

[0029] Tipicamente, quando uma figura é codificada em relação a uma figura de referência, dispositivos de codificação de vídeo sinalizam a figura de referência como um membro de uma lista de figura de referência. Assim, para a codificação de intervisualização, as figurações de referência adicionadas à lista de figura de referência podem incluir figurações de outras visualizações, além de outas figurações da visualização atual, por exemplo, para a previsão temporal. O processo de construção de lista de figura de referência pode dispor de maneira flexível as referências de previsão temporal e de visualização. Isso pode fornecer não apenas ganho de eficiência de codificação potencial, porém também resiliência de erro, visto que os mecanismos de seleção de figura de referência e de figura redundante podem ser estendidos, em seguida, à dimensão de visualização.

[0030] No contexto atual de 3DV, os codificadores de vídeo podem codificar tanto informações de textura (por exemplo, luminância/brilho e crominância/color) quanto informações de profundidade para cada visualização de um fluxo de bits de multivisualização. Desse modo, os codificadores de vídeo de 3DV podem ser configurados para codificar mapas de profundidade. Os mapas de profundidade são figurações cujo valor de pixel representam as profundidades tridimensionais de objetos mostrados em figurações de textura correspondentes. Um mapa

de profundidade pode ser codificado como outra visualização de uma unidade acesso. Ou seja, além de figurações de textura para diferentes visualizações, uma ou mais visualizações representadas em uma unidade acesso podem incluir um include mapa de profundidade respectivo. Um mapa de profundidade para uma visualização também pode ser denominado de "componente de visualização de profundidade" da unidade acesso. O termo "componente de visualização" pode ser usado para se referir tanto ao componente de visualização de textura quanto ao componente de visualização de profundidade uma visualização particular de uma unidade acesso.

[0031] Por conseguinte, um fluxo de bits de 3DV pode incluir uma representação codificada de múltiplas visualizações, em que cada uma pode conter tanto informações de textura (componente de visualização de texturas) quanto informações de profundidade (componentes de visualização de profundidade). Embora o 3DV também contenha textura de múltiplas visualizações, em alguns casos, um fluxo de bits de 3DV pode ser compatível a um fluxo de bits de multivisualização.

[0032] Mediante o fornecimento de informações de profundidade, um dispositivo pode render figurações para várias outras visualizações além daquelas fornecidas no fluxo de bits. Por exemplo, caso duas visualizações com informações de profundidade sejam fornecidas em um fluxo de bits, um dispositivo de cliente pode usar as informações de profundidade para gerar uma a figura para uma terceira visualização, por exemplo, entre as duas visualizações codificadas do fluxo de bits. A geração de tais figurações é denominada de renderização com base em imagem de profundidade (DIBR). A DIBR pode ser usada para gerar uma visualização com base em uma textura disponível e imagens

de profundidade. Tal processo de geração é chamado de síntese de visualização.

[0033] A síntese de visualização pode ser usada para processamento posterior em seguida da decodificação das visualizações (também denominadas de síntese de visualização “pós-ciclo”). De maneira adicional ou alternativa, a síntese de visualização pode ser usada para gerar as figurações de referência usadas para predição de intervisualização (também denominada de síntese de visualização “em ciclo”). Deve-se entender que a DIBR pode ser usada para render uma figura de textura, uma figura de profundidade (tipicamente, denominada de mapa de profundidade), ou ambas a figurações. Em geral, os mapas de profundidade são representados semelhantes aos dados de luminância (luma) de figurações de textura, sem incluir crominância informações, com a exceção de que o valor de pixel do mapa de profundidade representa a profundidade para as informações de textura correspondentes, em vez das informações de brilho.

[0034] Em relação à síntese de visualização em ciclo, uma figura de referência gerada com o uso de síntese de visualização, às vezes, denominada de figura de referência de síntese de visualização (VSRP) (ou, alternativamente, previsão de figura de síntese de visualização), pode ser usada como uma figura de referência semelhante a uma figura de referência de interprevisão temporal ou a uma figura de referência de intervisualização. Uma VSRP pode estar incluída em uma lista de figura de referência. Em alguns projetos específicos de técnicas de codificação convencionais, múltiplas visualizações (com profundidade) podem ser usadas para gerar apenas uma VSRP através da DIBR. Desse modo, para uma figura que é codificada no momento, apenas uma

VSRP pode estar disponível nessas técnicas de codificação convencionais.

[0035] A previsão de síntese de visualização em ciclo (VSP) é suportada na no padrão atual de codificação de vídeo 3D com base em H.264/AVC para codificação textura aprimorada (WD4, verificado acima). A fim de possibilitar que a VSP para codificar a visualização atual, a textura codificada anteriormente e os componentes de visualização de profundidade da mesma unidade acesso podem ser usados para a síntese de visualização. Uma figura sintetizada que resulta da VSP está incluída nas listas de figura de referência iniciais (por exemplo, RefPicList0 e/ou RefPicList1) em seguidas de quadros de referência de intervisualização e temporais.

[0036] O documento “3D-CEl .a: Generalized View Syntheses Prediction (GVSP) mode”, Equipe Colaborativa Conjunta em Desenvolvimento de Extensão de Codificação de Vídeo 3D de ITU-T SG 16 WP 3 de ISO/IEC JTC 1/SC 29/WG 11 2ª reunião: Xangai, CN, 13 a 19 de outubro de 2012, JCT3V-B0054 propôs um processo de síntese de visualização denominado de previsão de síntese de visualização generalizado (GVSP). Na GVSP, um sinalizador é introduzido em múltiplos níveis, macrobloco ou inferior, para indicar se um bloco é codificado como um modo de pulo e previsto a partir da visualização figura sintetizada.

[0037] A sinalização de previsão de síntese de visualização pode se basear em um indicie de referência tradicional, que é transparente ao projeto de H.264/AVC, em termos de sinalização de nível de MB. Por exemplo, para cada MB, um codificador de vídeo pode determinar se deve ser usada a use síntese de visualização com base em um indicie de referência que identifica a figura de síntese de visualização em uma lista de figura de referência, e a

síntese de visualização indicie de referência pode ser sinalizada de maneira explícita em um cabeçalho de slice. Esse projeto pode a suportar previsão bidirecional, em que uma direção é prevista com uma figura de síntese de visualização e a outra direção é prevista a partir de uma figura temporal normal ou de uma figura de referência de intervisualização. Esse projeto também suporta múltiplas figurações de síntese de visualização a serem adaptadas no nível de fatia, assim como no nível de macrobloco.

[0038] No entanto, no projeto atual (conforme apresentado no atual software de referência de AVC 3D, denominado de 3D-ATM, conforme verificado acima), previsão de síntese de visualização com base em bloco é feita em alta velocidade. A realização da previsão de síntese de visualização com base em bloco “de alta velocidade” significa que um bloco de referência de síntese de visualização não é gerado até que a síntese de visualização bloco seja necessária. Consequentemente, uma previsão de figura de síntese de visualização (conforme a figura de referência) não está presente, por exemplo, pré-gerada e incluída em uma lista de figura de referência.

[0039] No entanto, um determinado hardware para realizar a compensação de moção pode ter um módulo de compensação de moção herdado (por exemplo, mature) que usa um quadro de referência e um vetor de moção como uma admissão. Consequentemente, pelo menos alguns módulos de compensação de moção de H.264/AVC atualmente em circulação podem não ser usados para previsão de síntese de visualização de alta velocidade, devido ao fato de que não há figura de síntese de visualização pré-gerada como uma admissão inicial.

[0040] A fim de compensar o hardware herdado, em vez de realizar a síntese de visualização de alta

velocidade, conforme descrito acima, um codificador de vídeo pode sempre gerar uma figura de síntese de visualização. Ou seja, o codificador de vídeo pode gerar uma figura de síntese de visualização para uma sequência de figurações e atribuir à figura de síntese de visualização um valor de indicie figura de referência em uma lista de figura de referência. No entanto, caso uma previsão de figura de síntese de visualização seja sempre gerada, não é possível obter a vantagem da previsão de síntese de visualização com base em bloco (“de alta velocidade”), o que pode impactar eficiência computacional e de memória. Em típicos casos, tal processo de sinalização pode levar à degradação de eficiência de codificação. Além disso, o mecanismo com base em sinalização de indicie de referência pode exigir um comando de modificação de lista de figura de referência explicito a fim de dispor a figura de síntese de visualização em uma posição desejável de uma lista de figura de referência, o que pode impactar adicionalmente a eficiência computacional.

[0041] A presente revelação reconhece determinados problemas que podem surgir com técnicas de codificação convencionais, por exemplo, de uma extensão de 3DV de H.264/AVC. Por exemplo, não é possível desligar os mecanismos de sinalização com base em indicie de referência descritos acima para a previsão de síntese de visualização. Por exemplo, pode não ser possível desligar os mecanismos de sinalização com base em indicie de referência atuais de 3D-ATM sem interromper metas de retrocompatibilidade da extensão de 3D-ATM.

[0042] Consequentemente, um novo módulo hardware que suporta a previsão de síntese de visualização com base em bloco pode ser necessário para codificadores de vídeo. Além disso, quando uma sequência de figurações tem

uma figura de síntese de visualização associada (seq_view_synthesis_flog é igual a 1), um codificador de vídeo precisa gerar um valor de indicie de referência adicional para figurações de síntese de visualização, de modo que o valor de indicie de referência esteja disponível para ser atribuído em ocorrências nas quais uma figura de síntese de visualização é gerada e usada para propósitos de previsão. Consequentemente, o codificador de vídeo gera um valor de indicie de referência adicional, até mesmo quando uma figura de síntese de visualização não é realmente gerada durante a codificação, impactando a eficiência de codificação e/ou computacional.

[0043] De acordo com aspectos da presente revelação, um indicie de referência de uma figura de VSP não é indicado especificamente em um cabeçalho de slice ou adicionado durante uma construção de lista de figura de referência. De preferência, um ou mais elementos de sintaxe (por exemplo, um sinalizador) podem ser sinalizados no MB ou no nível de partição de MB para cada indicie de referência que corresponde a uma figura de referência de intervisualização. Esse sinalizador pode indicar se a VSP é aplicada ao MB ou à partição de MB para esse indicie de referência. Em outras palavras, por exemplo, o sinalizador pode indicar se deve ser usada uma figura de referência de intervisualização para propósitos de previsão, ou se deve ser usada uma figura de VSP (por exemplo, gerada a partir da figura de referência de intervisualização) para propósitos de previsão.

[0044] O sinalizador descrito acima é diferente do vsp_flag tradicional de 3D-ATM. Por exemplo, o sinalizador descrito acima pode substituir o vsp_flag de 3D-ATM. De acordo com aspectos da presente revelação, sinalizando-se um indicie de referência da maneira descrita

acima, os diferentes indicies de referência podem ser usados para a VSP. Ou seja, por exemplo, qualquer indicie de referência de intervisualização pode ser usado para VSP, em vez de uma VSP que é vinculada a uma e um indicie de referência de VSP específico sinalizado em um cabeçalho de slice.

[0045] Em um exemplo, de acordo com aspectos da presente revelação, um Vsp_flag está presente apenas quando um indicie de referência para o bloco que é codificado no momento corresponde a uma figura de referência de intervisualização. Em outro exemplo, o Vsp_flag pode estar presente para qualquer modo interprevisto. Em outro exemplo, adicional ou alternativamente, o sinalizador pode estar presente apenas para um indicie de referência corresponde a RefPlicList0. Em outro exemplo, adicional ou alternativamente, o Vsp_flag pode estar presente para um indicie de referência correspondente tanto a RefPlicList0 quanto a RefPicList1 (se houver).

[0046] A Figura 1 é um diagrama de blocos que ilustra um sistema de codificação e decodificação de vídeo exemplificativo 10 que pode utilizar as técnicas da presente revelação para a síntese de visualização. Conforme mostrado na Figura 1, o sistema 10 inclui um dispositivo de fonte 12 que fornece dados de vídeo codificados a serem decodificados em um momento posterior por um dispositivo destino 14. Em particular, o dispositivo de fonte 12 fornece os dados de vídeo ao dispositivo destino 14 por meio de uma mídia legível por computador 16. O dispositivo de fonte 12 e o dispositivo destino 14 pode compreender qualquer dentre uma ampla faixa de dispositivos, incluindo computadores do tipo desktop, computadores do tipo notebook (isto é, do tipo laptop), computadores do tipo tablet,

caixas adaptadoras, fones de telefones, tais como, então chamados telefones “inteligentes”, então chamados teclados “inteligentes”, televisões, câmera, dispositivos de exibição, reprodutores de mídias digitais, consoles de jogo eletrônico, dispositivo de transmissão contínua de vídeo, ou semelhantes. Em alguns casos, o dispositivo de fonte 12 e o dispositivo destino 14 podem ser equipados para comunicação sem fio.

[0047] O dispositivo destino 14 pode receber os dados de vídeo codificados a ser decodificados por meio de uma mídia legível por computador 16. A mídia legível por computador 16 pode compreender qualquer tipo de mídia ou de dispositivo com capacidade para mover os dados de vídeo codificados do dispositivo de fonte 12 para o dispositivo destino 14. Em um exemplo, a mídia legível por computador 16 pode compreender uma mídia de comunicação para possibilitar que o dispositivo de fonte 12 transmita os dados de vídeo codificados diretamente ao dispositivo destino 14 em tempo real.

[0048] Os dados de vídeo codificados podem ser modulados de acordo um padrão de comunicação, tal como, um protocolo de comunicação sem fio, e transmitidos ao dispositivo destino 14. A mídia de comunicação pode compreender qualquer mídia de comunicação sem fio ou cabeada, tal como, um espectro de radiofrequência ou uma ou mais linhas de transmissão física. A mídia de comunicação pode formar parte de uma rede com base em pacote, tal como, uma rede área local, uma rede área ampla ou uma rede global, tal como, a Internet. A mídia de comunicação pode incluir roteadores, comutadores, estações-base, ou qualquer outro equipamento que pode ser útil para facilitar a comunicação do dispositivo de fonte 12 ao dispositivo destino 14.

[0049] Em alguns exemplos, dados codificados podem ser emitidos da interface de emissão 22 a um dispositivo de armazenamento. De modo semelhante, os dados codificados podem ser acessados do dispositivo de armazenamento por uma interface de admissão. O dispositivo de armazenamento pode incluir qualquer um dentre uma variedade mídias de armazenamento de dados acessados localmente ou distribuídas tal como, uma unidade disco rígido, discos Blu-ray, DVDs, CD-ROMs, memória flash, memória volátil não volátil, ou quaisquer outras mídias de armazenamento digital adequadas para armazenar dados de vídeo codificados. Em um exemplo adicional, o dispositivo de armazenamento pode corresponder a um servidor de arquivo ou a outro dispositivo de armazenamento intermediário que pode armazenar o vídeo codificado gerado pelo dispositivo de fonte 12.

[0050] O dispositivo destino 14 pode acessar os dados de vídeo armazenados do dispositivo de armazenamento por meio de transmissão contínua ou de transferência por download. O servidor de arquivo pode ser qualquer tipo de servidor com capacidade para armazenar dados de vídeo codificados e para transmitir esses dados de vídeo codificados ao dispositivo destino 14. Um servidor de arquivos exemplificativo inclui um servidor da web (por exemplo, para um site da web), um servidor FTP, dispositivos de armazenamento conectados à rede (NAS), ou uma unidade disco local. O dispositivo destino 14 pode acessar os dados de vídeo codificados através de qualquer conexão de dados armazenados, incluindo uma conexão à Internet. Isso pode inclui um canal sem fio (por exemplo, uma conexão Wi-Fi), uma conexão cabeada (por exemplo, modem DSL, modem a cabo, etc.) ou uma combinação dos mesmos que é adequada para acessar os dados de vídeo codificados

armazenados em um servidor de arquivo. A transmissão de dados de vídeo codificados do dispositivo de armazenamento pode ser uma transmissão contínua, uma transmissão por download, ou uma combinação das mesmas.

[0051] As técnicas da presente revelação não se limitam necessariamente a aplicações ou regulagens sem fio. As técnicas podem ser aplicadas à codificação de vídeo para apoiar qualquer uma dentre uma variedade aplicações multimídia, tal como, difusão área de televisão, transmissões de televisão a cabo, transmissões de televisão a satélite, transmissão contínua de vídeo pela Internet, tal como, transmissão contínua de adaptação dinâmica sobre HTTP (DASH), vídeo digital que é codificado em uma mídia de armazenamento de dados, decodificação de vídeo digital armazenado em uma mídia de armazenamento de dados, ou outras aplicações. Em alguns exemplos, o sistema 10 pode ser configurado para suportar uma transmissão de vídeo unidirecional ou bidirecional a fim de suportar aplicações, tais como, transmissão contínua de vídeo, reprodução de vídeo, difusão de vídeo, e/ou vídeo telefonia.

[0052] No exemplo da Figura 1, o dispositivo de fonte 12 inclui uma fonte de vídeo 18, um codificador de vídeo 20 e uma interface de emissão 22. O dispositivo destino 14 inclui uma interface de admissão 28, um decodificador de vídeo 30 e um dispositivo de exibição 32. Em conformidade com a presente revelação, o codificador de vídeo 20 de dispositivo de fonte 12 pode ser configurado para aplicar as técnicas para a previsão de vetor de moção na codificação de multivisualização. Em outros exemplos, um dispositivo de fonte e um dispositivo destino podem incluir outros componentes ou disposições. Por exemplo, o dispositivo de fonte 12 pode receber dados de vídeo de uma fonte de vídeo externa 18, tal como, uma câmera externa. De

igual modo, o dispositivo destino 14 pode fazer interface com um dispositivo de exibição externo, em vez de incluir um dispositivo de exibição integrado.

[0053] O sistema ilustrado 10 da Figura 1 é meramente um exemplo. As técnicas para a síntese de visualização podem ser realizadas por qualquer dispositivo de codificação e/ou decodificação de vídeo digital. Embora, em geral, as técnicas da presente revelação sejam realizadas por um dispositivo de codificação de vídeo, as técnicas também podem ser realizadas por um codificador/decodificador de vídeo, denominado, tipicamente, de "CODEC." Ademais, as técnicas da presente revelação também podem ser realizadas por um pré-processador de vídeo. O dispositivo de fonte 12 e o dispositivo destino 14 são meros exemplos de tais dispositivos de codificação nos quais o dispositivo de fonte 12 gera os dados de vídeo codificados para a transmissão ao dispositivo destino 14. Em alguns exemplos, os dispositivos 12, 14 podem operar de maneira substancialmente simétrica de modo que cada um dentre os dispositivos 12, 14 incluam componentes de codificação e decodificação de vídeo. Por conseguinte, o sistema 10 pode suportar uma transmissão de vídeo unidirecional ou bidirecional entre dispositivos de vídeo 12, 14, por exemplo, para a transmissão contínua de vídeo, reprodução de vídeo, difusão de vídeo, ou vídeo telefonia.

[0054] A fonte de vídeo 18 do dispositivo de fonte 12 pode incluir um dispositivo de captura de vídeo, tal como, uma câmera de vídeo, um arquivo de vídeo que contém um vídeo capturado anteriormente e/ou uma interface de alimentação de vídeo para receber vídeo de um fornecedor de conteúdo de vídeo. Com uma alternativa adicional, uma fonte de vídeo 18 pode gerar dados com base em gráfica de

computador como o vídeo de fonte, ou uma combinação de vídeo ao vivo, vídeo arquivado e vídeo gerado por computador. Em alguns casos, caso a fonte de vídeo 18 seja uma câmera de vídeo, um dispositivo de fonte 12 e um dispositivo destino 14 pode formar então chamados fones de câmera ou fones de vídeo. Conforme mencionado acima, no entanto, as técnicas descritas na presente revelação podem ser aplicáveis à codificação de vídeo em geral e podem ser aplicadas a aplicações sem fio e/ou cabeadas. Em cada caso, o vídeo capturado, pré-capturado ou gerado por computador pode ser codificado por um codificador de vídeo 20. As informações de vídeo codificadas podem ser, então, emitidas por uma interface de emissão 22 a uma mídia legível por computador 16.

[0055] A mídia legível por computador 16 pode incluir mídias transientes, tais como, uma transmissão por difusão sem fio ou por rede cabeada, ou mídias de armazenamento (ou seja, mídias de armazenamento não transitórias), tais como, um disco rígido, unidade flash, um disco compacto, disco de vídeo digital, disco Blu-ray ou outras mídias legíveis por computador. Em alguns exemplos, um servidor de rede (não mostrado) pode receber os dados de vídeo codificados do dispositivo de fonte 12 e fornecer os dados de vídeo codificados ao dispositivo destino 14, por exemplo, por meio de transmissão de rede. De modo semelhante, um dispositivo de computação de uma instalação de produção mídia, tal como, uma instalação de estampagem de disco, pode receber os dados de vídeo codificados do dispositivo de fonte 12 e produzir um disco que contém os dados de vídeo codificados. Portanto, pode-se entender que a mídia legível por computador 16 inclui uma ou mais mídias legíveis por computador de várias formas, em vários exemplos.

[0056] A interface de admissão 28 de um dispositivo destino 14 recebe informações da mídia legível por computador 16. As informações de mídia legível por computador 16 podem incluir informações de sintaxe definidas por um codificador de vídeo 20, que é também usado pelo decodificador de vídeo 30, que inclui elementos de sintaxe que descrevem as características e/ou o processamento de blocos e outras unidades codificadas, por exemplo, GOPs. O dispositivo de exibição 32 exibe os dados de vídeo decodificados a um usuário e pode compreender qualquer um dentre variedade dispositivos de exibição tal como, um tubo de rádio catódico (CRT), a visor de cristal líquido (LCD), um visor de plasma, um visor com diodo de emissor de luz orgânica (OLED), ou outro tipo de dispositivo de exibição.

