



República Federativa do Brasil

Ministério do Desenvolvimento, Indústria,
Comércio e Serviços

Instituto Nacional da Propriedade Industrial

(11) BR 112015017192-3 B1

(22) Data do Depósito: 17/01/2014

(45) Data de Concessão: 07/11/2023

(54) Título: MÉTODO E DISPOSITIVO PARA ENCAPSULAR DADOS DE MÍDIA PROGRAMADA PARTICIONADOS, MEIO DE ARMAZENAMENTO E MEIO DE GRAVAÇÃO

(51) Int.Cl.: H04N 21/854; H04N 21/845; H04N 21/4728.

(30) Prioridade Unionista: 09/04/2013 GB 1306451.4; 18/01/2013 GB 1300953.5.

(73) Titular(es): CANON KABUSHIKI KAISHA.

(72) Inventor(es): FRÉDÉRIC MAZE; HERVÉ LE FLOCH; FRANCK DENOUAL; CYRIL CONCOLATO; JEAN LE FEUVRE.

(86) Pedido PCT: PCT EP2014050952 de 17/01/2014

(87) Publicação PCT: WO 2014/111547 de 24/07/2014

(85) Data do Início da Fase Nacional: 17/07/2015

(57) Resumo: MÉTODO, DISPOSITIVO E PROGRAMA DE COMPUTADOR PARA ENCAPSULAR DADOS DE MÍDIA PROGRAMADA PARTICIONADOS. A invenção se refere a encapsular dados de mídia temporizada lado a lado em um servidor e fornecer um fluxo de bits de dados de mídia temporizada a partir de mídia temporizada lado a lado encapsulada em um dispositivo de cliente. Os dados de mídia temporizada lado a lado compreendem amostras temporizadas e cada amostra temporizada compreende uma pluralidade de subamostras espaciais. Após ter selecionado pelo menos uma subamostra espacial dentre a pluralidade de subamostras espaciais de uma das amostras temporizadas, uma trilha de item lado a lado que compreende a subamostra espacial selecionada e uma subamostra espacial correspondente de cada uma dentre as outras amostras temporizadas é criada para cada subamostra espacial selecionada. Uma trilha de composição que compreende pelo menos um extrator que identifica pelo menos uma das trilhas lado a lado é criada. Cada trilha criada é encapsulada de modo independente em pelo menos um arquivo de segmento de mídia.

“MÉTODO E DISPOSITIVO PARA ENCAPSULAR DADOS DE MÍDIA
PROGRAMADA PARTICIONADOS, MEIO DE ARMAZENAMENTO E MEIO DE
GRAVAÇÃO”

CAMPO DA INVENÇÃO

[0001] A invenção refere-se em geral ao campo de encapsulação de dados de mídia temporizada, por exemplo, de acordo com o Formato de Arquivo de Mídia Base conforme definido pela organização de padronização de MPEG, para aprimorar a entrega de fluxo, em particular, em relação ao fluxo contínuo de HTTP (*Protocolo de Transferência de Hipertexto*) de regiões de usuário selecionado de interesse em fluxos de vídeo compactados. Mais particularmente, a invenção se refere a um método, dispositivo e instruções executáveis para encapsulação de um fluxo elementar contendo dados particionados como tiles espaciais permitindo fluxo contínuo eficaz de dados, em particular, de um ou mais tiles.

ANTECEDENTES DA INVENÇÃO

[0002] A codificação de vídeo é uma forma de transformação de uma série de imagens de vídeo em um fluxo de bits digitalizado compacto de modo que as imagens de vídeo possam ser transmitidas ou armazenadas. Um dispositivo de codificação é usado para codificar as imagens de vídeo, em que um dispositivo de decodificação associado está disponível para reconstruir o fluxo de bits para exibição e visualização. Uma meta geral é formar o fluxo de bits de modo a ter um tamanho menor do que as informações de vídeo original. Isso reduz vantajosamente a capacidade necessária de uma rede de transferência ou de um dispositivo de armazenamento, para transmitir ou armazenar o código de fluxo de bits. Para ser transmitido, um fluxo de bits de vídeo é, em geral, encapsulado de acordo com um protocolo de transmissão que adiciona tipicamente cabeçalhos e bits de inspeção.

[0003] Recentemente, o Grupo de Especialistas em Imagens com Movimento (MPEG) publicou um novo padrão para unificar e superar soluções de fluxo contínuo existentes através de HTTP (*Protocolo de Transferência de Hipertexto*). Esse novo padrão, chamado de “Fluxo contínuo adaptativo dinâmico sobre HTTP (DASH)”, destina-se a suportar um modelo de fluxo contínuo de mídia sobre HTTP com base

em servidores da web padrão, em que o serviço de informações (isto é, a seleção de dados de mídia para o fluxo e a adaptação dinâmica dos fluxos de bits para escolhas de usuário, condições de rede e capacidades de cliente) se baseia exclusivamente em dispositivos e escolhas de cliente.