[0057] O codificador de vídeo 20 e o decodificador de vídeo 30 podem ser implantados, cada um, com qualquer um dentre uma variedade conjunto de circuitos codificadores decodificadores adequados, conforme aplicável, tais como, um ou mais microprocessadores, processadores de sinal digital (DSPs), circuitos integrados de aplicação específica (ASICs), matrizes de porta programáveis em campo (FPGAs), conjunto de circuitos de lógica distinta, software, hardware, firmware ou combinações dos mesmos. Cada um dentre o codificador de vídeo 20 e o decodificador de vídeo 30 pode ser incluído em um ou mais codificadores ou decodificadores, em que cada um pode ser integrado como parte de um codificador/decodificador de vídeo combinado (CODEC). Um dispositivo que inclui um codificador de vídeo 20 e/ou decodificador de vídeo 30 pode compreender um circuito integrado, um microprocessador e/ou um dispositivo de comunicação sem fio, tal como, um telefone celular.

[0058] Embora não mostrado na Figura 1, em alguns aspectos, o codificador de vídeo 20 e o decodificador de vídeo 30 pode ser integrados, cada um, a um codificador de áudio e a um decodificador e podem incluir unidade MUX-DEMUX apropriadas, ou outro hardware e software, a fim de manipular o codificação tanto do áudio quanto do vídeo em um fluxo de dados comum ou em fluxos de dados separados. Caso aplicável, as unidades de MUX-DEMUX podem se conformar ao protocolo multiplexador de ITU H.223, ou a outros protocolos, tais como, o protocolo de datagrama de usuário (UDP).

[0059] A presente revelação pode se referir, em geral, ao codificador de vídeo 20 que “sinaliza” determinadas informações para outro dispositivo, tal como, o decodificador de vídeo 30. Deve-se entender, no entanto, que o codificador de vídeo 20 pode sinalizar informações associando-se determinados de sintaxe com várias porções codificadas de dados de vídeo. Ou seja, o codificador de vídeo 20 pode “sinalizar” os dados armazenando-se determinados elementos de sintaxe aos cabeçalhos de várias porções codificadas de dados de vídeo. Em alguns casos, tais elementos de sintaxe podem ser codificados e armazenados (por exemplo, armazenados ao dispositivo de armazenamento 24) antes de serem recebidos e decodificados pelo decodificador de vídeo 30. Assim, o termo “sinalização” pode se referir, em geral, à comunicação de sintaxe ou a outros dados para decodificar os dados de vídeo comprimidos, caso tal comunicação ocorra em tempo real ou quase real ou em um período de tempo, como pode ocorrer durante o armazenamento de elementos de sintaxe em uma mídia ao mesmo tempo do codificação, que podem, então, ser recuperados por um dispositivo decodificação a qualquer momento após ser armazenado nessa mídia.

[0060] O codificador de vídeo 20 e o decodificador de vídeo 30 podem operar de acordo com uma variedade padrões de codificação de vídeo. Por exemplo, o codificador de vídeo 20 e o decodificador de vídeo 30 pode operar de acordo com o padrão ITU-T H.264/MPEG-4 (AVC), que foi formulado pelo Grupo de Especialistas de Codificação de Vídeo (VCEG) ITU-T junto do Grupo de Especialistas de Figura Móvel ISO/IEC (MPEG) como o produto de uma parceira conhecida como Equipe de Vídeo Conjunta (JVT). Em alguns aspectos, as técnicas descritas na presente revelação podem ser aplicadas aos dispositivos que se conforma, em geral, ao padrão H.264. O padrão H.264 é descrito na Recomendação de ITU-T H.264, Codificação de Vídeo Avançada para serviços audiovisual genéricos, pelo Grupo de Estudo ITU-T, e datado de março de 2005, que pode ser denominado no presente documento de padrão H.264 ou especificação de H.264, ou de padrão ou especificação de H.264/AVC.

[0061] A Equipe de vídeo conjunta (JVT) continua a trabalhar em extensões para H.264/MPEG-4 AVC. Em particular, um padrão potencial é uma codificação de vídeo 3D com base no padrão H.264/AVC que fornece codificação tanto da textura quando da profundidade para cada visualização de dados de vídeo. As ferramentas de codificação principais para WD4 de 3DV podem ser descrias no documento JCT3V-B1002, conforme verificado acima. Além disso, o software de referência mais recente pode ser denominado no presente documento de ATM ou 3D-ATM, conforme verificado acima.

[0062] As técnicas da presente revelação, no entanto, não se limitam a qualquer padrão de codificação particular. Outros exemplos de padrões de codificação de vídeo incluem MPEG-2 e ITU-T H.263, e o padrão de Codificação de Vídeo de Alta Eficiência (HEVC). Por

exemplo, em alguns exemplos, o codificador de vídeo 20 e o decodificador de vídeo 30 pode realizar a codificação de vídeo de multivisualização (MVC), e pode implantar uma extensão de multivisualização de HEVC, então chamada MV-HEVC, ou um codec de 3DV completo com base em HEVC com profundidade aprimorada, 3D-HEVC.

[0063] Em H.264/AVC, cada macrobloco interprevisto (MB) pode ser particionado de uma dentre quatro maneira diferentes, incluindo uma partição de MB 16x16, duas partições de MB 16x8, duas partições de MB 8x16 e quatro partições de MB 8x8. As diferentes partições de MB em um MB ode ter um valor de indicie de referência diferente para cada direção (RefPicList0 ou RefPicList1). Quando um MB não é particionado e quatro partições de MB 8x8, o MB pode ter apenas um vetor de moção para toda a partição de MB em cada direção. Quando um MB é particionado em quatro partições de MB 8x8, cada partição de MB 8x8 pode ser particionada adicionalmente em sub-blocos, em que cada um pode ser um vetor de moção diferente em cada direção. Pode haver quatro maneiras de particionar uma partição de MB 8x8 em sub-blocos, incluindo um sub-bloco 8x8, duas sub-blocos 8x4, dois sub-blocos4x8 e quatro sub-blocos 4x4. Cada sub-bloco pode ter um vetor de moção diferente e0m cada direção.

[0064] Em geral, o ciclo de compensação de moção de H.264/AVC e HEVC são iguais. Por exemplo, a reconstrução de um quadro atual / no ciclo de compensação de moção pode ser igual aos coeficientes dequantizados r mais a previsão temporal P:

$$\hat{I} = r + P.$$

Na fórmula acima, P indica interprevisão uni-previsível para P quadros ou interprevisão biprevisível

para B quadros.

[0065] No entanto, a unidade de compensação de moção em HEVC é diferente daquela em padrões de codificação de vídeo anteriores. Por exemplo, o conceito de um macrobloco em padrões de codificação de vídeo anteriores não existe em HEVC. Preferencialmente, os macrobloco são substituídos por uma estrutura hierárquica flexível com base em um esquema de árvore quádrupla genérico. Dentro desse esquema, três tipos de blocos, isto é, unidades de codificação (CUs), unidades de previsão (PUs) e unidades de transformada (TUs), são definidos. Uma CU é uma unidade básica de divisão de região. O conceito e uma CU é análogo ao conceito de um macrobloco, porém uma CU não é restrita a um tamanho máximo e uma CU permite e divisão recursiva em quatro CUs igualmente dimensionadas para melhorar o conteúdo de maneira adaptada. Uma PU é uma unidade básica de inter/intraprevisão. Em alguns exemplos, uma PU pode conter múltiplas partições conformadas de maneira arbitrária em uma única PU a fim de codificar de maneira eficaz modelos de imagens irregulares. Uma TU é uma unidade básica de transformada. As TUs de uma CU pode ser definida independentemente das PUs da CU. No entanto, um tamanho de uma TU é limitado à CU à qual a TU pertence. Essa separação da estrutura de bloco em três diferentes conceitos pode permitir que cada um seja otimizado de acordo com a função do mesmo, o que pode resultar em uma eficiência de codificação melhorada.

[0066] Em qualquer evento, nas especificações de codificação de vídeo, uma sequência de vídeo inclui tipicamente uma série de figurações. As figurações podem também ser chamadas de "quadros." Uma figura pode incluir três amostras arranjos, denotados de S_L, S_{Cb} e $S_{Cr.}$ O S_L é um arranjo bidimensional (isto é, um bloco) de amostras de

luma. O S_{Cb} é um arranjo bidimensional de amostras de crominância de Cb. S_{Cr} é um arranjo bidimensional de amostras de crominância de Cr. As amostras de crominância podem ser também referidas no presente documento como amostras “croma”. Em outros exemplos, uma figura pode ser monocromática e pode incluir apenas um arranjo de amostras de luma.

[0067] Em alguns exemplos, as figurações podem ser divididas em um número de fatias, em que cada fatia que inclui um número de macroblocos ou codificação de três unidades (CTUs, em relação ao HEVC). Uma fatia pode incluir um número inteiro de macroblocos ou CTUs organizados consecutivamente em uma organização de varredura de dados raster. Uma fatia codificada pode compreender um cabeçalho de slice e dados de fatia. Um cabeçalho de slice de uma fatia pode ser a estrutura de sintaxe que inclui elementos de sintaxe que fornecem informações sobre a fatia.

[0068] Essa revelação pode usar o termo "unidade de vídeo" ou "bloco de vídeo" ou "bloco" para se referir a um ou mais blocos de amostra e estruturas de sintaxe usadas para amostras de código do um ou mais blocos de amostras. Os tipos de exemplo de unidades de vídeo ou blocos pode incluir macroblocos, partições de macrobloco, CTUs, CUs, PUs, unidades transformadas (TUs) e assim por diante. Blocos de dados de vídeo podem geralmente ser referidos como no "nível de bloco," contra níveis mais altos tal como um nível de fatia, nível de figura ou nível de sequência.

[0069] Codificador de vídeo 20 pode usar intraprevisão ou intraprevisão para gerar um bloco de previsão. Se o codificador de vídeo 20 usar a intraprevisão para gerar o bloco de previsão, o codificador de vídeo 20 pode gerar o bloco de previsão com base em amostras de

figura associado ao bloco que é codificado. Nessa revelação, a frase "com base em" pode indicar "com base em pelo menos uma parte".

[0070] Se o codificador de vídeo 20 usar a intraprevisão para gerar os blocos de previsão, o codificador de vídeo 20 pode gerar os blocos de previsão com base nas amostras decodificadas de uma ou mais figurações além da figura que é codificada atualmente.

Quando a intraprevisão for usada para gerar os blocos de previsão de um bloco, essa revelação pode se referir ao bloco as "intracodificado" ou "interprevisto". A intraprevisão pode ser uniprevista (isto é, uniprevisão) ou biprevista (isto é, biprevisão). Para realizar a uniprevisão ou a biprevisão, o codificador de vídeo 20 pode gerar uma primeira lista de referência de figura (RefPicList0) e uma segunda lista de referência de figura (RefPicList1) para uma figura atual. Cada uma das listas de figura de referências pode incluir uma ou mais figurações de referência. Após uma lista de referência de figura ser construída (nomeadamente RefPicList0 e RefPicList1 se disponível), um índice de referência a uma lista de referência de figura pode ser usada para identificar qualquer figura de referência incluída na lista de referência de figura.

[0071] Quando se usa a uniprevisão, o codificador de vídeo 20 pode pesquisar as figurações de referência em qualquer um ou em ambos RefPicList0 e RefPicList1 para determinar uma local de referência dentro de uma figura de referência. Adicionalmente, quando se usa a uniprevisão, o codificador de vídeo 20 pode gerar, com base em pelo menos parte das amostras que correspondem à local de referência, os blocos de previsão. Além disso, quando se usa a uniprevisão, o codificador de vídeo 20 pode

gerar um vetor único de movimento que indica um deslocamento espacial entre um bloco de previsão e a local de referência. O vetor de movimento pode incluir a componente horizontal que especifica um deslocamento horizontal entre um bloco de previsão e a local de referência e pode incluir um componente vertical que especifica um deslocamento vertical entre um bloco de previsão e um local de referência.

[0072] Quando se usa a biprevisão para codificar um bloco, o codificador de vídeo 20 pode determinar uma primeira local de referência em uma figura de referência em RefPicList0 e uma segunda local de referência em uma figura de referência in RefPicList1. O codificador de vídeo 20 pode gerar, com base em pelo menos parte das amostras que correspondem a primeira e a segunda local de referências, os blocos de previsão. Além disso, quando se usa a biprevisão, o codificador de vídeo 20 pode gerar um primeiro vetor de movimento que indica um deslocamento espacial entre um bloco de previsão e a primeira local de referência e um segundo vetor de movimento que indica um deslocamento espacial entre o bloco de previsão e a segunda local de referência.

[0073] Seguindo a intraprevisão ou a codificação de interprevisão, o codificador de vídeo 20 pode calcular os dados residuais para o bloco que é codificado atualmente. Os dados residuais podem corresponder às diferenças de pixel entre os pixels de figura não codificada e os blocos de previsão. O codificador de vídeo 20 pode, então, transformar os dados residuais. Por exemplo, o codificador de vídeo 20 pode transformar os dados de vídeo do domínio espacial ao domínio de transformada pela aplicação, por exemplo, de uma transformada discreta de cosseno (DCT), uma transformada

inteira, uma transformada de wavelet ou uma transformada conceitualmente similar.

[0074] Seguindo quaisquer transformadoras para produzir os coeficientes de transformada, o codificador de vídeo 20 pode realizar a quantização dos coeficientes de transformada. A quantização se refere geralmente a um processo em que os coeficientes de transformada são quantizados para reduzir possivelmente a quantidade de dados usada para representar os coeficientes, que fornece um entendimento adicional. O processo de quantização pode reduzir a profundidade de bit associada com alguns ou todos os coeficientes. Por exemplo, um valor de n-bit pode ser arredondado a um valor de m-bit durante a quantização, em que n é maior que m.

[0075] Seguindo a quantização, o codificador de vídeo 20 pode varrer os coeficientes de transformada, que produz um vetor unidimensional da matriz bidimensional que inclui os coeficientes de transformada quantizada. A varredura pode ser projetada para colocar coeficientes de energia mais alta (e, portanto, frequência menor) na parte dianteira do arranjo e para colocar os coeficientes de energia menor (e, portanto, frequência maior) na parte traseira do arranjo. Em alguns exemplos, o codificador de vídeo 20 pode utilizar uma organização de varredura predefinida para varrer os coeficientes de transformada quantizada para produzir um vetor serializado que pode ser codificado por entropia. Em outros exemplos, o codificador de vídeo 20 pode realizar uma varredura adaptativa.

[0076] Após a varredura os coeficientes de transformada quantizada para formar um vetor unidimensional, o codificador de vídeo 20 pode ser codificar por entropia o vetor unidimensional, por exemplo, de acordo com codificação por comprimento variável

adaptativo ao contexto (CAVLC), codificação de aritmética binária de contexto adaptativo (CABAC), codificação de aritmética binária de contexto adaptativo com base em sintaxe (SBAC), codificação de Entropia de Partição de Intervalo de Probabilidade (PIPE) ou outra metodologia de codificação por entropia. O codificador de vídeo 20 pode ser também elementos de sintaxe de código por entropia associados aos dados de vídeo codificados para o uso pelo decodificador de vídeo 30 na decodificação dos dados de vídeo.

[0077] O codificador de vídeo 20 pode adicionalmente enviar dados de sintaxe, tal como dados de sintaxe com base em bloco, dados de sintaxe com base em figura e dados de sintaxe com base em GOP para o decodificador de vídeo 30, por exemplo, em uma figura cabeçalho, um bloco cabeçalho, um cabeçalho de slice ou um Cabeçalho GOP. Os dados de sintaxe GOP podem descrever uma diversidade de figurações no GOP respectivo e os dados de sintaxe de figura podem indicar um modo de codificação/previsão usado para codificar a figura correspondente.

[0078] O codificador de vídeo 20 pode emitir um fluxo de bits que inclui uma sequência de bits que forma uma representação de dados de vídeo (isto é, figurações codificadas e dados associados). O fluxo de bits pode compreender uma sequência de unidades de camada de abstração de rede (NAL). Uma Unidade de NAL é uma estrutura de sintaxe que contém uma indicação do tipo de dados na Unidade de NAL e dos bytes que contém tais dados na forma de carga de sequência de byte bruto (RBSP) intercalada conforme necessário com os bits de prevenção de emulação. Cada uma das unidades de NAL inclui um cabeçalho de unidade de NAL e encapsula um RBSP. O cabeçalho de unidade de NAL

pode incluir um elemento de sintaxe que indica um código de tipo de unidade NAL. O código de tipo de unidade NAL especificado pelo cabeçalho de unidade de NAL de uma unidade de NAL indica o tipo da unidade de NAL. Uma RBSP pode ser uma estrutura de sintaxe que contém um número inteiro de bytes que é encapsulado dentro de uma unidade de NAL. Em alguns exemplos, uma RBSP que inclui zero bit.

[0079] Tipos diferentes de unidades de NAL podem encapsular tipos diferentes tipos diferentes de RBSPs. Por exemplo, tipos diferentes de unidades de NAL podem encapsular RBSPs diferentes para conjuntos de parâmetro de sequência (SPSs), conjuntos de parâmetro de figura (PPSs), fatias codificadas, mensagens de informações de melhoria suplementar (SEI) e assim por diante. As unidades de NAL que encapsulam as RBSPs para dados de codificação de vídeo (conforme oposto as RBSPs para conjuntos de parâmetro e mensagens de SEI) podem chamadas de unidades de NAL de camada de codificação de vídeo (VCL).

[0080] O decodificador de vídeo 30 pode receber um fluxo de bits gerado pelo codificador de vídeo 20. Além disso, o decodificador de vídeo 30 pode analisar o fluxo de bits para obter os elementos de sintaxe do fluxo de bits. O decodificador de vídeo 30 pode reconstruir as figurações dos dados de vídeo com base em pelo menos partes dos elementos de sintaxe obtidos do fluxo de bits. O processo para reconstruir os dados de vídeo pode ser geralmente recíproco ao processo pré-formado pelo codificador de vídeo 20.

[0081] Por exemplo, o decodificador de vídeo 30 pode usar os vetores de movimento para determinar os blocos de previsão para um bloco que é decodificado atualmente (um bloco atual). Além disso, decodificador de vídeo 30 pode inverter os coeficientes quantizados de

transformada do bloco atual. O decodificador de vídeo 30 pode aplicar uma transformada inversa nos coeficientes para determinar os dados residuais para o bloco atual. O decodificador de vídeo 30 pode reconstruir o bloco atual pela adição das amostras do bloco de previsão para corresponder às amostras residuais. Reconstruindo-se cada um dos blocos para uma figura, o decodificador de vídeo 30 pode reconstruir a figura.

[0082] Em geral, para realizar a previsão de síntese de visualização, um codificador de vídeo, tal como o codificador de vídeo 20 ou o decodificador de vídeo 30, pode usar as informações de textura e profundidade de uma ou mais visualizações para gerar um componente de visualização sintetizada. Por exemplo, o VSP pode se referir à geração de um quadro de referência sintetizado adicional para o propósito de previsão dos dados de vídeo. Uma versão virtual de um quadro pode ser sintetizada para ser codificado a partir dos quadros codificados anteriormente de outras câmeras e pode usar o virtual quadro como uma referência de previsão.

[0083] Esse processo é substancialmente similar ao processo para gerar um componente de visualização sintetizada com o uso das informações de textura e profundidade para propósitos de exibição. No entanto, na previsão de síntese de visualização, um componente de visualização sintetizada (isto é, uma figura) pode ser adicionado a uma lista de referência de figura para o uso como uma figura de referência quando a codificação de uma visualização diferente. A previsão de síntese de visualização pode ser usada para gerar as figurações de referência para informações de textura ou informações de profundidade de codificação. Evidentemente, as figurações de síntese de visualização geradas podem

também ser exibidas por um dispositivo de cliente como parte de reprodução. Adicional ou alternativamente, o decodificador de vídeo 30 ou outra unidade de dispositivo de destino 14, tal como uma unidade de pós-processamento (não mostrado), pode realizar a síntese de visualização para outros propósitos da geração de figurações sintetizadas para a reprodução. Deve-se compreender que, embora as figurações de síntese de visualização usadas para a previsão possam ser exibidas, em alguns exemplos, uma figura separada pode ser gerada com o uso de DIBR para os propósitos de exibição. As imagens de síntese de visualização para os propósitos de previsão podem ser geradas durante a codificação ou “em ciclo”, ao mesmo tempo em que imagens de síntese de visualização para os propósitos de exibição podem ser geradas “em ciclo” o “pós-ciclo” isto é, seguindo a codificação.

[0084] Para gerar uma imagem de referência adicional, um processo de distorção 3D ou um processo de interpolação de visualização podem ser usados. Por exemplo, um processo de distorção 3D pode incluir realizar uma projeção com o uso de informações de profundidade e os parâmetros de câmera. Um processo de interpolação de visualização pode incluir informações de disparidade de exploração entre imagens de ponto de vista adjacente. Em geral, para o VSP, há dois tipos de técnicas de distorção de imagem, a saber, distorção para frente e distorção para trás, dependendo da disponibilidade do mapa de profundidade da visualização atual. A distorção para frente gera uma visualização sintética quando o mapa de profundidade do ponto de vista de referência está disponível. Nesse exemplo, o mapa de profundidade do ponto de vista de referência é codificado/decodificado antes de codificação/decodificação do componente de textura da

visualização atual. Uma derivação similar pode ser obtida para distorção para trás com o uso da profundidade da visualização a ser sintetizada.

[0085] O VSP em ciclo é suportado padrão de codificação de vídeo 36 com base em H.264/AVC atual para codificação de textura melhorada codificação de textura (WD4, observado acima). Para permitir que o VSP codifique a visualização atual, a textura codificada anteriormente e os componentes de visualização de profundidade da mesma unidade de acesso podem ser usadas para a síntese de visualização. Uma figura sintetizada que resulta do VSP é incluída na lista de referência de figurações inicial (por exemplo, RefPicList0 e/ou RefPicList1) que segue os quadros de referência de visualização temporal e intravisualização.

[0086] Conforme observa-se acima, o documento "3D-CEl.a: Generalized View Synthesis Prediction (GVSP) mode," Joint Collaborative Team on 3D Video Coding Extension Development of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 2ª Reunião: Shanghai, CN, a 19 de outubro de 2012, no qual o JCT3V-B0054 propôs introduzir um sinalizador em múltiplos níveis, macrobloco ou menores, para indicar se um bloco é codificado como o modo de pulo e previsto a partir da visualização figura sintetizada.

[0087] Um exemplo de conjunto de elemento de sintaxe é estabelecido na Tabela 1 a Tabela 3 abaixo. Por conveniência editorial, os sinalizadores podem ser chamados como sinalizadores de vsp. Um exemplo de sintaxe de camada de macrobloco é mostrado na Tabela 1 abaixo:

TABELA 1 - SINTAXE DE CAMADA DE MACROBLOCO

macrobloco_layer() {		Descritor
if (slice_vsp_flag && VspRefExist)		

vsp_mb_flag		ae(v)
if (! vsp_mb_flag)		
mb_type		ue(v)|ae(v)
if mb_type = = I_PCM) {		
while(!byte_aligned())		
pcm_alignment_zero_bit		f(1)
for(i = 0; i < 256; i++)		
pcm_sample_luma[i]		u(v)
for(i = 0: i < 2 * MbWidthC * MbHcightC: i++)		
pcm_sample_croma[i]		u(v)
} else {		
noSubMbPartSizeLessThan8x8Flag = 1		
if(mb_type != I_NxN && MbPartPredMode(mb_type, 0) != Intra_16x16 && NumMbPart(mb_type) = = 4){		
if (!vsp_mb_flag)		
sub_mb_pred(mb_type)		
for(mbPartIdx = 0; mbPartldx <4; mbPkutldx++)		
if(sub mb_type[mbPartldx] != B_Direct_8x8) {		
it NumSubMbPart(sub_mb_type[mbPartldx]) >1)		
noSubMbPartSizeLessThan8x8Flag = 0		
} else if(!direct_8x8_inference_flag)		
noSublvlbPartSizeLessThan8x8Flag = 0		
} else {		
if(transform_8x8_mode_flag && mb_type = I NxN)		
transform_size_8x8_flag		u(1)|ae(v)

if(!vsp_mb_flag)		
mb_pred(mb_type)		
}		
if(MbPartPredMode(mb_type, 0)!= Intra_16x16){		
Coded_block_pattern		me(v)|ae(v)
if(CodedBlockPatternLuma>0 && transform_8x8_mode_flag && mb_type !=I_NxN && noSubMbPartSizeLessThan8x8Flag && (vsp_mb_flag || (!vsp_mb_flag && (mb_type != B_Direct_16x16 || direct_8x8_inference_flag))))		
transform_size_8x8_flag		u(1)| ae(v)
}		
if(CodedBlockPattemLuma > 0 || CodedBlockPatternCroma > 0 || MbPartPredMode(mb type, 0) = = Intra_16x16){		
mb_qp_delta		se(v)|ae(v)
residual(0, 15)		
}		
}		
}		

[0088] No exemplo da Tabela 1 acima, o vsp_mb_flag igual a 1 indica que o MB inteiro é previsto a partir da figura de VSP. Esse sinalizador igual a 0 indica que o MB inteiro pode ser previsto por outros modos. Quando esse sinalizador é igual a 1, o MB_type não é sinalizado. Quando não está presente, esse é inferido como igual a 0.

[0089] Um exemplo de sintaxe de previsão de macrobloco conforme mostrado na Tabela 2 abaixo:

TABELA 2 - Sintaxe de previsão de macrobloco

mb_pred(mb_type) {	C	Descritor
if(MbPartPredMode(mb_type, 0) == Intra_4x4 || MbPartPredMode(mb_type, 0) == Intra_16x16) {		
if(MbPartPredMode(mb_type, 0) == Intra_4x4)		
for(luma4x4BlkIdx=0; luma4x4BlkIdx<16; luma4x4BlkIdx++) {		
prev_intra4x4_pred_mode_flag[luma4x4BlkIdx]	2	u(1) | ae(v)
if(!prev_intra4x4_pred_mode_flag[luma4x4BlkIdx])		
rem_intra4x4_pred_mode[luma4x4BlkIdx]	2	u(3) | ae(v)
}		
intra_chroma_pred_mode	2	ue(v) | ae(v)
} else if(MbPartPredMode(mb_type, 0) != Direct) {		
for(mbPartIdx = 0; mbPartIdx < NumMbPart(mb_type); mbPartIdx++)		

if (slice_vsp_flag && NumMbPart(mb_type) != 1 && VspRefExist) //vsp_mb_flag is not 1		
mb_part_vsp_flag[mbPartIdx]	2	ae(v)
if(!mb_part_vsp_flag[mbPartIdx] &&(num_ref_idx_l0_activc_minus1 > 0 || mb_field_decoding_flag) && MbPartPredMode(mb_type, mbPartIdx) != Pred_L1)		
ref_idx_l0[mbPartIdx]	2	te(v) | ae(v)
for (mbPartIdx = 0; mbPartIdx < NumMbPart(mb_type) && !mb_part_vsp_flag[mbPartIdx]; mbPartIdx++)		
if((num_ref_idx_l1_active_minus1 > 0 || mb_field_decoding_flag) && MbPartPredMode(mb_type, mbPartIdx) != Pred_L0)		
ref_idx_l1[mbPartIdx]	2	te(v) | ae(v)
for(mbPartIdx = 0; mbPartIdx < NumMbPart(mb_type); mbPartIdx++)		
if(!mb_part_vsp_flag[mbPartIdx] && MbPartPredMode (mb_type, mbPartIdx) != Pred_L1)		
for(compIdx = 0; compIdx < 2; compIdx++)		
mvd_l0[mbPartIdx][0][compIdx]	2	se(v) | ae(v)
for(mbPartIdx = 0; mbPartIdx < NumMbPart(mb_type); mbPartIdx++)		
if(!mb_part_vsp_flag[mbPartIdx] && MbPartPredMode(mb_type, mbPartIdx) != Pred_L0)		
for(compIdx = 0; compIdx < 2; compIdx++)		
mvd_l1[mbPartIdx][0][compIdx]	2	se(v) | ae(v)
}		
}		

[0090] No exemplo da Tabela 2 acima, mb_part_vsp_flag[mbPartldx] igual a 1 indica que a partição de MB atual é prevista de uma figura de VSP. Esse sinalizador igual a 0 indica que a partição de MB inteira não é prevista de uma figura de VSP. Quando esse sinalizador não for apresentado, esse sinalizador será inferido como igual a 0.

[0091] Em alguns exemplos, se o número de partição de MB for menor que 4, noSubMbPartSizeLessThan8x8Flag será verdadeira (a mesma que

em H.264/AVC). Se o número de partição de MB for igual a 4, um elemento de sintaxe noSubMbPartSizeLessThan8x8Flag pode ser derivado pela verificação de partições de MB de 8x8 com mb_part_vsp_flag igual a 0, conforme segue:

1. Se qualquer Partição de MB tiver uma subpartição de MB menor que 8x8, noSubMbPartSizeLessThan8x8Flag será falso, retorno

2. noSubMbPartSizeLessThan8x8Flag será verdadeiro.

[0092] Um exemplo da Sintaxe de previsão de sub-macrobloco é mostrado na Tabela 3 abaixo:

TABELA 3 - Sintaxe de previsão de sub-macrobloco

sub_mb_pred(mb_type) {	C	Descritor
for(mbPartIdx = 0; mbPartIdx < 4; mbPartIdx++) {		
if (slice_vsp_flag && VspRefExist) // vsp_mb_flag is not 1		
sub_mb_vsp_flag[mbPartIdx]		
if (!sub_mb_vsp_flag[mbPartIdx])		
sub_mb_type[mbPartIdx]	2	ue(v) | ae(v)
}		
for(mbPartIdx = 0; mbPartIdx < 4; mbPartIdx++)		
if(!sub_mb_vsp_flag[mbPartIdx] &&(num_ref_idx_l0_active_minus1 > 0 || mb_field_decoding_flag) && mb_type != P_8x8ref0 && sub_mb_type[mbPartIdx] != B_Direct_8x8 && SubMbPredMode(sub_mb_type[mbPartIdx]) != Pred_L1)		
ref_idx_l0[mbPartIdx]	2	te(v) | ae(v)
for(mbPartIdx = 0; mbPartIdx < 4; mbPartIdx++)		
if(!sub_mb_vsp_flag[mbPartIdx] && (num_ref_idx_l1_active_minus1 > 0 || mb_field_decoding_flag) && sub_mb_type[mbPartIdx] != B_Direct_8x8 && SubMbPredMode(sub_mb_type[mbPartIdx]) != Pred_L0)		
ref_idx_l1[mbPartIdx]	2	te(v) | ae(v)
for(mbPartIdx = 0; mbPartIdx < 4; mbPartIdx++)		
if(!sub_mb_vsp_flag[mbPartIdx] &&(sub_mb_type[mbPartIdx] != B_Direct_8x8 && SubMbPredMode(sub_mb_type[mbPartIdx]) != Pred_L1)		
for(subMbPartIdx = 0; subMbPartIdx < NumSubMbPart(sub_mb_type[mbPartIdx]); subMbPartIdx++)		
for(compIdx = 0; compIdx < 2; compIdx++)		
mvd_l0[mbPartIdx][subMbPartIdx][compIdx]	2	se(v) | ae(v)
for(mbPartIdx = 0; mbPartIdx < 4; mbPartIdx++)		
if(!sub_mb_vsp_flag[mbPartIdx] &&sub_mb_type[mbPartIdx] != B_Direct_8x8 && SubMbPredMode(sub_mb_type[mbPartIdx]) != Pred_L0)		
for(subMbPartIdx = 0; subMbPartIdx < NumSubMbPart(sub_mb_type[mbPartIdx]); subMbPartIdx++)		
for(compIdx = 0; compIdx < 2; compIdx++)		
mvd_l1[mbPartIdx][subMbPartIdx][compIdx]	2	se(v) | ae(v)
}		

[0093] No exemplo da Tabela 3 acima, o sub_mb_vsp_flag[mbPartIdx] igual a 1 indica que a partição de MB atual (8x8) é prevista a partir de uma Figura de VSP. Esse sinalizador igual a 0 indica que a

partição de MB inteira não é predicada a partir de uma figura de VSP. Quando não estiver presente, esse sinalizador é inferido como igual a 0.

[0094] Para simplificar a síntese de visualização, em alguns exemplos, a previsão de síntese de visualização pode ser unificada com o vetor de movimento previsão. Em tais exemplos, a previsão de síntese de visualização pode ser simplificara para realizar a distorção para trás apenas e não para realizar quaisquer técnicas de síntese de visualização adicional tal como preenchimento de orifício. Além disso, um valor de disparidade ou um valor de profundidade é aplicado ao bloco 4x4 inteiro. Consequentemente, é possível simplificar a previsão de síntese de visualização para usar as técnicas de compensação de movimento tradicionais, de modo que o vetor de movimento seja derivado da profundidade ou da disparidade e são associadas com cada bloco 4x4 de um MB, uma partição de MB ou uma subpartição de MB. Essa síntese de visualização processo sintetizada pode ser chamada de modo de VSP unificado.

[0095] Conforme observado acima, a sinalização da previsão de síntese de visualização pode ter como base um índice de referência tradicional, que é transparente ao projeto H.264/AVC nos termos de Sinalização de nível de MB. Por exemplo, para cada MB, um codificador de vídeo pode determinar se o uso da síntese de visualização com base em um índice de referência que identifica a figura de síntese de visualização em uma lista de referência de figura e se a síntese de visualização índice de referência pode ser sinalizada explicitamente em um cabeçalho de slice. Esse projeto pode suportar a previsão bidirecional, onde uma direção é prevista com uma figura de síntese de visualização e a outra direção é

prevista a partir de uma figura temporal normal ou figura de referência de intervisualização. Esse projeto também suporte múltiplas figurações de síntese de visualização a serem adaptadas no nível de slice, bem como no nível de macrobloco.

[0096] Não é possível desligar o índice de referência atual com base no mecanismo de sinalização de 3D-ATM. Por exemplo, não é possível desligar o índice de referência atual com base no mecanismo de sinalização de 3D-ATM sem interromper os objetivos de compatibilidade para trás da extensão de 3D-ATM. Consequentemente, um módulo de hardware novo que suporte o bloco com base na previsão de síntese de visualização pode ser necessário para o codificador de vídeos. Além disso, quando uma sequência de figurações tem uma figura de síntese de visualização associada (a seq_view_synthesis_flag é igual a 1), um codificador de vídeo deve gerar os valores de índice de referência adicionais para as figurações de síntese de visualização, de modo que os valores de índice de referência estão disponíveis para serem atribuídos em exemplos em que uma figura de síntese de visualização é gerada e usada para propósitos de previsão. Consequentemente, o codificador de vídeo gera valores de índice de referência adicionais, mesmo quando uma figura de síntese de visualização não é gerada atualmente durante a codificação, impactar a eficácia computacional e/ou de codificação.

[0097] Os aspectos dessa revelação incluem técnicas para indicar se a previsão de síntese de visualização deve ser aplicada quando houver codificação de um bloco de dados de vídeo. Em um exemplo, um codificador de vídeo, tal como o codificador de vídeo 20 e/ou o decodificador de vídeo 30, podem determinar se um índice de

referência para um bloco atual que corresponde a uma figura de referência de intervisualização e quando o índice de referência para o bloco atual corresponde à figura de referência de intervisualização, determinar a indicação de dados de um modo de VSP do bloco atual, onde o Modo de VSP para o índice de referência indica se o bloco atual é previsto com o uso da previsão de síntese de visualização da figura de intervisualização (por exemplo, a partir de pelo menos uma porção de uma figura sintetizada da figura de intervisualização). Em exemplos em o codificador de vídeo é configurado como um decodificador de vídeo 30, o decodificador de vídeo 30 pode determinar a indicação de dados do modo de VSP pela obtenção de dados para um fluxo de bits codificado.

[0098] Consequentemente, no exemplo acima, um índice de referência de uma figura de VSP não é especificamente indicado em um cabeçalho de slice ou adicionado durante a construção de lista de referência de figura. Ao invés disso, o codificador de vídeo 20 pode codificar (e o decodificador de vídeo 30 pode obter, a partir de um fluxo de bits codificado) um ou mais elementos de sintaxe em nível de MB ou de partição de MB para cada índice de referência que corresponde a uma figura de referência interprevista. Conforme descrito no presente documento, uma figura de referência de interprevisão é geralmente associada a um modo de interprevisão. Isto é, uma figura de interprevisão pode incluir uma figura de referência de intravisualização ou uma figura de referência de intervisualização.

[0099] Em alguns exemplos, um ou mais elementos de sintaxe podem ser um Sinalizador de VSP. Por exemplo, o sinalizador pode indicar se o VSP é aplicado ao MB ou à partição de MB para um índice de referência

particular. Em outras palavras, por exemplo, o sinalizador pode indicar usar uma figura de referência de intervisualização para os propósitos da previsão ou usar uma Figura de VSP para os propósitos da previsão.

[0100] Em um exemplo, de acordo com os aspectos dessa revelação, o codificador de vídeo 20 pode apenas codificar um Sinalizador de VSP quando um índice de referência para codificação de previsibilidade de um bloco corresponder a uma figura de referência de intervisualização. Em outro exemplo, o codificador de vídeo 20 pode codificar o sinalizador de VSP para qualquer modo interprevisto. Em outro exemplo, adicional ou alternativamente, o codificador de vídeo 20 pode codificar o sinalizador apenas para um índice de referência que corresponde à RefPlicList0. Em ainda outro exemplo, o codificador de vídeo 20 pode codificar o sinalizador para um índice de referência que corresponde tanto à RefPlicList0 quanto à RefPicList1 (se estiver presente).

[0101] De acordo com outro aspecto dessa revelação, quando a previsão de síntese de visualização é permitida, o codificador de vídeo 20 pode codificar (e o decodificador de vídeo 30 pode obter, a partir de um fluxo de bits codificado) um ou mais elementos de sintaxe em um conjunto de parâmetro, tal como uma SPS ou uma extensão de SPS que indica se a referência com base na sinalização é usada para o VSP (por exemplo, em que um índice de figura de referência é gerado para uma a figura de síntese de visualização de referência e o VSP é realizado com base no índice de referência), por exemplo, um primeiro modo de VSP ou se GVSP é usado para o VSP (por exemplo, em que o VSP é realizado com base na sinalização de macrobloco descrita acima, por exemplo, em relação às Tabelas 1 a 3 acima), por exemplo, um segundo modo de VSP. Em alguns exemplos, um ou

mais elementos de sintaxe podem incluir um sinalizador em uma Extensão de SPS.

[0102] De acordo com outros aspectos dessa revelação, quando o índice de referência com base na sinalização não for usado, um processo de modificação de lista de referência de figura (RPLM) não deve conter qualquer comando de RPLM que corresponda a uma figura de VSP. Por exemplo, a construção de lista de referência de figura pode incluir as seguintes etapas: (1) o processo de inicialização de lista de referência de figura para figurações de referência temporais (intravisualização) conforme especificado no AVC é aplicado (2) a intervisualização e/ou as figurações de referência sintetizadas são anexas ao fim da lista, e (3) um processo de RPLM processo é aplicado para reorganizar as figurações de referência na lista de referência de figura. O codificador de vídeo 20 pode emitir um ou mais comandos de RPLM que instruem o decodificador de vídeo 30 a aplicar o processo de RPLM. O decodificador de vídeo 30 pode aplicar o processo de RPLM em resposta aos comandos de RPLM.

[0103] De acordo com os aspectos dessa revelação, quando o índice de referência com base na sinalização não for usado, o codificador de vídeo 20 e o decodificador de vídeo 30 não poderão emitir um comando de RPLM que corresponda a uma Figura de VSP. Isto é, o codificador de vídeo 20 e o decodificador de vídeo 30 poderão ser impedidos de usar um RPLM para alterar a posição de uma Figura de VSP, devido ao fato de que nenhuma Figura de VSP foi gerada.

[0104] De acordo com outros aspectos dessa revelação, quando o GVSP for usado, o codificador de vídeo 20 e o decodificador de vídeo 30 poderão sintetizar uma previsão de figura de síntese de visualização a partir

apenas da primeira figura de referência de intervisualização para RefPicList0. Nesse exemplo, nenhuma sinalização da previsão de figura de síntese de visualização é necessária e é determinada automaticamente (por exemplo, inferida) como o Modo de GVSP que emite um vetor de movimento de disparidade que aponta à primeira figura de referência de intervisualização, por exemplo, RefPicList0. Alternativamente, em outro exemplo, o codificador de vídeo 20 e o decodificador de vídeo 30 podem realizar previsão de síntese de visualização com base na figura de intervisualização em RefPicList1. Alternativamente, em outro exemplo, o codificador de vídeo 20 e o decodificador de vídeo 30 podem realizar a previsão de síntese de visualização com base nas figurações de intervisualização, se presente, tanto em RefPicList0 quanto em RefPicList1.

[0105] A Figura 2 é um diagrama em bloco que ilustra um exemplo de codificador de vídeo 20 que pode implantar as técnicas descritas nessa revelação para síntese de visualização. O codificador de vídeo 20 pode realizar a intracodificação e a intercodificação de blocos de vídeo dentro de fatias de vídeo. A intracodificação depende da previsão espacial para reduzir ou remover a redundância espacial no vídeo dentro de uma figura dada. A intercodificação depende da previsão temporal para reduzir ou remover a redundância temporal no vídeo dentro das figurações adjacentes ou figurações de uma sequência de vídeo. O intramodo (Modo 1) pode se referir a quaisquer diversos espaços com base nos modos de compreensão. Os intermodos, tal como a previsão unidirecional (modo P) ou biprevisão (modo B), pode se referir a qualquer um dos modos de compreensão diversos com base temporal.

[0106] Conforme observa-se acima, o

codificador de vídeo 20 pode ser adaptado para realizar a codificação de vídeo de multivisualização. Além dos mapas de textura de codificação (isto é, os valores de luma e de croma) para cada visualização, o codificador de vídeo 20 pode adicionalmente codificar um mapa de profundidade para cada visualização.

[0107] Conforme mostrado na Figura 2, o codificador de vídeo 20 recebe um bloco atual de vídeo dentro de um quadro de vídeo a ser codificado. No exemplo da Figura 2, o codificador de vídeo 20 inclui a unidade de seleção de modo 40, a memória de figura de referência 64, o somador 50, a unidade de processamento de transformada 52, a unidade de quantização 54 e a unidade de codificação de entropia 56. A unidade de seleção de modo 40, por sua vez, inclui a unidade de compensação de movimento 44, a unidade de estimativa de movimento 42, a unidade de intraprevisão 46, e a unidade de partição 48. Para a reconstrução de bloco de vídeo, o codificador de vídeo 20 inclui também uma unidade de quantização inversa 58, uma unidade de transformada inversa 60 e o somador 62.