[0004] Nesse modelo, organiza-se uma apresentação de mídia em segmentos de dados e em um manifesto chamado de “Descrição de Apresentação de Mídia (MPD)” que representa a organização de dados de mídia temporizada a ser apresentados. Em particular, um manifesto compreende identificadores de recurso para o uso na transferência por download de segmentos de dados e fornece o contexto para selecionar e combinar esses segmentos de dados para obter uma apresentação de mídia válida. Os identificadores de recurso são tipicamente HTTP-URLs (*Localizador de Recurso Uniforme*), possivelmente combinado com trilhas de byte. Com base em um manifesto, um dispositivo de cliente determina a qualquer momento quais segmentos de mídia devem ser transferidos por download de um servidor de dados de mídia de acordo com suas necessidades, suas capacidades (por exemplo, codecs suportados, tamanho de visor, taxa de quadro, nível de qualidade, etc.) e, dependendo de condições de rede (por exemplo, largura de banda disponível).

[0005] Além disso, a resolução de vídeo cresce continuamente, indo de uma definição padrão (SD) para alta definição (HD) e para definição ultra alta (por exemplo, 4K2K ou 8K4K, ou seja, vídeo que compreende imagens de 4.096 x 2.400 pixels ou 7.680 x 4.320 pixels). No entanto, nem todos os dispositivos de recebimento e de decodificação de vídeo têm recursos (por exemplo, largura de banda de acesso de rede ou CPU (*Unidade de Processamento Central*)) para acessar vídeo em resolução total, em particular, quando o vídeo for de definição ultra alta e nem todos os usuários precisam acessar tal vídeo. Em tal contexto, é particularmente vantajoso fornecer a capacidade de acessar apenas algumas Regiões de Interesse (ROIs), ou seja, acessar apenas subpartes espaciais de uma sequência de vídeo completa.

[0006] Um mecanismo conhecido para acessar subpartes espaciais de quadros que pertencem a um vídeo consiste na organização de cada quadro do vídeo como uma disposição de áreas espaciais independentemente decodificáveis chamadas, em

geral, de tiles. Alguns formatos de vídeo como SVC (*Codificação de Vídeo Escalonável*) ou HEVC (*Codificação de Vídeo de Alta Eficácia*) fornecem suporte para definição de tile. Uma ROI de usuário definido pode cobrir um ou diversos tiles contíguos.

[0007] Consequentemente, para fluxo contínuo de ROIs de usuário selecionado de acordo com protocolo HTTP, é importante fornecer encapsulação de dados de mídia temporizada de um fluxo de bits de vídeo codificado de forma a permitir acesso espacial a um ou mais tiles e permitir a combinação de tiles acessados.

[0008] Deve-se lembrar que os fluxos de bits de vídeo codificados são, em geral, construídos como um conjunto de amostras temporais contíguas que correspondem a quadros completos, em que as amostras temporais são organizadas como uma função da ordem de decodificação. Os formatos de arquivo são usados para encapsular e descrever tais fluxos de bits codificados.

[0009] Por uma questão de ilustração, o Formato de Arquivo de Mídia Base de Organização Internacional para Padronização (ISO BMFF) é um formato flexível e extensível bem conhecido que descreve fluxos de bits de dados de mídia temporizada codificado para armazenamento local ou para transmissão através de uma rede ou através de outro mecanismo de entrega de fluxo de bits. Esse formato de arquivo está orientado para objeto. É composto de blocos de construção chamados de caixas que são sequencial ou hierarquicamente organizadas e que definem parâmetros do fluxo de bits de dados de mídia temporizada codificado como parâmetros de temporização e de estrutura. De acordo com esse formato de arquivo, o fluxo de bits de dados de mídia temporizada está contido em uma estrutura de dados chamada de *caixa de mdat* que é definido em outra estrutura de dados chamada de *caixa de trilha*. A trilha representa uma sequência temporizada de amostras, em que uma amostra corresponde a todos os dados associados a um único carimbo de data e hora, ou seja, todos os dados associados a um único quadro ou todos os dados associados a vários quadros que compartilham o mesmo carimbo de data e hora.