[0108] Um filtro de desbloqueio (não mostrado na Figura 2) pode ser também incluído para filtrar os primeiros limites de bloco para remover artefatos de blocagem do vídeo reconstruído. Se desejado, o filtro de desbloqueio seria um filtro típico de saída de somador 62. Filtros adicionais (em ciclo ou pós-ciclo) podem também ser usados, além do filtro de desbloqueio. Tais filtros não são mostrados por motivos de brevidade, mas se desejado, podem filtrar a saída do somador 50 (como um filtro em ciclo).

[0109] Durante o processo de codificação, o codificador de vídeo 20 recebe um quadro de vídeo ou a fatia a ser codificada. O quadro ou a fatia podem ser divididos em múltiplos blocos de vídeo. A unidade de

estimativa de movimento 42 e a unidade de compensação de movimento 44 realizam a codificação de interprevisão do bloco de vídeo recebido em relação a um ou mais blocos em um ou mais quadros de referência para fornecer previsão temporal. A unidade de intraprevisão 46 pode realizar alternativamente a codificação de intraprevisão do bloco de vídeo recebido em relação a um ou mais blocos vizinhos no mesmo quadro ou fatia como o bloco a ser codificado para fornecer previsão espacial. O codificador de vídeo 20 pode realizar múltiplos passos de codificação, por exemplo, para selecionar um modo de codificação adequado para cada um dos blocos de dados de vídeo.

[0110] Além disso, a unidade de partição 48 pode partir os blocos de dados de vídeo em sub-blocos, com base na avaliação dos esquemas de partição anteriores em passes de codificação anteriores. Por exemplo, a unidade de partição 48 pode partir inicialmente um quadro ou uma fatia em blocos (por exemplo, macroblocos ou LCUs) e partir cada um dos sub-blocos de blocos com base na análise de taxa de distorção (por exemplo, aprimoramento de taxa de distorção).

[0111] A unidade de seleção de modo 40 pode selecionar um dos modos de codificação, tal como previsão de intratemporal, intertemporal ou intervisualização, por exemplo, com base em resultados de erros e fornecer um bloco previsto ao somador 50 para gerar dados de bloco residual e ao somador 62 para reconstruir o bloco codificado para o uso como um quadro de referência. A unidade de seleção de modo 40 fornece também elementos de sintaxe, tal como vetores de movimento, indicadores de intramodo, informações de partição e outras informações de sintaxe para a unidade de codificação de entropia 56.

[0112] Em alguns exemplos, a unidade de

seleção de modo 40 pode ser configurada para selecionar a previsão de intervisualização, por exemplo, em relação a uma visualização codificada anteriormente ou a uma visualização sintetizada para a previsão de síntese de visualização. Conforme discutido em maiores detalhes abaixo, a unidade de síntese de visualização 66 pode ser configurada para sintetizar uma visualização (isto é, sintetizar as figurações, que inclui valores de pixel de textura e/de profundidade) para a previsão de síntese de visualização. Por exemplo, a unidade de síntese de visualização 66 pode ser configurada para realizar técnicas substancialmente similares à renderização com base em profundidade de imagem (DIBR).

[0113] Deve-se observar que onde a previsão de síntese de visualização for permitida, a unidade de seleção de modo 40 poderá ser ainda selecionada dentre outros modos de codificação disponíveis, por exemplo, a intraprevisão, interprevisão temporal ou previsão de intervisualização em uma figura codificada anteriormente. Desse modo, o codificador de vídeo 20 pode fornecer uma indicação de qual modo de codificação é selecionado para um bloco de dados de vídeo, separado da indicação de informações se a previsão de síntese de visualização está disponível.

[0114] A unidade de estimativa de movimento 42 e a unidade de compensação de movimento 44 (bem como uma ou mais outras unidades do codificador de vídeo 20) podem ser altamente integradas, porém são ilustradas separadamente para fins conceituais. A estimativa de movimento, pré-formada pela unidade de estimativa de movimento 42, é o processo de geração de vetores de movimento, que estimam o movimento para os blocos de vídeo. Um vetor de movimento, por exemplo, pode indicar o

deslocamento de um bloco de vídeo dentro de um quadro de vídeo ou figura atual em relação a um bloco de previsão dentro de um quadro de referência (ou outra unidade de codificação) em relação ao bloco atual que é codificado dentro da armação atual (ou outra unidade de codificação).

[0115] Um bloco de previsão é um bloco que é encontrado quase nivelado com o bloco a ser codificado, em termos de diferença de pixel, que pode ser determinada pela soma da diferença absoluta (SAD), a soma das diferenças do quadro (SSD) ou métricas diferente. Os vetores de movimento podem incluir vetores de movimento temporais, que descrevem o movimento de um bloco em relação a um bloco codificado anteriormente da mesma visualização em uma figura distinta temporalmente e vetores de movimento de disparidade, que descrevem a disparidade entre os blocos similares em visualizações diferentes (que tem perspectivas de câmera horizontais diferentes), porém que podem ter a mesma posição temporal. Em alguns casos, os vetores de movimento podem descrever o movimento em relação a uma figura distinta temporariamente que é também uma visualização diferente.

[0116] Em alguns exemplos, o codificador de vídeo 20 pode calcular os valores para posições de pixel sub-integral das figurações de referência armazenadas na memória de figura de referência 64. Por exemplo, o codificador de vídeo 20 pode interpolar valores de posições de pixel de um quarto, posições de pixel de um oitavo ou outras posições de pixel fracionadas de figura de referência. Portanto, a unidade de estimativa de movimento 42 pode realizar uma pesquisa de movimento em relação às posições de pixel completas e posições de pixel fracionárias e saídas de vetor de movimento com precisão de pixel fracionária.

[0117] Quando se pré-forma a previsão de intervisualização, a unidade de estimativa de movimento 42 pode calcular vetores de movimento de disparidade em relação às figurações codificadas anteriores ou um visualização diferente ou figurações sintetizadas para previsão de síntese de visualização. Desse modo, a unidade de estimativa de movimento 42 pode também ser referida como uma unidade de estimativa de movimento/disparidade.

[0118] Em geral, as figurações sintetizadas podem ser armazenadas na memória de figura de referência 64 pela unidade de síntese de visualização 66 e, desse modo, a unidade de estimativa de movimento 42 e a unidade de compensação de movimento44 não precisam ser configuradas para determinar se uma figura de referência é uma figura codificada anteriormente de uma visualização diferente ou uma figura sintetizada de um processo de síntese de visualização. O processo para pesquisa por um vetor de movimento de disparidade pode ser restringida para pesquisar horizontalmente, ao invés de tanto horizontal quanto verticalmente, e devido ao fato das figurações de visualizações diferentes na mesma posição temporal incluírem tipicamente apenas diferenças horizontais e não diferenças verticais devido ao fato das figurações de uma cena serem capturadas ou geradas a partir de perspectivas de câmeras horizontal distintas.

[0119] A unidade de estimativa de movimento 42 calcula um vetor de movimento (por exemplo, um vetor de movimento temporal ou um vetor de movimento de disparidade) para um bloco de vídeo em uma fatia intercodificada, comparando-se a posição do bloco à posição de um bloco de previsão de uma figura de referência. A figura de referência pode ser selecionada a partir de uma primeira lista de figura de referência (Lista 0) ou uma segunda

lista de figura de referência (Lista 1), cada uma das quais identifica uma ou mais figurações de referência armazenadas na memória de figura de referência 64. A unidade de estimativa de movimento 42 envia o vetor de movimento calculado para a unidade de codificação de entropia 56 e para a unidade de compensação de movimento 44.

[0120] A compensação de movimento, realizada pela unidade de compensação de movimento 44, pode envolver buscar ou gerar o bloco de previsão com base no vetor de movimento determinado pela unidade de estimativa de movimento 42, o qual novamente pode ser um vetor de movimento temporal ou um vetor de movimento de disparidade. Novamente, a unidade de estimativa de movimento 42 e a unidade de compensação de movimento 44 podem ser funcionalmente integradas em alguns exemplos. Mediante o recebimento do vetor de movimento para o bloco de vídeo atual, a unidade de compensação de movimento 44 pode localizar o bloco de previsão para o qual o vetor de movimento aponta em uma das listas de figura de referência. O somador 50 forma um bloco de vídeo residual, subtraindo-se valores de pixel do bloco de previsão dos valores de pixel do bloco de vídeo atual que é codificado, formando valores de diferença de pixel, conforme discutido abaixo.

[0121] Em geral, a unidade de estimativa de movimento 42 realiza estimativa de movimento em relação aos componentes de luma e a unidade de compensação de movimento 44 usa vetores de movimento calculados com base nos componentes de luma para ambos os componentes de croma e os componentes de luma. A unidade de estimativa de movimento 42 e a unidade de compensação de movimento 44 podem ser configuradas para reusar os vetores de movimento dos componentes de luma para codificar mapas de profundidade ou para calcular independentemente os vetores de movimento

para mapas de profundidade. Assim, em alguns casos, a estimativa de movimento 42 e a unidade de compensação de movimento 44 podem ser configuradas para prever mapas de profundidade de maneira similar a componentes de croma ou de maneira similar a componentes de luma. A unidade de seleção de modo 40 também pode gerar elementos de sintaxe associados aos blocos de vídeo e à fatia de vídeo para uso pelo decodificador de vídeo 30 na decodificação dos blocos de vídeo da fatia de vídeo.

[0122] A unidade de intraprevisão 46 pode intraprever um bloco atual, como uma alternativa à interprevisão (por exemplo, interprevisão temporal e/ou previsão de intervisualização) realizada pela unidade de estimativa de movimento 42 e a unidade de compensação de movimento 44, conforme descrito acima. Em particular, a unidade de intraprevisão 46 pode determinar um modo de intraprevisão para codificar um bloco atual. Em alguns exemplos, a unidade de intraprevisão 46 pode codificar um bloco atual com o uso de vários modos de intraprevisão, por exemplo, durante passos de codificação separadas, e a unidade de intraprevisão 46 (ou unidade de seleção de modo 40, em alguns exemplos) pode selecionar um modo de intraprevisão apropriado para usar a partir dos modos testados.

[0123] Por exemplo, a unidade de intraprevisão 46 pode calcular valores de taxa de distorção com o uso de uma análise de taxa de distorção para os vários modos de intraprevisão testados e selecionar o modo de intraprevisão que tem as melhores características de taxa de distorção entre os modos testados. A análise de taxa de distorção, em geral, determina em quantidade de distorção (ou erro) entre um bloco codificado e um original, um bloco não codificado que foi codificado para

produzir o bloco codificado, assim como uma taxa de bits (isto é, um número de bits) usada para produzir o bloco codificado. A unidade de intraprevisão 46 pode calcular razões a partir das distorções e taxas para os vários blocos codificados para determinar qual modo de intraprevisão exibe o melhor valor de taxa de distorção para o bloco.

[0124] Após selecionar um modo de intraprevisão para um bloco, a unidade de intraprevisão 46 pode fornecer informações indicativas do modo de intraprevisão selecionado para o bloco à unidade de codificação de entropia 56. A unidade de codificação de entropia 56 pode codificar as informações que indicam o modo de intraprevisão selecionado. O codificador de vídeo 20 pode incluir nos dados de configura de fluxo de bits transmitido, que pode incluir uma pluralidade de tabelas de índice de modo de intraprevisão e uma pluralidade de tabelas de índice de modo de intraprevisão modificadas (também denominadas tabelas de mapeamento de palavra código), definições de contextos de codificação para vários blocos e indicações de um modo de intraprevisão mais provável, uma tabela de índice de modo de intraprevisão e uma tabela de índice de modo de intraprevisão modificada para uso em cada um dos contextos.

[0125] Além disso, conforme observado acima, a unidade de seleção de modo 40 pode determinar a utilização ou não de uma previsão de síntese de visualização para codificar um bloco particular de uma visualização de não base. A unidade de seleção de modo 40 pode receber dados de configura que indicam se a previsão de síntese de visualização está habilitada para um fluxo de bits ou para um ponto de operação particular de um fluxo de bits de multivisualizações. Por exemplo, um usuário pode

fornecer dados de configura que indicam que o fluxo de bits deve ser compatível com dispositivos que não são equipados para realizar previsão de síntese de visualização, caso em que a unidade de seleção de modo 40 pode desabilitar a previsão de síntese de visualização para o fluxo de bits total. Alternativamente, os dados de configura podem indicar um subconjunto de visualizações que formam um ponto de operação para o qual a previsão de síntese de visualização é desabilitada, caso em que a unidade de seleção de modo 40 pode desabilitar a previsão de síntese de visualização para cada uma das visualizações no subconjunto, mas testar a previsão de síntese de visualização para outras visualizações do fluxo de bits não incluídas no subconjunto.

[0126] Em geral, supondo-se que a previsão de síntese de visualização é permitida por uma visualização de não base particular, a unidade de seleção de modo 40 pode testar a previsão de síntese de visualização para blocos de figura da visualização de não base para determinar se a previsão de síntese de visualização rende melhor desempenho que outros modos de codificação, por exemplo, interprevisão temporal, intraprevisão e/ou previsão de intervisualização em relação a uma visualização codificada anteriormente (não sintetizada). A unidade de seleção de modo 40 pode testar o desempenho de vários modos de codificação com o uso de otimização de taxa de distorção (RDO) e/ou com o uso de taxas Bjontegaard-Delta (taxas BD). Em geral, a otimização de taxa de distorção é uma medida do número de bits necessário para codificar um conjunto de dados de vídeo com o uso de um modo particular ou de uma combinação de modos em relação à quantidade de distorção introduzida pelo modo ou pela combinação de modos. A unidade de seleção de modo 40 pode selecionar o modo ou a combinação de modos que

rendem o melhor desempenho, conforme indicado por essas medidas.

[0127] Além disso, a unidade de seleção de modo 40 pode determinar um número de blocos em uma fatia, figura, sequência ou outra unidade codificada (por exemplo, um recorte de imagem ou frente de onda) para a qual a previsão de síntese de visualização rende o melhor desempenho e determina se esse número de blocos é suficientemente alto para justificar a despesa de processamento de usar a previsão de síntese de visualização. Por exemplo, sintetizar uma visualização por um decodificador de vídeo pode requerer uma quantidade bastante grande de recursos de processamento e, portanto, a unidade de seleção de modo 40 pode selecionar um modo de codificação diferente para um bloco particular além da previsão de síntese de visualização, se não houver um número suficientemente grande de blocos na unidade codificada para a qual a previsão de síntese de visualização deve ser usada.

[0128] A unidade de síntese de visualização 66 representa uma unidade que sintetiza visualizações para previsão de intervisualização. Quando a previsão de síntese de visualização é habilitada, a unidade de síntese de visualização 66 pode sintetizar uma figura para uma visualização sintetizada com o uso de informações de textura e profundidade armazenadas na memória de figura de referência 64. A unidade de síntese de visualização 66 pode usar uma ou mais visualizações de referência para sintetizar outra visualização.

[0129] Em geral, para sintetizar uma figura, a unidade de síntese de visualização 66 pode usar informações de textura e profundidade de uma ou mais visualizações codificadas anteriormente. A unidade de

síntese de visualização 66 pode calcular a disparidade horizontal para as informações de textura com base nas informações de profundidade correspondentes (por exemplo, substancialmente co-situadas) de uma figura de textura e um mapa de profundidade correspondente. Em geral, objetos que devem aparecer na profundidade da exibição (por exemplo, no plano de convergência) podem ter uma disparidade de zero. Isto é, pixels representativos desse objeto podem ter substancialmente a mesma posição horizontal na visualização sintetizada, assim como na visualização de referência. Para objetos a serem exibidos na frente da tela, uma disparidade positiva pode ser designada, de modo que em uma figura de "olho esquerdo", pixels para o objeto são posicionados à direita dos pixels correspondentes na figura de "olho direito". Alternativamente, para objetos a serem exibidos atrás da tela, uma disparidade negativa pode ser designada, de modo que em uma figura de “olho esquerdo”, pixels para o objeto são posicionados à esquerda dos pixels correspondentes na figura de “olho direito”. A alternância horizontal relativa pode ser determinada como uma função da quantidade de profundidade a ser alcançada, do deslocamento horizontal relativo, da distância ao plano de convergência, das distâncias de mundo real, dos parâmetros de câmera e similares.

[0130] Desse modo, a unidade de síntese de visualização 66 pode sintetizar uma figura de uma visualização a ser usada como uma figura de referência. A unidade de síntese de visualização 66 pode usar processos similares para sintetizar um ou outro ou ambos dentre uma imagem de textura e/ou um mapa de profundidade. Assim, o codificador de vídeo 20 pode codificar uma imagem de textura e/ou um mapa de profundidade em relação a um componente de visualização sintetizada, por exemplo, uma

imagem de textura sintetizada ou um mapa de profundidade sintetizado. A unidade de síntese de visualização 66 pode armazenar dados para o componente de visualização sintetizada na memória de figura de referência 64. Desse modo, a unidade de seleção de modo 40, a unidade de estimativa de movimento 42 e a unidade de compensação de movimento 44 podem tratar o componente de visualização sintetizada como se o componente de visualização sintetizada fosse uma figura de referência convencional.

[0131] De acordo com aspectos dessa revelação, a unidade de seleção de modo 40 pode gerar um ou mais elementos de sintaxe que indicam se a VSP para codificar um bloco particular deve ser aplicado. Em alguns exemplos, a unidade de seleção de modo 40 pode gerar apenas os elementos de sintaxe para blocos que são interprevistos, tais como previstos por intervisualização. Assim, nesse exemplo, quando um índice de figura de referência para o bloco atual corresponde a uma figura de referência intervisualizada, a unidade de seleção de modo 40 pode enviar um sinalizador à unidade de codificação de entropia 56 para indicar se a VSP é usada para codificar o bloco atual.

[0132] No exemplo acima, um índice de referência de uma figura de VSP não é indicado especificamente em um cabeçalho de slice ou adicionado durante a construção da lista de figura de referência. Ao invés disso, o codificador de vídeo 20 pode codificar um sinalizador no MB ou nível de partição de MB para cada índice de referência que corresponde a uma figura de referência de intervisualização. Tal sinalização pode fornecer uma abordagem mais flexível à sinalização de VSP, em que a figura sintetizada não está vinculada a um índice de figura de referência particular.

[0133] Em outro exemplo, a unidade de seleção de modo 40 pode gerar o sinalizador de VSP para qualquer modo interprevisto. Em outro exemplo, adicional ou alternativamente, a unidade de seleção de modo 40 pode gerar o sinalizador apenas para um índice de referência correspondente a RefPlicList0. Em ainda outro exemplo, a unidade de seleção de modo 40 pode gerar o sinalizador para um índice de referência correspondente a ambos RefPlicList0 e RefPicList1 (se presente).

[0134] De acordo com outros aspectos dessa revelação, quando a previsão de síntese de visualização for habilitada, a unidade de seleção de modo 40 pode gerar um ou mais elementos de sintaxe para um conjunto de parâmetro, tais como um SPS ou extensão de SPS que indica se a referência com base na sinalização é usada para a VSP (por exemplo, em que um índice de figura de referência é gerado para uma figura de referência de síntese de visualização e a VSP é realizada com base no índice de referência) ou se a GVSP é usada para a VSP (por exemplo, em que a VSP é realizada com base na sinalização de macrobloco descrita acima, por exemplo, em relação às Tabelas 1 a 3 acima). Em alguns exemplos, o um ou mais elementos de sintaxe podem incluir um sinalizador em uma extensão de SPS.

[0135] De acordo com outros aspectos dessa revelação, quando o índice de referência com base na sinalização não é usado, o codificador de vídeo 20 pode não emitir nenhum comando RPLM que corresponde a uma figura de VSP.

[0136] De acordo com outros aspectos dessa revelação, quando a GVSP é usada, a unidade de síntese de visualização 66 pode sintetizar uma figura de previsão de síntese de visualização a partir apenas da primeira figura de referência de intervisualização para RefPicList0. Nesse

exemplo, nenhuma sinalização da figura de previsão de síntese de visualização é necessária e é automaticamente determinado (por exemplo, inferido) que o modo de GVSP emite um vetor de movimento de disparidade que aponta para a primeira figura de referência intervisualizada, por exemplo, RefPicList0. Alternativamente, em outro exemplo, a unidade de síntese de visualização 66 pode realizar a previsão de síntese de visualização com base em uma figura de intervisualização em RefPicList1. Alternativamente, em outro exemplo, a unidade de síntese de visualização 66 pode realizar a previsão de síntese de visualização com base na figura de intervisualização, se estiver presente, em ambos RefPicList0 e RefPicList1.

[0137] O codificador de vídeo 20 forma um bloco de vídeo residual, subtraindo-se os dados de previsão da unidade de seleção de modo 40 do bloco de vídeo original que é codificado. O somador 50 representa o componente ou os componentes que realizam essa operação de subtração. A unidade de processamento de transformada 52 aplica uma transformada, tal como uma transformada discreta de cosseno (DCT) ou uma transformada conceitualmente similar, ao bloco residual, produzindo um bloco de vídeo que compreende valores de coeficiente de transformada residual. A unidade de processamento de transformada 52 pode realizar outras transformadas que são conceitualmente similares à DCT. Transformadas Wavelet, transformadas inteiras, transformadas de sub-banda ou outros tipos de transformadas também podem ser usadas.

[0138] Em todo caso, a unidade de processamento de transformada 52 aplica a transformada ao bloco residual, produzindo um bloco de coeficientes de transformada residual. A transformada pode converter as informações residuais a partir de um domínio de valor de

pixel para um domínio de transformada, tal como um domínio de frequência. A unidade de processamento de transformada 52 pode enviar os coeficientes de transformada resultantes para a unidade de quantização 54. A unidade de quantização 54 quantiza os coeficientes de transformada para reduzir ainda mais a taxa de bits. O processo de quantização pode reduzir a profundidade de bit associada a alguns ou a todos os coeficientes. O grau de quantização pode ser modificado, ajustando-se um parâmetro de quantização. Em alguns exemplos, a unidade de quantização 54 pode, então, realizar uma varredura da matriz que inclui os coeficientes de transformada quantizados. Alternativamente, a unidade de codificação de entropia 56 pode realizar a varredura.

[0139] Após a quantização, a unidade de codificação de entropia 56 codifica por entropia os coeficientes de transformada quantizados. Por exemplo, a unidade de codificação de entropia 56 pode realizar codificação por comprimento variável adaptativo ao contexto (CAVLC), codificação por aritmética binária adaptativa ao contexto (CABAC), codificação por aritmética binária baseada em sintaxe adaptativa ao contexto (SBAC), codificação de entropia por partição de probabilidade de intervalo (PIPE) ou outra técnica de codificação de entropia. No caso da codificação de entropia baseada em contexto, o contexto pode ser baseado nos blocos próximos. Após a codificação de entropia pela unidade de codificação de entropia 56, o fluxo de bits codificado pode ser transmitido a outro dispositivo (por exemplo, um decodificador de vídeo 30) ou alcançado para transmissão ou recuperação posterior.

[0140] A unidade de quantização inversa 58 e a unidade de transformada inversa 60 aplicam quantização inversa e transformação inversa, respectivamente, para

reconstruir o bloco residual no domínio de pixel, por exemplo, para uso posterior como um bloco de referência. A unidade de compensação de movimento 44 pode calcular um bloco de referência, adicionando-se o bloco residual a um bloco de previsão de um dos quadros de memória de figura de referência 64. A unidade de compensação de movimento 44 também pode aplicar um ou mais filtros de interpolação ao bloco residual reconstruído para calcular valores de pixel subinteiros para uso na estimativa de movimento.

[0141] O somador 62 adiciona o bloco residual reconstruído ao bloco de previsão compensado por movimento produzido pela unidade de compensação de movimento 44 para produzir um bloco de vídeo reconstruído para armazenamento na memória de figura de referência 64. O bloco de vídeo reconstruído pode ser usado pela unidade de estimativa de movimento 42 e pela unidade de compensação de movimento 44 como um bloco de referência para intercodificar um bloco em um quadro de vídeo subsequente.

[0142] De acordo com aspectos dessa revelação, o codificador de vídeo 20 pode realizar em método que inclui determinar de um índice de referência para um bloco atual corresponde a uma figura de referência intervisualizada e quando o índice de referência para o bloco atual corresponde à figura de referência intervisualizada, gerar dados que indicam um modo de previsão de síntese de visualização (VSP) do bloco atual, em que o modo de VSP para o índice de referência indica se o bloco atual é previsto com previsão de síntese de visualização a partir da figura de referência intervisualizada. O codificador de vídeo 20 também é configurado para codificar os dados que indicam o modo de VSP em um fluxo de bits.

[0143] A Figura 3 é um diagrama de blocos que

ilustra um exemplo de decodificador de vídeo 30 que pode implantar técnicas para codificar informações relacionadas à previsão de síntese de visualização. No exemplo da Figura 3, o decodificador de vídeo 30 inclui uma unidade de decodificação de entropia 70, uma unidade de compensação de movimento 72, uma unidade de intraprevisão 74, uma unidade de quantização inversa 76, uma unidade de transformação inversa 78, uma memória de figura de referência 82 e um somador 80. O decodificador de vídeo 30 pode, em alguns exemplos, realizar um passo de decodificação em geral recíproco ao passo de codificação descrito em relação ao codificador de vídeo 20 (Figura 2). A unidade de compensação de movimento 72 pode gerar dados de previsão com base em vetores de movimento recebidos da unidade de decodificação de entropia 70, enquanto a unidade de intraprevisão 74 pode gerar dados de previsão com base em indicadores de modo de intraprevisão recebidos da unidade de decodificação de entropia 70.

[0144] Durante o processo de decodificação, o decodificador de vídeo 30 recebe um fluxo de bits de vídeo codificado que representa blocos de vídeo de uma fatia de vídeo codificada e associada a elementos de sintaxe do codificador de vídeo 20. A unidade de decodificação de entropia 70 decodifica por entropia o fluxo de bits para gerar coeficientes quantizados, vetores de movimento ou indicadores de modo de intraprevisão e outros elementos de sintaxe. A unidade de decodificação de entropia 70 transfere os vetores de movimento e outros elementos de sintaxe à unidade de compensação de movimento 72. O decodificador de vídeo 30 pode receber os elementos de sintaxe no nível de fatia de vídeo e/ou no nível de bloco de vídeo.

[0145] Em alguns exemplos, o decodificador de

vídeo 30 pode receber informações de sintaxe que indicam se a previsão de síntese de visualização é habilitada para qualquer uma ou para todas dentre uma sequência de figura, uma figura individual, uma fatia, um recorte de vídeo ou uma frente de onda. A unidade de decodificação de entropia 70 pode decodificar as informações de sintaxe e enviar informações que indicam se a previsão de síntese de visualização é habilitada à unidade de síntese de visualização 84.

[0146] A unidade de síntese de visualização 84 pode ser configurada para operar de maneira substancialmente similar àquela da unidade de síntese de visualização 66 (Figura 2). Por exemplo, quando a previsão de síntese de visualização for habilitada, a unidade de síntese de visualização 84 pode usar informações de textura e profundidade da figura decodificada anteriormente, armazenada na memória de figura de referência 82, para gerar uma figura de referência com o uso de síntese de visualização, isto é, uma figura de referência de síntese de visualização. A unidade de síntese de visualização 84 pode armazenar a figura sintetizada na memória de figura de referência 82. Quando a previsão de síntese de visualização não for habilitada, a unidade de síntese de visualização 84 não precisa gerar uma figura, o que pode conservar recursos de processamento e/ou potência de bateria.

[0147] De acordo com aspectos dessa revelação, a unidade de decodificação de entropia 70 pode receber e decodificar um ou mais elementos de sintaxe que indicam se a VSP para codificar um bloco particular deve ser aplicada. Em alguns exemplos, os elementos de sintaxe podem ser gerados apenas para blocos que são interprevistos, tais como previstos por intervisualização. Assim, nesse exemplo, quando um índice de figura de

referência para o bloco atual corresponde a uma figura de referência de intervisualização, a unidade de síntese de visualização 84 pode receber um sinalizador que indica se a VSP é usada para codificar o bloco atual.

[0148] No exemplo acima, um índice de referência de uma figura de VSP não é indicado especificamente em um cabeçalho de slice ou adicionado durante a construção da lista de figura de referência. Ao invés disso, o decodificador de vídeo 30 decodifica um sinalizador no MB ou no nível de partição de MB para cada índice de referência que corresponde a uma figura de referência de intervisualização. Tal sinalização pode fornecer uma abordagem mais flexível à sinalização de VSP, em que a figura sintetizada não é vinculada a um índice de figura de referência particular.

[0149] Em outro exemplo, a unidade de decodificação de entropia 70 pode receber e decodificar o sinalizador de VSP para qualquer modo interprevisto. Em outro exemplo, adicional ou alternativamente, a unidade de decodificação de entropia 70 pode receber e decodificar o sinalizador apenas para um índice de referência que corresponde a RefPlicList0. Em ainda outro exemplo, a unidade de decodificação de entropia 70 pode receber e decodificar o sinalizador para um índice de referência que corresponde a ambos RefPlicList0 e RefPicList1 (se estiver presente).

[0150] De acordo com outros aspectos dessa revelação, quando a previsão de síntese de visualização é habilitada, o decodificador de vídeo 30 pode receber e decodificar um ou mais elementos de sintaxe em um conjunto de parâmetro, tal como uma SPS ou uma extensão de SPS que indica se a sinalização com base em referência é usada para a VSP (por exemplo, em, que um índice de figura de

referência é gerado para uma figura de referência de síntese de visualização e a VSP é realizada com base no índice de referência) ou se a GVSP é usada para a VSP (por exemplo, em que a VSP é realizada com base na sinalização de macrobloco descrita acima, por exemplo, em relação às Tabelas 1 a 3 acima). Em alguns exemplos, o um ou mais elementos de sintaxe pode incluir um sinalizador em uma extensão de SPS.

[0151] De acordo com outros aspectos dessa revelação, quando a sinalização com base em índice de referência não é usada, o decodificador de vídeo 30 pode não realizar nenhum comando de RPLM que corresponde a uma figura de VSP.

[0152] De acordo com outros aspectos dessa revelação, quando a GVSP é usada, a unidade de síntese de visualização 84 pode sintetizar uma figura de previsão de síntese de visualização a partir apenas da primeira figura de referência de intervisualização para RefPicList0. Nesse exemplo, nenhuma sinalização da figura de previsão de síntese de visualização é necessária e é automaticamente determinado (por exemplo, inferido) que o modo de GVSP emite um vetor de movimento de disparidade que aponta para a primeira figura de referência intervisualizada, por exemplo, RefPicList 0. Alternativamente, em outro exemplo, a unidade de síntese de visualização 84 pode realizar a previsão de síntese de visualização com base na figura de intervisualização em RefPicList1. Alternativamente, em outro exemplo, a unidade de síntese de visualização 84 pode realizar a previsão de síntese de visualização com base na figura de intervisualização, se estiver presente, em ambos RefPicList0 e RefPicList1.

[0153] Quando a fatia de vídeo é codificada como uma fatia intracodificada (I), a unidade de

intraprevisão 74 pode gerar dados de previsão para um bloco de vídeo da fatia de vídeo atual com base em um modo de intraprevisão sinalizado e dados dos blocos decodificados anteriormente do quadro ou figura atual. Quando o quadro de vídeo é codificado como uma fatia intercodificada (por exemplo, B, P ou GPB), a unidade de compensação de movimento 72 produz blocos de previsão para um bloco de vídeo da fatia de vídeo atual com base nos vetores de movimento e outros elementos de sintaxe recebidos da unidade de decodificação de entropia 70. Os blocos de previsão podem ser produzidos a partir de uma das figurações de referência dentro uma das listas de figura de referência. O decodificador de vídeo 30 pode construir as listas de quadro de referência, Lista 0 e Lista 1, com o uso de técnicas de construção padrão com base em figurações de referência armazenadas na memória de figura de referência 82.

[0154] A unidade de compensação de movimento 72 determina informações de previsão para um bloco de vídeo da fatia de vídeo atual, analisando-se os vetores de movimento e outros elementos de sintaxe, e os usos das informações de previsão para produzir os blocos de previsão para o bloco de vídeo atual que é decodificado. Por exemplo, a unidade de compensação de movimento 72 usa alguns dos elementos de sintaxe recebidos para determinar um modo de previsão (por exemplo, intraprevisão, interprevisão temporal ou previsão por intervisualização) usado para codificar os blocos de vídeo da fatia de vídeo, um tipo de fatia de interprevisão (por exemplo, fatia B, fatia P ou fatia GPB), informações de construção para um ou mais dentre as listas de figura de referência para a fatia, vetores de movimento (por exemplo, vetores de movimento de disparidade e/ou vetores de movimento temporais) para cada

bloco de vídeo intercodificado da fatia, estado de interprevisão para cada bloco de vídeo intercodificado da fatia e outras informações para decodificar os blocos de vídeo na fatia de vídeo atual.

[0155] A Unidade de compensação de movimento 72 também pode realizar interpolação com base em filtros de interpolação. A unidade de compensação de movimento 72 pode usar filtros de interpolação, conforme usados pelo codificador de vídeo 20 durante o codificação dos blocos de vídeo, para calcular valores interpolados para pixels subinteiros do bloco de referências. Nesse caso, a unidade de compensação de movimento 72 pode determinar os filtros de interpolação usados pelo codificador de vídeo 20 a partir dos elementos de sintaxe recebidos e usar os filtros de interpolação para produzir blocos de previsão.

[0156] Em alguns exemplos, a unidade de compensação de movimento 72 pode realizar previsão por intervisualização com o uso de vetores de movimento de disparidade. Tal previsão de intervisualização pode ser relativa a uma figura decodificada anteriormente de noutra visualização ou relativa a uma figura de referência gerada com o uso de síntese de visualização, supondo-se que a previsão de síntese de visualização seja habilitada. Consequentemente, a unidade de compensação de movimento 72 pode ser denominada de unidade de compensação de movimento/disparidade 72.

[0157] A unidade de quantização inversa 76 quantiza ao inverso, isto é, dequantiza, os coeficientes de transformada quantizados fornecidos no fluxo de bits e decodificados pela unidade de decodificação de entropia 70. O processo de quantização inverso pode incluir o uso de um parâmetro de quantização QPy calculado pelo decodificador de vídeo 30 para cada bloco de vídeo na fatia de vídeo para

determinar um grau de quantização e, de modo similar, um grau de quantização inversa que deve ser aplicado.

[0158] A unidade de transformada inversa 78 aplica uma transformada inversa, por exemplo, uma DCT inversa, uma transformada inteira inversa ou um processo de transformada inversa conceitualmente similar aos coeficientes de transformada, a fim de produzir blocos residuais no domínio de pixel.

[0159] Depois que a unidade de compensação de movimento 72 ou a unidade de intraprevisão 74 gera o bloco de previsão para o bloco de vídeo atual com base nos vetores de movimento e outros elementos de sintaxe, o decodificador de vídeo 30 forma um bloco de vídeo decodificado, somando-se os blocos residuais da unidade de transformada inversa 78 com o bloco de previsão correspondente. O somador 80 representa o componente ou os componentes que realizam essa operação de soma.

[0160] Se desejado, um filtro de desbloqueio também pode ser aplicado para filtrar os blocos, a fim de remover artefatos de blocagem. Outros filtros em ciclo (no ciclo de codificação ou após o ciclo de codificação) também podem ser usados para transições suaves de pixel ou, de outra forma, para aprimorar a qualidade do vídeo. Os blocos de vídeo decodificados em um determinado quadro ou figura são, então, armazenados na memória de figura de referência 82 que armazena as figurações de referência usadas para compensação de movimento subsequente. A memória de figura de referência 82 também armazena o vídeo decodificado para apresentação posterior em um dispositivo de exibição, tal como o dispositivo de exibição 32 da Figura 1.

[0161] Desse modo, o decodificador de vídeo 30 da Figura 3 representa um exemplo de um decodificador de vídeo configurado para determinar se um índice de

referência para o bloco atual corresponde a uma figura de referência intervisualizada e, quando o índice de referência para o bloco atual corresponder à figura de referência de intervisualização, obter, a partir de um fluxo de bits codificado, dados que indicam um modo de previsão de síntese de visualização (VSP) do bloco atual, em que o modo de VSP para o índice de referência indica se o bloco atual é previsto com previsão de síntese de visualização a partir da figura de referência intervisualizada.

[0162] A Figura 4 ilustra uma ordem de decodificação de MVC típica (isto é, ordem de fluxo de bits). A disposição de ordem de decodificação é denominada de codificação de primeiro tempo. Cada unidade de acesso é definida para conter a figura codificada de todas as visualizações para uma instância de tempo de saída. A ordem de decodificação de unidades de acesso pode não ser idêntica à ordem de saída ou exibição.

[0163] A Figura 5 é um diagrama conceitual que ilustra um exemplo padrão de previsão de MVC. A codificação de vídeo de multivisualizações (MVC) é uma extensão de ITU-T H.264/AVC. Uma técnica similar pode ser aplicada a HEVC. No exemplo da Figura 4, oito visualizações (que têm IDs de visualização de "S0" a "S7") são ilustradas e doze locais temporais ("T0" a "T11”) são ilustradas para cada visualização. Isto é, cada fileira na Figura 4 corresponde a uma visualização, enquanto cada coluna indica um local temporal.

[0164] Embora a MVC tenha uma assim chamada visualização de base que pode ser decodificada por decodificadores H.264/AVC e um par de visualizações estéreo que também pode ser sustentado por MVC, uma vantagem da MVC é que a mesma pode sustentar um exemplo que usa mais que

duas visualizações como uma entrada de vídeo 3D e decodifica esse vídeo 3D representado pelas múltiplas visualizações. Um prestador de serviço de um cliente que tem um Decodificador MVC pode esperar um conteúdo de vídeo 3D com múltiplas visualizações.

[0165] As figurações na Figura 5 são indicadas com o uso de um bloco sombreado que inclui uma letra que designa se a figura correspondente é intracodificada (isto é, um quadro I), ou intercodificada em uma direção (isto é, como um quadro P) eu em múltiplas direções (isto é, como um quadro B). Em geral, as previsões são indicadas por flechas, em que a figura apontada usa o objeto a partir do qual se aponta como referência de previsão. Por exemplo, o quadro P da visualização S2 no local temporal T0 é previsto a partir do quadro I da visualização S0 no local temporal T0.

[0166] Em relação ao codificação de vídeo de visualização única, a figura de uma sequência de vídeo de múltiplas visualizações pode ser codificada de modo previsível em relação à figura em diferentes locais temporais. Por exemplo, o quadro b da visualização S0 no local temporal T1 tem uma flecha apontada para o mesmo a partir do quadro I da visualização S0 no local temporal T0, o que indica que o quadro b é previsto a partir do quadro I. Adicionalmente, porém, no contexto do codificação de vídeo de múltiplas visualizações, a figura pode ser prevista por intervisualização. Isto é, um componente de visualização pode usar os componentes de visualização em outras visualizações para referência. Na MVC, por exemplo, a previsão de intervisualização é realizada como se o componente de visualização em outra visualização fosse uma referência de interprevisão. As referências de intervisualização potenciais podem ser sinalizadas na

extensão de MVC de Conjunto de Parâmetro de Sequência (SPS) e podem ser modificadas pelo processo de construção da lista de figura de referência, o que habilita a ordenação flexível das referências de interprevisão ou de previsão de intervisualização.

[0167] A Figura 5 fornece vários exemplos de previsão de intervisualização. As figurações da visualização S1, no exemplo da Figura 5, são ilustradas como previstas a partir da figura em diferentes locais temporais da visualização S1, assim como previstas por intervisualização a partir da figura de figura das visualizações S0 e S2 nos mesmos locais temporais. Por exemplo, o quadro b da visualização S1 no local temporal T1 é previsto a partir de cada um dos quadros B da visualização S1 nos locais temporais T0 e T2, assim como os quadros b das visualizações S0 e S2 no local temporal T1.

[0168] No exemplo da Figura 5, a letra maiúscula "B" e letra minúscula "b" destinam-se a indicar diferentes relacionamentos hierárquicos entre as figurações, em vez de metodologias de codificação diferentes. Em geral, os quadros de "B" maiúsculo são relativamente mais altos na hierarquia de previsão que os quadros de "b" minúsculo. A Figura 5 também ilustra variações na hierarquia de previsão com o uso de níveis diferentes de sombreamento, em que as figurações com uma quantidade maior de sombreamento (isto é, relativamente mais escuro) são mais altas na hierarquia de previsão que aquelas figurações que têm menos sombreamento (isto é, relativamente mais claras). Por exemplo, todos os quadros I na Figura 5 são ilustrados com sombreamento total, enquanto os quadros P têm um sombreamento um pouco mais claro e os quadros B (e os quadros de b minúsculo) têm vários níveis de sombreamento em relação entre si, mas sempre mais claros

que o sombreamento dos quadros P e dos quadros I.

[0169] Em geral, a hierarquia de previsão se relaciona a índices de ordem de visualização, em que as figurações relativamente mais altas na hierarquia de previsão devem ser decodificadas antes das figurações de decodificação que são relativamente mais baixas na hierarquia, de modo que aquelas figurações relativamente mais altas na hierarquia possam ser usadas como figurações de referência durante a decodificação das figurações relativamente mais baixas na hierarquia. Um índice de ordem de visualização é um índice que indica que a ordem de decodificação de componentes de visualização é uma unidade de acesso. Os índices de ordem de decodificação podem estar implícitos em um conjunto de parâmetro, tal como um SPS.

[0170] Desse modo, as figurações usadas como figurações de referência podem ser decodificadas antes da decodificação das figurações que são codificadas em relação às figurações de referência. Um índice de ordem de visualização é um índice que indica a ordem de decodificação de componentes de visualização em uma unidade de acesso. Para cada índice de ordem de visualização i, a view_id correspondente é sinalizada. A decodificação dos componentes de visualização segue a ordem ascendente dos índices de ordem de visualização. Se todas as visualizações forem apresentadas, então, o conjunto de índices de ordem de visualização compreende um conjunto ordenado consecutivamente de zero a um menor que o número total de visualizações.

[0171] Na MVC, um subconjunto de um fluxo de bits total pode ser extraído para formar um subfluxo de bits que ainda se conforma à MVC. Há muitos subfluxos de bits possíveis que as aplicações específicas podem requerer, com base, por exemplo, em um serviço fornecido

por um servidor, a capacidade, suporte e capacidade de decodificadores de um ou mais clientes e/ou a preferência de um ou mais clientes. Por exemplo, um cliente pode requerer apenas três visualizações e pode haver dois cenários. Em um exemplo, um cliente pode requerer uma experiência de visualização estável e pode preferir visualizações com valores de view_id S0, S1 e S2, enquanto outro cliente pode requerer escalabilidade de visualização e preferir visualizações com valores de view_id S0, S2 e S4. Observa-se que esses subfluxos de bits podem ser decodificados como fluxos de bits de MVC independentes e podem ser sustentados simultaneamente.

[0172] Consequentemente, uma estrutura de previsão de MVC típica (incluindo tanto a previsão por interfigura dentro de cada visualização quanto à previsão por intervisualização) para a codificação de vídeo de multivisualizações é mostrada na Figura 5, em que as previsões são indicadas por flechas, em que o objeto apontado usa o objeto a partir do qual se aponta como referência de previsão. Na MVC, a previsão de intervisualização é sustentada pela compensação de movimento de disparidade que usa a sintaxe da compensação de movimento H.264/AVC, mas permite uma figura em uma visualização diferente a ser usada como uma figura de referência.

[0173] A codificação das duas visualizações também pode ser sustentada por MVC e uma das vantagens potenciais da MVC é que um codificador MVC pode tomar mais que duas visualizações como uma entrada de vídeo 3D e um decodificador MVC pode decodificar essa representação de multivisualização. Consequentemente, qualquer prestador de serviço com um decodificador MVC pode esperar conteúdos de vídeo 3D com mais de duas visualizações.

[0174] Em relação à previsão de intervisualização, na MVC, a previsão de intervisualização é permitida entre as figurações na mesma unidade de acesso (isto é, com a mesma instância de tempo). Ao codificar uma figura em uma das visualizações de não base, uma figura pode ser adicionada a uma lista de figura de referência, se a mesma estiver em uma visualização diferente, mas com a mesma instância de tempo. Uma figura de referência de previsão por intervisualização pode ser colocada em qualquer posição de uma lista de figura de referência, como qualquer outra figura de referência de interprevisão.

[0175] As vistas S0 a S7 da Figura 4 representam, exemplos de vistas codificadas, isto é, vistas para as quais informações codificadas são fornecidas em um fluxo de bits. Em alguns exemplos, podem ser sintetizadas vistas adicionais entre as vistas SO a S7. Por exemplo, uma vista pode ser sintetizada entre as vistas SO e S1. Para sintetizar uma figura de tal vista, informações de textura e/ou profundidade de figurações em vistas S0 e S2 podem ser usadas. Por exemplo, para sintetizar uma figura de referência no tempo T1, informações de textura e profundidade provenientes das figurações no tempo T1 a partir das vistas S0 e/ou S2 podem ser usadas. Um codificador de vídeo, como um codificador de vídeo 20 ou decodificador de vídeo 30, pode interpolar dados de pixel para tal figura de uma vista sintetizada para serem usados como uma referência para codificar outra figura, por exemplo, uma figura da vista S1 no tempo T1.

[0176] Aspectos dessa revelação incluem técnicas para indicar se é para aplicar a previsão de síntese de visualização quando codificar um bloco de dados de vídeo. Em um exemplo, um codificador de vídeo, como um codificador de vídeo 20 e/ou um decodificador de vídeo 30,

pode determinar se um índice de referência para um bloco atual corresponde a uma figura de referência de interprevisão, e quando o índice de referência para o bloco atual corresponde à figura de referência de interprevisão, determinar dados que indicam um modo de VSP do bloco atual, em que o modo de VSP para o índice de referência indica se o bloco atual é previsto com o uso de previsão de síntese de visualização, por exemplo, a partir de uma figura de referência de intervisualização. Em casos em que o codificador de vídeo é configurado como o decodificador de vídeo 30, o decodificador de vídeo 30 pode determinar os dados que indicam o modo de VSP obtendo-se os dados de um fluxo de bits codificado.

[0177] Em alguns exemplos, a figura de referência de interprevisão é geralmente associada a um modo interprevisão. Em tais exemplos, os dados que indicam o modo de VSP podem ser determinados para qualquer intermodo, por exemplo, para figurações de referência intravisualização e figurações de referência intervisualização. Em outros exemplos, os dados que indicam o modo de VSP podem ser determinados apenas para figurações de referência intervisualização. Isto é, por exemplo, o decodificador de vídeo 30 pode decodificar apenas os dados que indicam o modo de VSP quando um índice de figura de referência corresponde a uma figura de referência de intervisualização.

[0178] No exemplo acima, um índice de referência de uma figura de VSP não é indicado de modo específico em um cabeçalho de slice ou adicionado durante a construção da lista de figura de referência. Em vez disso, o codificador de vídeo pode codificar um ou mais elementos de sintaxe no nível de MB ou de partição MB para cada índice de referência que corresponde a uma figura de

referência de interprevisão. Em alguns exemplos, os um ou mais elementos de sintaxe podem ser um sinalizador de VSP. Por exemplo, o sinalizador pode indicar se a VSP é aplicada para MB ou partição de MB para um índice de referência particular.

[0179] Em relação à semântica de camada de macrobloco, um sinalizador de VSP igual a 1 pode especificar que o MB ou a partição de MB atual é previsto de modo unidirecional a partir de uma figura de referência de intervisualização, identificado por non_anchor_ref_10[VOIdx][0] quando anchor_pic_flag for 0 ou anchor_ref_10[VOIdx][0] quando anchor_pic_flag for 1, em que VOIdx é o índice de organização de visualização do componente de visualização atual. Os vetores de movimento do MB ou da partição de MB podem ser derivados conforme especificado na subcláusula J.8.3.1.3 (por exemplo, de WD 4). Adicionalmente, um sinalizador de VSP igual a 0 pode especificar que a previsão de síntese de visualização não é usada para o MB ou para a partição de MB atual. De modo alternativo, um sinalizador de VSP igual a 0 pode especificar que a previsão de síntese de visualização sob GVSP seja desabilitada, mas o mecanismo com base em previsão de síntese de visualização com o uso de sinalização do índice de referência pode continuar permitido.

[0180] Em outro exemplo, quando o sinalizador de VSP for igual a 1, o MB ou a partição de MB pode ser previsto a partir da primeira referência intervisualização em RefPicListO. Em outro exemplo, quando o sinalizador de VSP for igual a 1, o MB ou a previsão de MB podem ser previstos a partir da primeira entrada em um RPLM que tenha modification_of_pic_nums_idc igual a 6, quando vsp_pic_flag não for usado para restringir o presente de

modification_of_pic_nums_idc. Nesse caso, o comando com modification_of_pic_nums_idc igual a 6 no RPLM não indica uma nova entrada em uma lista de figura de referência.

[0181] Em relação a restrições de nível, quando o MB ou a partição de MB for codificada com um sinalizador de VSP e tiver vetores diferentes de movimento em pelo menos um bloco 8x8 dentro do MB ou da partição de MB, um codificador de vídeo, como um codificador de vídeo 20 ou um decodificador de vídeo 30, pode contar a quantidade de sub-blocos menores que 8x8 com base no MB ou na partição de MB. A quantidade de sub-blocos não deve exceder MaxSubMbRectSize. Nesse exemplo, o codificador de vídeo pode determinar que um MB previsto por VSP tenha um ou mais sub-blocos menores que 8x8.

[0182] De modo alternativo, em outro exemplo, após prever um MB com o uso de VSP (com um sinalizador de VSP), o codificador de vídeo pode determinar que o MB tenha sub-blocos menores que 8x8. De modo alternativo, ainda em outro exemplo, quando um sinalizador de VSP está presente para um MB ou uma partição de MB, para cada partição de MB, apenas um vetor de movimento intervista é derivado, de modo que o MB ou a partição de MB nunca têm uma partição de bloco menor que 8x8. De modo alternativo, ainda em outro exemplo, quando um sinalizador de VSP está presente para um MB ou uma partição de MB, para cada partição de MB, apenas um vetor de movimento intervista é derivado para cada um dos 4 blocos 8x8, de modo que o MB ou a partição de MB nunca têm uma partição de bloco menor que 8x8.

[0183] De acordo com outros aspectos dessa revelação, quando a previsão de síntese de visualização é habilitada, um codificador de vídeo pode codificar um ou mais elementos de sintaxe em um conjunto de parâmetros, como um SPS ou extensão SPS que indica se a sinalização com

base em referência é usada para a VSP (por exemplo, na qual um índice de figura de referência seja gerado para uma figura de referência de síntese de visualização e a VSP seja realizada com base no índice de referência) ou se a GVSP é usada para a VSP (por exemplo, na qual a VSP é realizada com base na sinalização de macrobloco descrita acima, por exemplo, em relação às Tabelas 1 a 3 acima). Em alguns exemplos, o um ou mais elementos de sintaxe podem incluir um sinalizador em uma extensão de SPS.

[0184] Por exemplo, de acordo com aspectos dessa revelação, o codificador de vídeo 20 e/ou o decodificador de vídeo 30 podem usar a sintaxe de extensão de conjunto de parâmetros de sequência em relação a WD4 (notada acima), conforme mostrado na Tabela 4 abaixo:

TABELA 4 - Sintaxe de Extensão 3DVC de Conjunto de Parâmetros de Sequência

seq_parameter_set_3dvc_extension( ) {		Descritor
for(i = 0; i <= num_level_values_signalled_minus1; i++)		
for(j = 0; j <= num_applicable_ops_minus1[i]; j++)		
applicable_op_num_depth_views_minus1[i][j]		ue(v)
depth_info_present_flag		u(1)
if(depth_info_present_flag) {		
3dv_acquisition_idc		ue(v)
for(i = 0; i < num_views_minus1; i++)		
view_id_3dv[i]		ue(v)
if(3dv_acquisition_idc) {		
depth_ranges(2, 0)		
if(profile_idc == 139 && num_views_minus1)		
vsp_param(2, 0)		
}		
if (profile_idc == 139)		
for(i = 0; i<= num_views_minus1; i++)		
depth_preceding_texture_flag[i]		u(1)
}		
if(profile_idc == 139 && depth_info_present_flag) {		
reduced_resolution_flag		u(1)
slice_header_prediction_flag		u(1)
inside_view_mvp_flag		u(1)
seq_view_synthesis_flag		u(1)
if(seq_view_synthesis_flag)		
vsp_pic_flag		u(1)
disp_flag		u(1)
psip_flag		u(1)
nonlinear_depth_representation_num		ue(v)
for(i = 1; i <= nonlinear_depth_representation_num; i++)		
nonlinear_depth_representation_model[i]		ue(v)
}		
if(profile_idc == 139 && !depth_info_present_flag)		
alc_sps_enable_flag		u(1)
}		

[0185] No exemplo da Tabela 4 acima, vsp_pic_flag igual a 1 indica que a figura de previsão de síntese de visualização pode ser gerada e referenciada por um índice de referência em uma lista de figura de referência. Adicionalmente, vsp_pic_flag igual a 0 indica que a figura de previsão de síntese de visualização não é gerada e a previsão de síntese de visualização sempre se refere a uma figura de referência de intervisualização. Quando não está presente, vsp_pic_flag pode ser inferido como igual a 0.

[0186] De acordo com outros aspectos dessa

revelação, quando o índice de sinalização com base em referência não é usado, um RPLM não deve conter qualquer comando de RPLM que corresponda a uma figura de VSP. Por exemplo, quando o índice de sinalização com base em referência não é usado, um codificador de vídeo (como o codificador de vídeo 20 e/ou o decodificador de vídeo 30) pode não emitir um comando de RPLM que corresponda a uma figura de VSP. Isto é, o codificador de vídeo pode evitar usar um RPLM para mudar a posição de uma figura de VSP, porque nenhuma figura de VSP foi gerada.

[0187] Em um exemplo, em relação às modificações da lista de figura de referência 3DVC, a semântica especificada na sub-cláusula 1.7.4.3.1.1 (por exemplo, em WD 4, notado acima) aplica-se com as seguintes adições:

A entrada de tabela adicional a seguir é inserida na Tabela H-1:

modification_of_pic_nums_idc	Modificação especificada
6	vsp_ref_idx está presente e corresponde a um índice índice de referência de VSP

em que modification_of_pic_nums_idc não deve ser igual a 6,

se seq_view_synthesis_flag ou vsp_pic_flag for igual a 0.

[0188] De acordo com outros aspectos dessa revelação, quando a GVSP é usada, o codificador de vídeo pode sintetizar uma figura de previsão de síntese de visualização apenas a partir da primeira figura de referência de intervisualização para RefPicListO. Nesse exemplo, nenhuma sinalização da figura de previsão de síntese de visualização é necessária e é automaticamente determinado (por exemplo, inferido) que o modo de GVSP emite um vetor de movimento de disparidade que aponta para

a primeira figura de referência de intervisualização, por exemplo, RefPicList 0. De modo alternativo, em outro exemplo, o codificador de vídeo pode realizar a previsão de síntese de visualização com base em uma figura intervisualização em RefPicListl . De modo alternativo, em outro exemplo, o codificador de vídeo pode realizar a previsão de síntese de visualização com base em figurações intervisualizações, caso haja, tanto em RefPicListO quanto em RefPicListl .

[0189] A Figura 6 é um fluxograma que ilustra um método exemplificativo para codificar informações relacionadas à previsão de síntese de visualização. O método da Figura 6 é explicado em relação ao codificador de vídeo 20 (Figuras 1 e 2). Entretanto, deve ser notado que outros dispositivos de codificação de vídeo podem ser configurados para realizar um método similar. Ademais, certas etapas no método podem ser realizadas em uma ordem diferente ou em paralelo. Do mesmo modo, certas etapas podem ser omitidas e outras etapas podem ser adicionadas, em vários exemplos.

[0190] Nesse exemplo, o codificador de vídeo 20 prevê inicialmente o bloco atual e o índice de referência do bloco de previsão (120). Nesse exemplo, admite-se que o codificador de vídeo 20 interpreveja o bloco atual. Por exemplo, a unidade de estimativa de movimento 42 pode calcular um vetor de movimento para o bloco atual realizando-se uma busca de movimento de figurações previamente codificadas, por exemplo, figurações intervisualização e figurações temporais. Assim, a unidade de estimativa de movimento 42 pode produzir um vetor de movimento temporal ou um vetor de movimento de disparidade para prever o bloco atual.

[0191] Em alguns casos, conforme notado

acima, o codificador de vídeo 20 pode prever o bloco atual a partir de um bloco sintetizado. Por exemplo, o codificador de vídeo 20 pode realizar um processo de síntese de visualização para gerar uma figura sintetizada, que pode ser adicionada a uma lista de figura de referência e ter um índice de figura de referência. O codificador de vídeo 20 pode realizar a VSP para prever o bloco atual em relação a um bloco da figura sintetizada.

[0192] No exemplo da Figura 6, o codificador de vídeo 20 pode determinar se o índice de referência da figura de referência que inclui o bloco de previsão é uma figura de referência de intervisualização (122). Em alguns exemplos, o codificador de vídeo 20 pode determinar se a figura de referência está incluída em uma visualização diferente da do bloco que é codificado com base em um identificador de visualização (view_id).

[0193] De acordo com aspectos dessa revelação, se o índice de referência corresponder a uma figura de referência de intervisualização (a ramificação SIM da etapa 122), o codificador de vídeo 20 pode codificar dados que indicam se é para aplicar a VSP para o índice de referência. Por exemplo, em casos em que o codificador de vídeo20 realizou a VSP para determinar o bloco de previsão, o codificador de vídeo 20 pode codificar dados que indicam que a VSP foi usada (por exemplo, definir um sinalizador de VSP igual a 1). Em casos em que o codificador de vídeo 20 simplesmente realizou a previsão intervisualização para determinar o bloco de previsão (no VSP), o codificador de vídeo 20 pode codificar dados que indicam que a VSP não foi usada (por exemplo, definir um sinalizador de VSP igual a 0).

[0194] O codificador de vídeo 20 pode calcular, então, um bloco residual para o bloco atual

(126). para calcular o bloco residual, o codificador de vídeo 20 pode calcular uma diferença entre o bloco original não codificado e o bloco de previsão, que pode ser um bloco na mesma visualização que o bloco que é codificado, um bloco em uma visualização diferente da do bloco que é codificado, ou um bloco sintetizado. O codificador de vídeo 20, então, pode transformar e quantizar coeficientes do bloco residual (128). Em seguida, o codificador de vídeo 20 pode varrer os coeficientes de transformada quantizados do bloco residual (130). Durante ou após a varredura, o codificador de vídeo 20 pode codificar por entropia os coeficientes (132). Por exemplo, o codificador de vídeo 20 pode codificar os coeficientes com o uso de CAVLC ou CAB AC. O codificador de vídeo 20, então, pode emitir os dados codificados por entropia do bloco assim como a indicação de VSP (134).

[0195] Desse modo, o método da Figura 6 representa um exemplo de um método para codificar dados de vídeo, sendo que o método inclui determinar se um índice de referência para um bloco atual corresponde a uma figura de referência de intervisualização, quando o índice de referência para o bloco atual corresponde à figura de referência de intervisualização, gerar dados que indicam um modo de previsão de síntese de visualização (VSP) do bloco atual, em que o modo de VSP para o índice de referência indica se o bloco atual está previsto com a previsão de síntese de visualização a partir da figura de referência de intervisualização, e codificar os dados que indicam o modo de VSP em um fluxo de bits.

[0196] A Figura 7 é um fluxograma que ilustra um método exemplificativo para codificar informações relacionadas à previsão de síntese de visualização. O método da Figura 7 é explicado em relação ao decodificador

de vídeo 30 (Figuras 1 e 3). Entretanto, deve ser notado que outros dispositivos de codificação de vídeo podem ser configurados para realizar um método similar. Ademais, certas etapas no método podem ser realizadas em uma ordem diferente ou em paralelo. Do mesmo modo, certas etapas podem ser omitidas e outras etapas podem ser adicionadas, em vários exemplos.

[0197] A unidade de decodificação por entropia 70 decodifica os dados para os coeficientes do bloco que está atualmente em decodificação, pelo menos um vetor de movimento ou vetor de disparidade e pelo menos um índice de referência correspondente (160). O exemplo da Figura 7 admite que o bloco atual é monoprevisto, mas deve ser entendido que em outros exemplos, o bloco atual pode ser biprevisto, conforme descrito no presente documento.

[0198] O vetor de movimento ou o vetor de disparidade podem identificar o bloco de previsão na figura de referência associada ao índice de figura de referência decodificado. Por exemplo, o bloco atual pode ser interprevisto a partir de um bloco na mesma visualização que o bloco atual, previsto intervisualização a partir de um bloco em uma visualização diferente da do bloco atual ou previsto intervisualização a partir de um bloco sintetizado com o uso de VSP. Independente do processo de previsão particular, o índice de referência decodificado identifica uma figura armazenada a uma memória de figura de referência.

[0199] No exemplo da Figura 7, o decodificador de vídeo 30 pode determinar se o índice de referência da figura de referência que inclui o bloco de previsão é uma figura de referência de intervisualização (162). Em alguns exemplos, o decodificador de vídeo 30 pode determinar se a figura de referência está incluída em uma

visualização diferente da do bloco que é codificado com base em um identificador de visualização (view_id).

[0200] De acordo com aspectos dessa revelação, se o índice de referência corresponder a uma figura de referência de intervisualização (a ramificação SIM da etapa 162), o decodificador de vídeo 30 pode decodificar dados que indicam se é para aplicar a VSP para o índice de referência. Por exemplo, em casos em que a VSP deve ser usada para determinar o bloco de previsão, o decodificador de vídeo 30 pode decodificar dados que indicam que a VSP deve ser usada (por exemplo, o decodificador de vídeo 30 pode obter, a partir de um fluxo de bits codificado, um sinalizador de VSP igual a 1). Em casos em que o decodificador de vídeo 30 deve realizar previsão intervisualização sem a VSP para determinar o bloco de previsão, o decodificador de vídeo 30 pode decodificar dados que indicam que a VSP não deve ser usada (por exemplo, o decodificador de vídeo 30 pode obter, a partir de um fluxo de bits codificado, um sinalizador de VSP igual a 0). Em alguns exemplos, se o índice de referência corresponde a uma figura de referência de intervisualização mas não há indicação de VSP incluída no fluxo de bits (por exemplo, nenhum sinalizador de VSP está presente), o decodificador de vídeo 30 pode determinar automaticamente (inferir) que o bloco atual é previsto intervisualização sem VSP.

[0201] O decodificador de vídeo 30 pode prever o bloco atual com o uso do movimento decodificado ou do vetor de disparidade e do índice de figura de referência (166). Em casos em que a VSP é usada, o decodificador de vídeo 30 pode gerar uma figura sintetizada para determinar o bloco de previsão. O decodificador de vídeo 30, então, pode varrer inversamente os coeficientes reproduzidos

(168), para criar um bloco de coeficientes de transformada quantizados. O decodificador de vídeo 30, então, pode, quantizar inversamente e transformar inversamente os coeficientes para produzir um bloco residual (170). O decodificador de vídeo 30, por fim, pode decodificar o bloco atual combinando-se o bloco previsto e o bloco residual (172).

[0202] Desse modo, o método da Figura 7 representa um exemplo de um método que inclui determinar se um índice de referência para um bloco atual corresponde a uma imagem de referência de intervisualização, e quando o índice de referência para o bloco atual corresponder à imagem de referência de intervisualização, obter, a partir de uma corrente de bits codificada, dados que indicam um modo de previsão de síntese de visualização (VSP) do bloco atual, em que o modo de VSP para o índice de referência indica se o bloco atual está previsto com a previsão de síntese de visualização a partir da imagem de referência de intervisualização.

[0203] As técnicas descritas acima em relação à síntese de visualização podem ser realizadas pelo codificador de vídeo 20 (Figuras 1 e 2) e/ou decodificador de vídeo 30 (Figuras 1 e 3), ambos os quais podem ser referidos de modo geral como um codificador de vídeo. Do mesmo modo, "codificar" e "codificar vídeo" podem se referir tanto ao codificação de vídeo (por exemplo, por um codificador de vídeo) ou à decodificação de vídeo (por exemplo, por um decodificador de vídeo), conforme aplicável.

[0204] Deve ser entendido que, dependendo-se do exemplo, certos atos ou eventos de qualquer um dos métodos descritos no presente documento podem ser realizados em uma sequência diferente, podem ser

adicionados, fundidos ou deixados de fora todos juntos (por exemplo, nem todos os atos e eventos descritos são necessários para a prática do método). Ademais, em certos exemplos, os atos e eventos podem ser realizados de modo concorrente, por exemplo, através de processamento multifiletado, processamento interrupto ou em múltiplos processadores, em vez de sequencialmente. Adicionalmente, embora certos aspectos dessa revelação sejam descritos como realizados por um único módulo ou unidade para propósitos de clareza, deve ser entendido que as técnicas dessa revelação podem ser realizadas por uma combinação de unidades ou módulos associados a um codificador de vídeo.

[0205] Embora combinações em particulares de vários aspectos das técnicas sejam descritas acima, essas combinações são fornecidas meramente para ilustrar exemplos das técnicas descritas nessa revelação. Portanto, as técnicas dessa revelação não devem ser limitadas a essas combinações exemplificativas e podem abranger qualquer combinação concebível dos vários aspectos das técnicas descritas nessa revelação.

[0206] Em um ou mais exemplos, as funções descritas podem ser implantadas em hardware, software, firmware ou qualquer combinação dos mesmos. Caso implantado em software, as funções podem ser armazenadas em, ou transmitidas sobre, como uma ou mais instruções ou código em um meio legível por computador e executadas por uma unidade de processamento com base em hardware. As mídias legíveis por computador podem incluir mídias de armazenamento legíveis por computador, que correspondem a uma mídia tangível como mídias de armazenamento de dados ou mídias de comunicação que incluem qualquer mídia que facilite a transferência de um programa de computador proveniente de um lugar para outro, por exemplo, de acordo

com um protocolo de comunicação.

[0207] Desse modo, as mídias legíveis por computador geralmente podem corresponder a (1) mídias tangíveis de armazenamento legíveis por computador que sejam não transitórias ou (2) uma mídia de comunicação como um sinal ou onda transportadora. As mídias de armazenamento de dados podem ser quaisquer mídias disponíveis que possam ser acessadas por um ou mais computadores ou um ou mais processadores para recuperar instruções, estruturas de código e/ou dados para a implantação das técnicas descritas nessa revelação. Um produto de programa de computador pode incluir uma mídia legível por computador.

[0208] A título de exemplo, e não de limitação, tais mídias de armazenamento legíveis por computador podem compreender RAM, ROM, EEPROM, CD-ROM ou outro armazenamento de disco óptico, armazenamento de disco magnético ou outros dispositivos de armazenamento magnético, ou qualquer outro meio que possa ser usado para armazenar o código de programa desejado na forma de instruções ou estruturas de dados e que possa ser acessado por um computador. Também, qualquer conexão é propriamente denominada uma mídia legível por computador. Por exemplo, se as instruções forem transmitidas a partir de um sítio da web, servidor ou outra fonte remota com o uso de um cabo coaxial, cabo de fibra óptica, par trançado, linha de inscrição digital (DSL) ou tecnologias sem fio como infravermelho, rádio e micro-onda, então, o cabo coaxial, o cabo de fibra óptica, o par trançado, a DSL ou as tecnologias sem fio como infravermelho, rádio e micro-onda estão incluídos na definição de mídia.

[0209] Deve ser entendido, entretanto, que as mídias de armazenamento legíveis por computador e as mídias de armazenamento de dados não incluem conexões, ondas

transportadoras, sinais ou outras mídias transitórias, mas são, em vez disso, direcionadas para mídias não transientes e tangíveis. Disco magnético e disco óptico, conforme usado no presente documento, inclui disco compacto (CD), disca laser, disco ótico, disco versátil digital (DVD), disquete e disco blu-ray onde discos magnéticos normalmente reproduzem dados magneticamente, enquanto discos ópticos reproduzem dados opticamente com lasers. As combinações dos supracitados também devem ser abrangidas pelo escopo de meios legíveis por computador.

[0210] As instruções podem ser executadas por um ou mais processadores, como um ou mais processadores de sinal digital (DSPs), microprocessadores para propósitos gerais, circuitos integrados de aplicação específica (ASICs), matrizes lógicas programáveis por campo (FPGAs) ou outro conjunto de circuitos lógicos equivalentes integrados ou discretos. Portanto, o termo "processador," conforme usado no presente documento pode se referir a qualquer uma das estruturas supracitadas ou qualquer outra estrutura adequada para a implantação das técnicas descritas no presente documento. Adicionalmente, em alguns aspectos, a funcionalidade descrita no presente documento pode ser fornecida dentro de módulos dedicados de hardware e/ou software configurados para codificar e decodificar ou incorporados em um codec combinado. Também, as técnicas podem ser totalmente implantadas em um ou mais circuitos ou elementos lógicos.

[0211] As técnicas dessa revelação podem ser implantadas em uma ampla variedade de dispositivos ou aparelhos, incluindo um monofone, um circuito integrado (IC) ou um conjunto de ICs (por exemplo, um conjunto de chips). Vários componentes, módulos ou unidades são descritos nessa revelação para enfatizar os aspectos

funcionais dos dispositivos configurados para realizar as técnicas reveladas, mas não exigem necessariamente a realização por diferentes unidades de hardware. Em vez disso, conforme descrito acima, várias unidades podem ser combinadas em uma unidade de hardware de codec ou fornecidas por uma coleção de unidades de hardware interoperativos, incluindo um ou mais processadores conforme descrito acima, em conjunto com software e/ou firmware adequados.

[0212] Vários aspectos da revelação foram descritos. Esses e outros aspectos estão dentro do escopo das reivindicações a seguir.

REIVINDICAÇÕES

1. Método de decodificação de dados de vídeo de multivisualização, **caracterizado pelo fato de que** compreende:

determinar se um índice de referência de uma figura de referência em uma lista de figura de referência para prever um bloco atual corresponde a uma figura de referência de intervisualização; e

com base no índice de referência para o bloco atual correspondendo à figura de referência de intervisualização, obter, a partir de um fluxo de bits codificado, dados que indicam se o bloco atual está previsto com a previsão de síntese de visualização, VSP, utilizando pelo menos uma porção de uma figura sintetizada a partir da figura de referência de intervisualização, utilizando renderização com base em imagem de profundidade;

em que obter os dados que indicam se o bloco atual é previsto com VSP, compreende obter um sinalizador de VSP em um nível de bloco apenas quando o índice de referência para o bloco atual corresponder a uma figura de referência de intervisualização; e

decodificar o bloco atual usando previsão de síntese de visualização, VSP, se o sinalizador de VSP indicar que VSP deve ser utilizado.

2. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** compreende adicionalmente, decodificar um cabeçalho de *slice* que não inclui uma indicação de um índice de referência de uma figura de referência de síntese de visualização para decodificar o bloco atual.

3. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** os dados que indicam se o bloco atual é previsto com VSP compreendem dados que

indicam um primeiro modo de VSP para o bloco atual, o método compreendendo adicionalmente, obter, a partir do fluxo de bits codificado, dados que indicam se é para realizar um dentre o primeiro modo de VSP e um segundo modo de VSP, em que realizar o segundo modo de VSP compreende determinar uma figura de referência de síntese de visualização com base em um índice de figura de referência da figura de referência de síntese de visualização.

4. Método, de acordo com a reivindicação 1, **caracterizado pelo fato de que** compreende adicionalmente:

determinar uma figura de referência de síntese de visualização a partir da figura de referência de intervisualização;

adicionar a figura de previsão de síntese de visualização à lista de figura de referência, em que um processo de modificação de lista de figura de referência não é usado para reorganizar a figura de previsão de síntese de visualização; e

decodificar o bloco atual em relação à figura de referência de síntese de visualização.

5. Método de codificação de dados de vídeo de multivisualização, **caracterizado pelo fato de que** compreende:

determinar se um índice de referência de uma figura de referência em uma lista de figura de referência para prever um bloco atual corresponde a uma figura de referência de intervisualização;

com base no índice de referência correspondendo à figura de referência de intervisualização, gerar dados que indicam se o bloco atual está previsto com a previsão de síntese de visualização, VSP, utilizando pelo menos uma porção de uma figura sintetizada a partir da figura de referência de intervisualização utilizando renderização com

base em imagem de profundidade; e

codificar os dados que indicam o modo de VSP em um fluxo de bits;

em que gerar os dados que indicam se o bloco atual é previsto com VSP compreende gerar um sinalizador de VSP em um nível de bloco apenas quando o índice de referência para o bloco atual corresponder a uma figura de referência de intervisualização; e

codificar o bloco atual utilizando previsão de síntese de visualização, VSP, se o sinalizador de VSP indicar que VSP deve ser utilizado.

6. Método, de acordo com a reivindicação 1 ou 5, **caracterizado pelo fato de que** o nível de bloco é um nível de macrobloco quando o bloco atual for um macrobloco, ou um nível de partição de macrobloco quando o bloco atual for uma partição de macrobloco.

7. Método, de acordo com a reivindicação 5, **caracterizado pelo fato de que** compreende adicionalmente codificar um cabeçalho de *slice* que não inclui uma indicação de um índice de referência de uma figura de referência de síntese de visualização para codificar o bloco atual.

8. Método, de acordo com a reivindicação 1 ou 5, **caracterizado pelo fato de que** a lista de figura de referência compreende somente uma lista de figura de referência zero.

9. Método, de acordo com a reivindicação 1 ou 5, **caracterizado pelo fato de que** a lista de figura de referência compreende uma dentre uma primeira lista de figura de referência e uma segunda lista de figura de referência.

10. Método, de acordo com a reivindicação 5, **caracterizado pelo fato de que** os dados que indicam se o

bloco atual é previsto com VSP compreendem dados que indicam um primeiro modo de VSP para o bloco atual, e que compreende, adicionalmente, codificar dados que indicam se é para realizar um dentre o modo de VSP e um segundo modo de VSP, em que realizar o segundo modo de VSP compreende determinar uma figura de referência de síntese de visualização com base em um índice de figura de referência da figura de referência de síntese de visualização.

11. Método, de acordo com a reivindicação 5, **caracterizado pelo fato de que** compreende adicionalmente:

determinar uma figura de referência de síntese de visualização a partir da figura de referência de intervisualização;

adicionar a figura de previsão de síntese de visualização à lista de figura de referência, em que um processo de modificação de lista de figura de referência não é usado para reorganizar a figura de previsão de síntese de visualização; e

codificar o bloco atual em relação à figura de referência de síntese de visualização.

12. Método, de acordo com a reivindicação 1 ou 5, **caracterizado pelo fato de que** compreende adicionalmente, sintetizar uma figura de referência de síntese de visualização a partir da figura de referência de intervisualização apenas quando a figura de referência de intervisualização for incluída na lista de figura de referência zero.

13. Método, de acordo com a reivindicação 1 ou 5, **caracterizado pelo fato de que** compreende adicionalmente, sintetizar uma figura de referência de síntese de visualização a partir da figura de referência de intervisualização apenas quando a figura de referência de intervisualização for incluída em lista de figura de

referência um.

14. Aparelho para codificar dados de vídeo de multivisualização, **caracterizado pelo fato de que** compreende:

meios para determinar se um índice de referência de uma figura de referência em uma lista de figura de referência para prever um bloco atual corresponde a uma figura de referência de intervisualização; e

com base no índice de referência correspondendo à figura de referência de intervisualização, meios para codificar dados que indicam se o bloco atual está previsto com a previsão de síntese de visualização, VSP, utilizando pelo menos uma porção de uma figura sintetizada a partir da figura de referência de intervisualização utilizando renderização com base em imagem de profundidade;

em que os meios para codificar os dados que indicam se o bloco atual é previsto com VSP compreendem meios para codificar um sinalizador de VSP em um nível de bloco apenas quando o índice de referência para o bloco atual corresponder a uma figura de referência de intervisualização; e

meios para codificar o bloco atual utilizando previsão de síntese de visualização, VSP, se o sinalizador de VSP indicar que VSP deve ser utilizado.

15. Memória **caracterizada pelo fato de que** compreende instruções armazenadas na mesma que, quando executadas, fazem com que um ou mais processadores realizem o método conforme definido em qualquer uma das reivindicações 1 a 13.

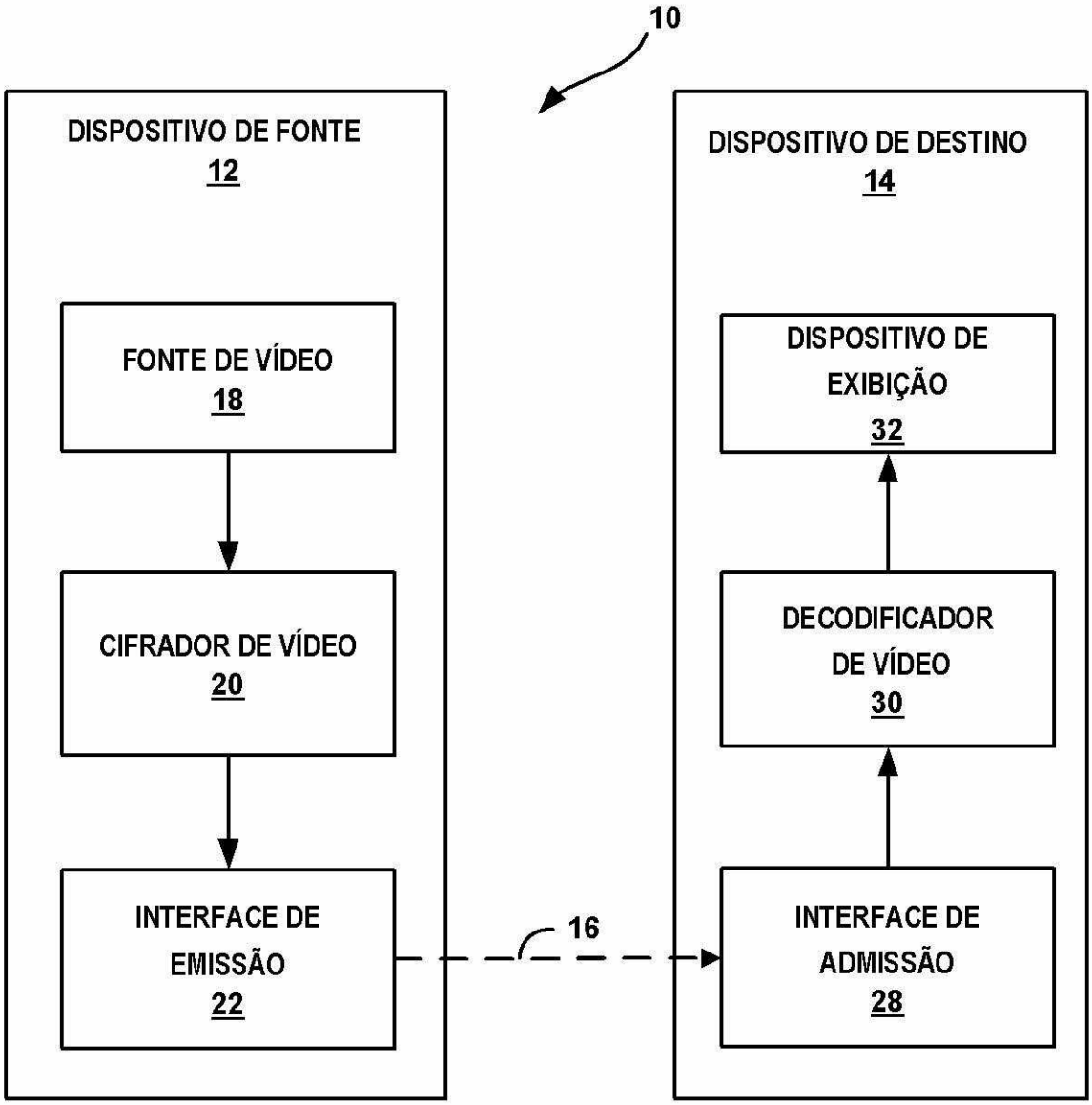

10
DISPOSITIVO DE FONTE
12
FONTE DE VÍDEO
18
CIFRADOR DE VÍDEO
20
INTERFACE DE EMISSÃO
22
16
DISPOSITIVO DE DESTINO
14
DISPOSITIVO DE EXIBIÇÃO
32
DECODIFICADOR DE VÍDEO
30
INTERFACE DE ADMISSÃO
28

FIG. 1

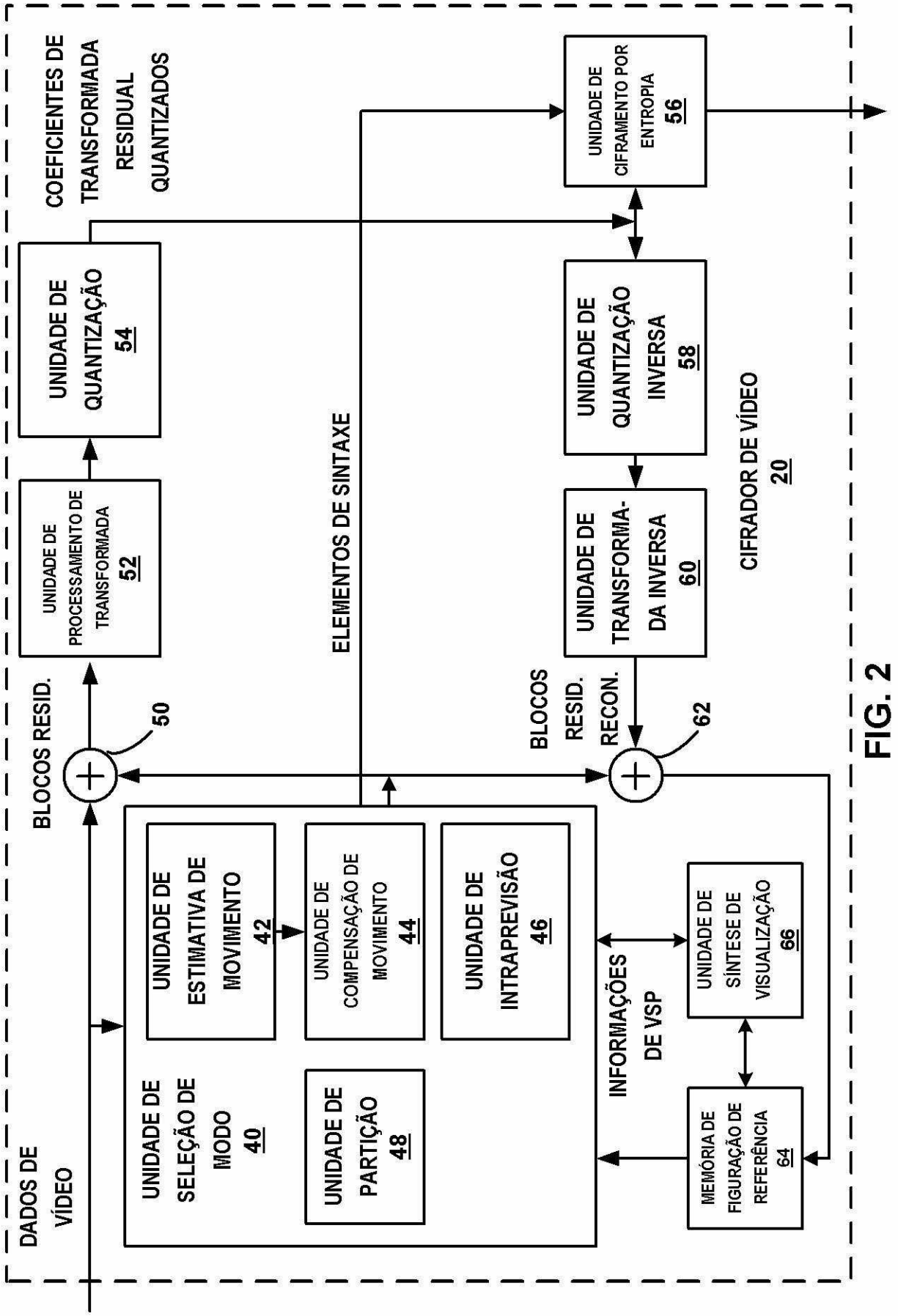

DADOS DE VÍDEO
UNIDADE DE SELEÇÃO DE MODO 40
UNIDADE DE ESTIMATIVA DE MOVIMENTO 42
UNIDADE DE COMPENSAÇÃO DE MOVIMENTO 44
UNIDADE DE INTRAPREVISÃO 46
UNIDADE DE PARTIÇÃO 48
BLOCOS RESID. 50
UNIDADE DE PROCESSAMENTO DE TRANSFORMADA 52
UNIDADE DE QUANTIZAÇÃO 54
COEFICIENTES DE TRANSFORMADA RESIDUAL QUANTIZADOS
UNIDADE DE CIFRAMENTO POR ENTROPIA 56
UNIDADE DE QUANTIZAÇÃO INVERSA 58
UNIDADE DE TRANSFORMADA INVERSA 60
BLOCOS RESID. RECON. 62
ELEMENTOS DE SINTAXE
INFORMAÇÕES DE VSP
UNIDADE DE SÍNTESE DE VISUALIZAÇÃO 66
MEMÓRIA DE FIGURAÇÃO DE REFERÊNCIA 64
CIFRADOR DE VÍDEO 20

FIG. 2

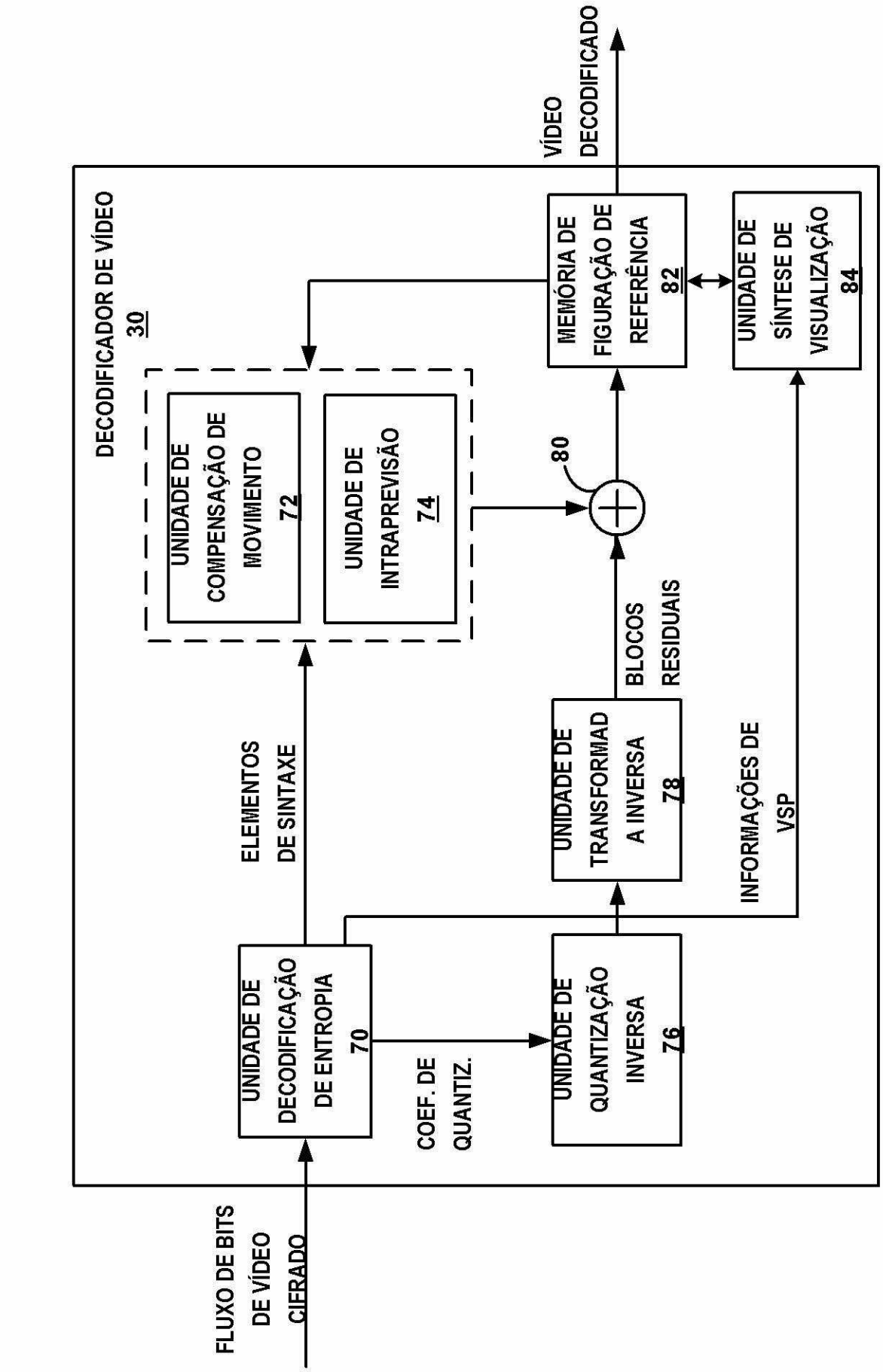

FLUXO DE BITS DE VÍDEO CIFRADO
DECODIFICADOR DE VÍDEO 30
UNIDADE DE DECODIFICAÇÃO DE ENTROPIA 70
UNIDADE DE COMPENSAÇÃO DE MOVIMENTO 72
UNIDADE DE INTRAPREVISÃO 74
UNIDADE DE QUANTIZAÇÃO INVERSA 76
UNIDADE DE TRANSFORMADA INVERSA 78
80
MEMÓRIA DE FIGURAÇÃO DE REFERÊNCIA 82
UNIDADE DE SÍNTESE DE VISUALIZAÇÃO 84
VÍDEO DECODIFICADO
ELEMENTOS DE SINTAXE
COEF. DE QUANTIZ.
BLOCOS RESIDUAIS
INFORMAÇÕES DE VSP

FIG. 3

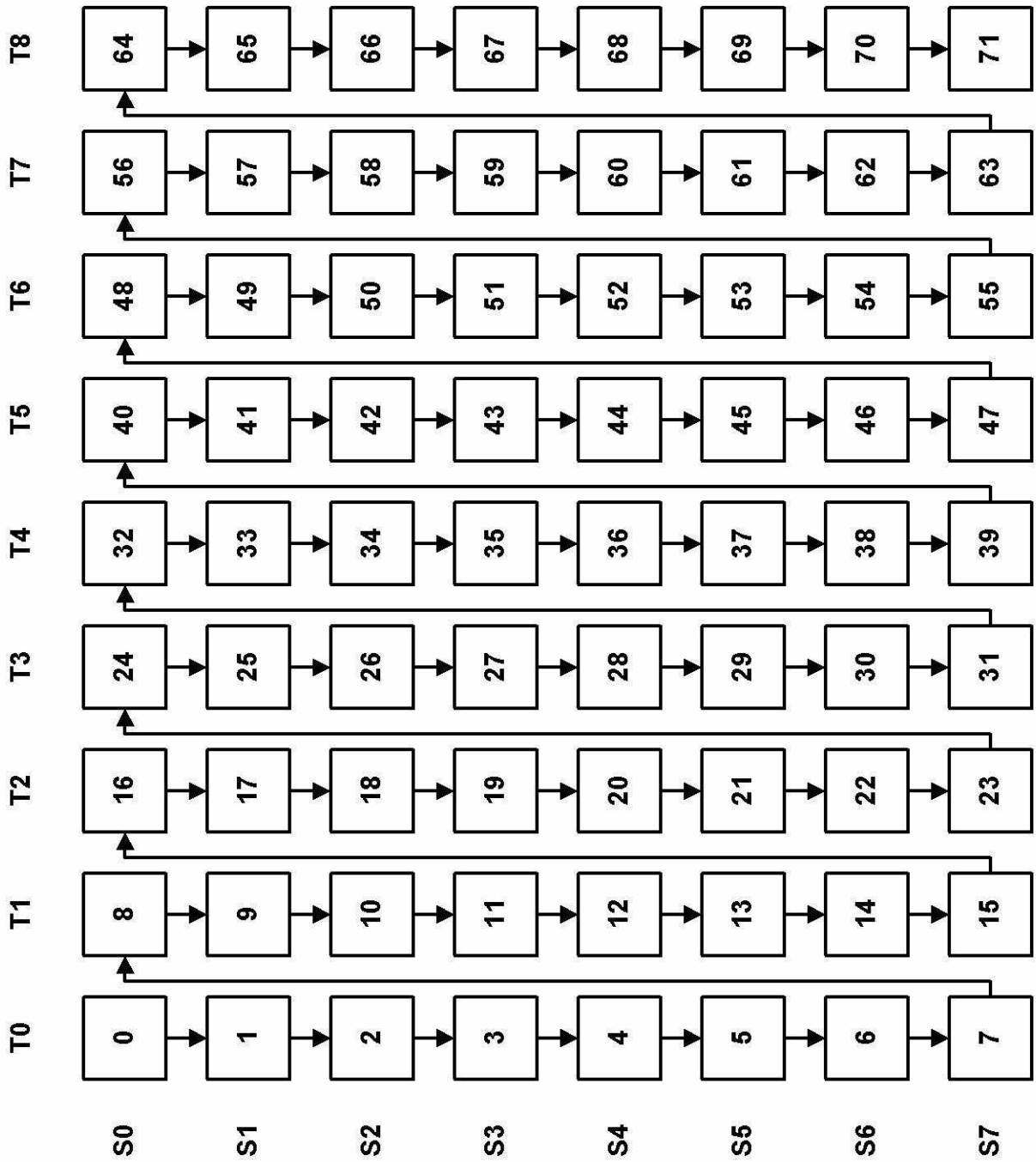

T0
T1
T2
T3
T4
T5
T6
T7
T8
S0
S1
S2
S3
S4
S5
S6
S7
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71

FIG. 4

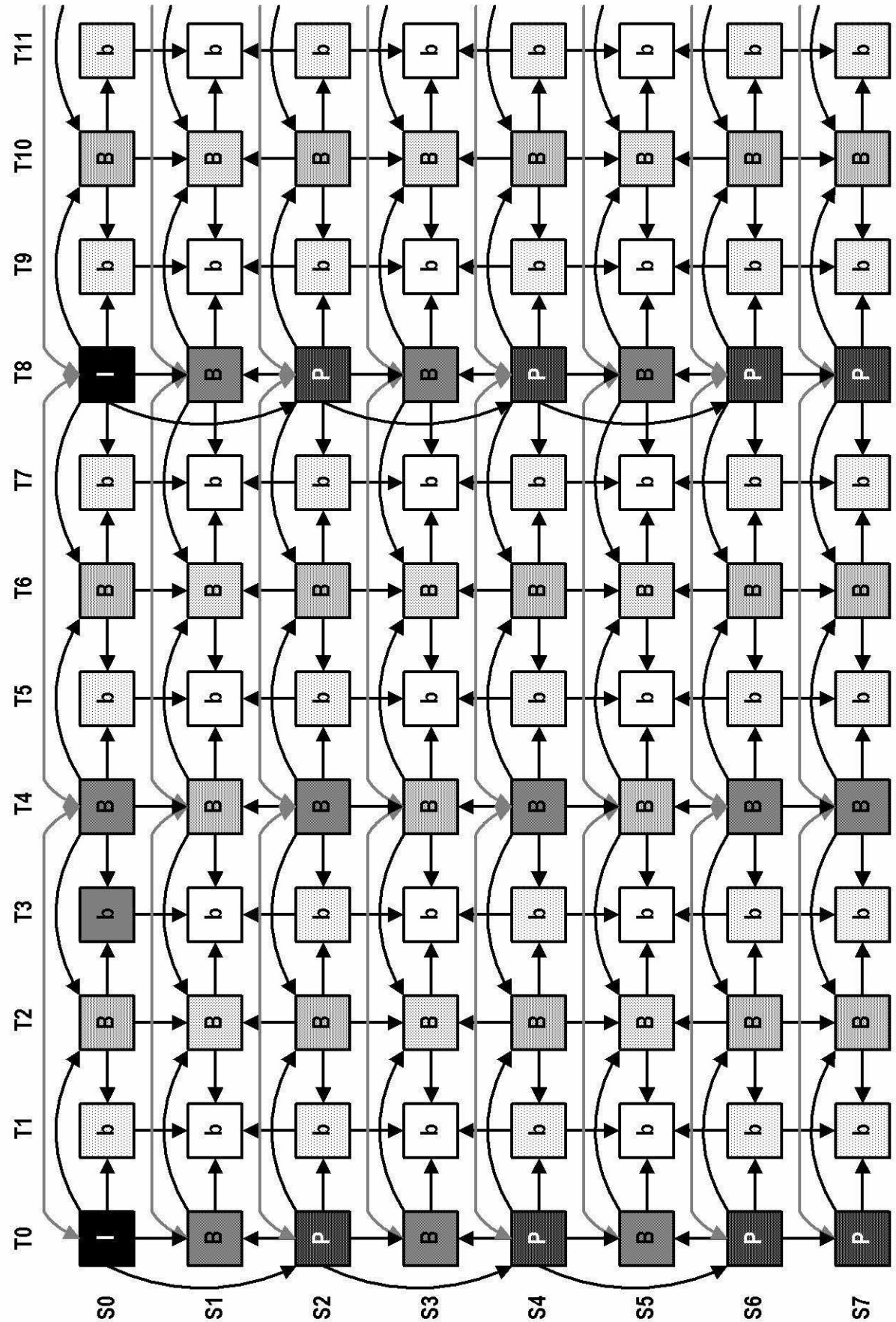


FIG. 5

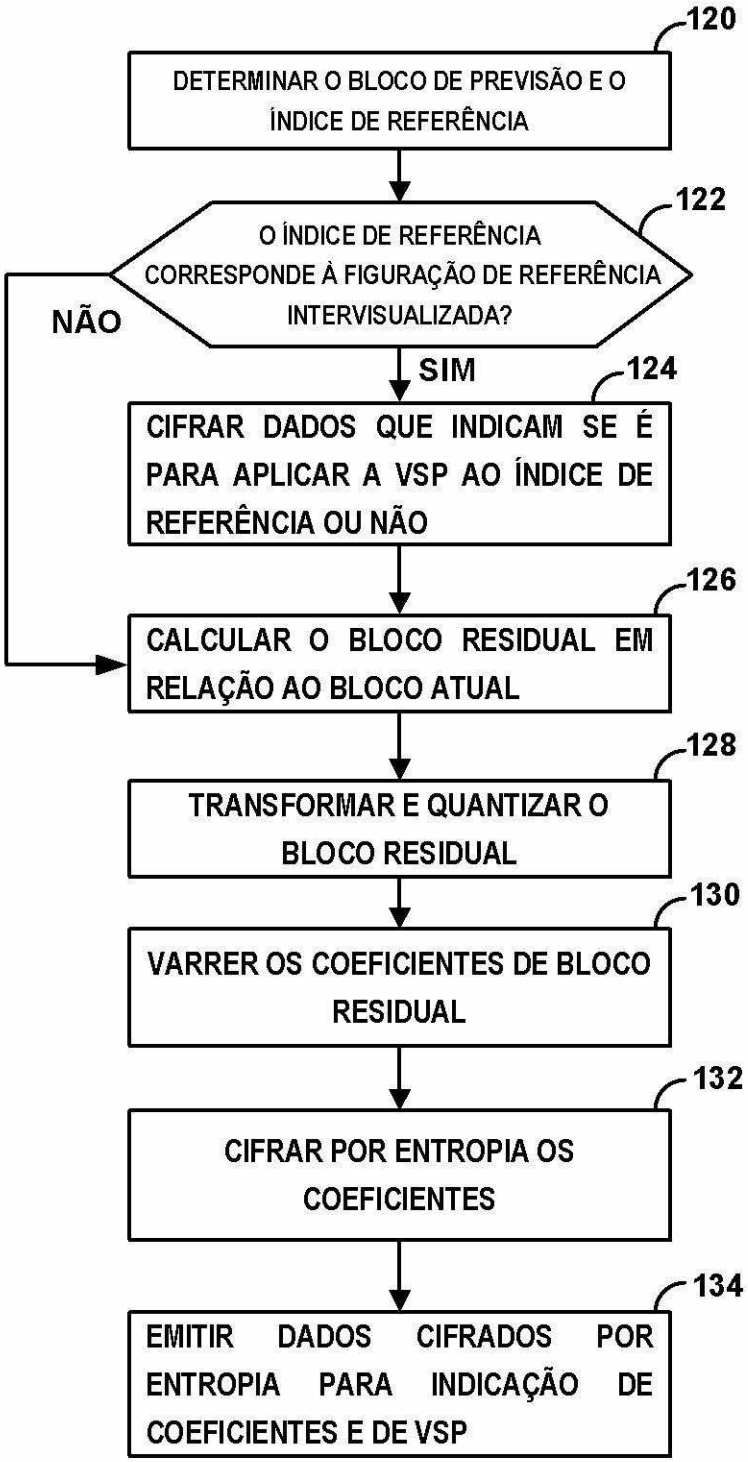

120
DETERMINAR O BLOCO DE PREVISÃO E O ÍNDICE DE REFERÊNCIA
122
O ÍNDICE DE REFERÊNCIA CORRESPONDE À FIGURAÇÃO DE REFERÊNCIA INTERVISUALIZADA?
NÃO
SIM
124
CIFRAR DADOS QUE INDICAM SE É PARA APLICAR A VSP AO ÍNDICE DE REFERÊNCIA OU NÃO
126
CALCULAR O BLOCO RESIDUAL EM RELAÇÃO AO BLOCO ATUAL
128
TRANSFORMAR E QUANTIZAR O BLOCO RESIDUAL
130
VARRER OS COEFICIENTES DE BLOCO RESIDUAL
132
CIFRAR POR ENTROPIA OS COEFICIENTES
134
EMITIR DADOS CIFRADOS POR ENTROPIA PARA INDICAÇÃO DE COEFICIENTES E DE VSP

FIG. 6

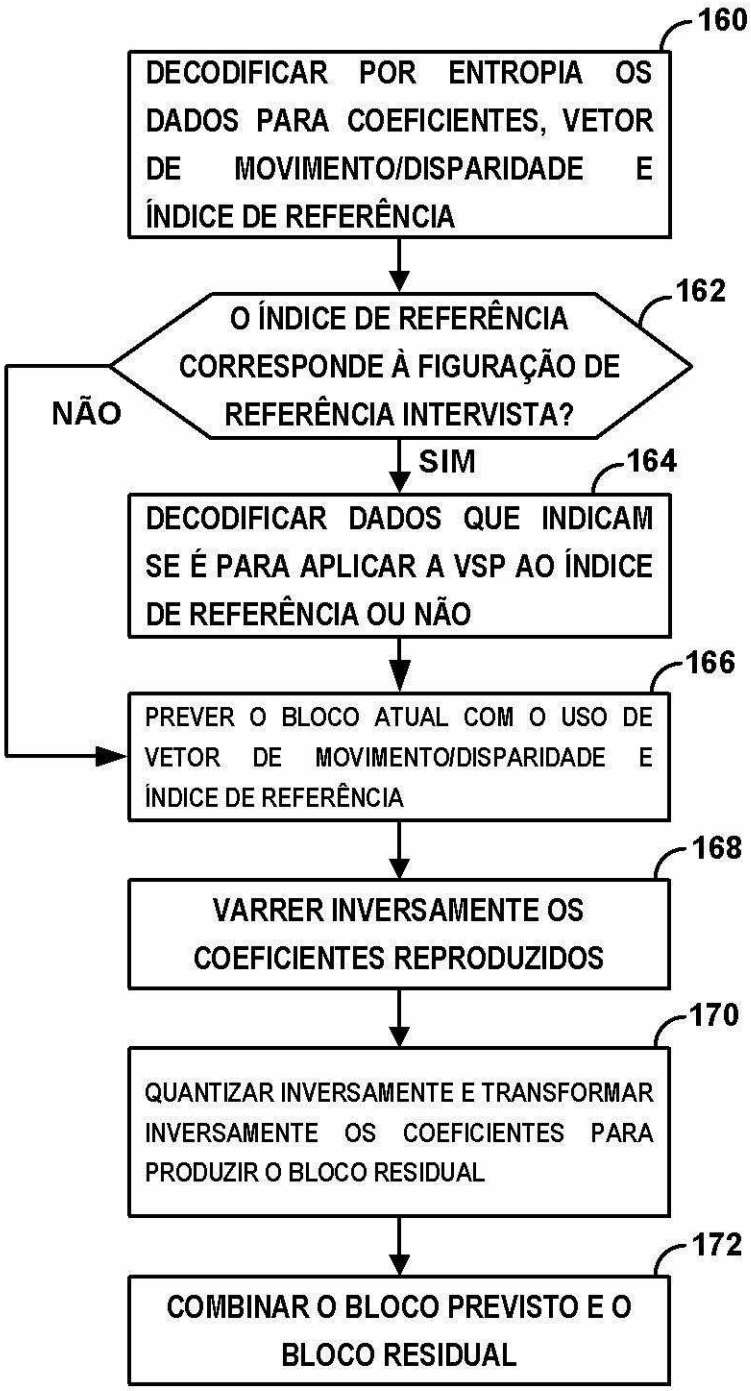

160
DECODIFICAR POR ENTROPIA OS DADOS PARA COEFICIENTES, VETOR DE MOVIMENTO/DISPARIDADE E ÍNDICE DE REFERÊNCIA
162
O ÍNDICE DE REFERÊNCIA CORRESPONDE À FIGURAÇÃO DE REFERÊNCIA INTERVISTA?
NÃO
SIM
164
DECODIFICAR DADOS QUE INDICAM SE É PARA APLICAR A VSP AO ÍNDICE DE REFERÊNCIA OU NÃO
166
PREVER O BLOCO ATUAL COM O USO DE VETOR DE MOVIMENTO/DISPARIDADE E ÍNDICE DE REFERÊNCIA
168
VARRER INVERSAMENTE OS COEFICIENTES REPRODUZIDOS
170
QUANTIZAR INVERSAMENTE E TRANSFORMAR INVERSAMENTE OS COEFICIENTES PARA PRODUZIR O BLOCO RESIDUAL
172
COMBINAR O BLOCO PREVISTO E O BLOCO RESIDUAL

FIG. 7