[0010] Para o vídeo escalonável como o vídeo do formato SVC, uma organização de dados de mídia dispostos em camadas pode ser representada de maneira eficaz

por meio do uso de múltiplas trilhas dependentes, em que cada trilha representa o vídeo em um nível particular de escalabilidade. A fim de evitar a duplicação de dados entre trilhas, podem ser usados extratores. De acordo com um formato de arquivo padrão, um extrator é uma estrutura de dados diretamente incluída em um fluxo de bits que permite a extração eficaz de unidades de camada de abstração de rede (NAL) de outros fluxos de bits. Por exemplo, o fluxo de bits de uma trilha de camada de aprimoramento pode compreender extratores que fazem referência a unidades de NAL de uma trilha de camada de base. Então, posteriormente, quando tal trilha de camada de aprimoramento é extraída do formato de arquivo, os extratores precisam ser substituídos pelos dados aos quais os mesmos faziam referência.

[0011] Várias estratégias podem ser adotadas durante o uso de ISO BMFF que incorporam esses mecanismos para descrever subinformações e para facilitar o acesso a essas subinformações ou para organizar de maneira eficaz os fluxos de bits formando múltiplos segmentos.

[0012] Por exemplo, no artigo intitulado “Implications of the ISO Base Media File Format on Adaptive HTTP Streaming of H.264/SVC”, os autores, Kofler et al., apresentam três estratégias diferentes para organizar um fluxo de bits de vídeo escalonável (H264/SVC) para fluxo contínuo de HTTP considerando possibilidades bem como limitações do ISO BMFF:

- a) um único arquivo contendo um cabeçalho de arquivo particular que compreende um arquivo tipo caixa “ftyp” e um arquivo tipo filme “moov” contendo todos os metadados de ISO BMFF (incluindo definições de trilha), em que o único arquivo também compreende uma única *caixa de mdat* contendo a totalidade de fluxo de bits codificado. Essa organização é adequada para armazenamento local, mas não é adaptada para fluxo contínuo de HTTP onde um cliente pode precisar apenas de uma parte da totalidade de fluxo de bits;
- b) um único arquivo contendo múltiplas caixas de moov/mdat adequadas para fragmentação. Esse formato permite transferência por download progressiva. A caixa moov é equivalente à caixa moov em nível de

fragmento. De acordo com esse esquema, com o uso de um arquivo de mídia fragmentado, o fluxo de bits escalonável é dividido em múltiplas trilhas dependentes que representam o vídeo em diferentes níveis de escalabilidade. Os extratores são usados para fazer referência a unidades de NAL de outras trilhas. No caso de ser usada uma trilha por tile, todas as trilhas endereçáveis precisam ser preparadas antecipadamente e as trilhas não podem ser selecionadas independentemente. Se vários tiles devem ser exibidos, diversos fluxos de bits precisam ser decodificados e a camada de base é decodificada várias vezes;

- c) arquivos de múltiplos segmentos, em que cada arquivo é acessível por meio de sua própria URL e é transferível por download independentemente. Cada segmento consiste tipicamente em uma caixa tipo segmento (styp), que atua como um tipo de cabeçalho de arquivo, uma caixa de índice de segmento opcional (sidx) e um ou múltiplos fragmentos. Novamente, cada fragmento consiste em uma caixa moof e mdat. De acordo com esse esquema, com o uso de um arquivo de mídia fragmentado, cada trilha é armazenada em seu próprio segmento com o fluxo de bits associado relacionado a um nível de escalabilidade. Caso seja necessário, são usados extratores para fazer referência a fluxo de bits necessários de trilhas dependentes. Tal esquema de decodificação é particularmente adequado para fluxo contínuo de trilhas independentemente. O mesmo é bem adaptado para o padrão DASH, mas não é adequado para fluxo contínuo de tile visto que vários fluxos de bits devem ser decodificados e, dessa forma, é necessário um decodificador por trilha. Ademais, existe uma duplicação potencial do fluxo de bits da camada de base durante a seleção de mais de um tile.

[0013] Quando aplicadas a tiles espaciais, nenhuma dessas estratégias permite acesso eficaz a tiles específicos no contexto de fluxo contínuo de HTTP. De fato, com a definição de formato de arquivo existente, seria necessário acessar inúmeras trilhas de byte não contínuas em um fluxo de bits codificado ou isso iria resultar na duplicação

de fluxo de bits a fim de exibir tiles espaciais de vários quadros que correspondem a um dado intervalo de tempo.

[0014] Para resolver esses problemas, é fornecido um esquema de descrição de trilha e organização de dados eficaz adequado para tiles espaciais que assegura, independente da combinação de trilha selecionada por um aplicativo de cliente, que o resultado da análise de ISO BMFF sempre leva a um fluxo de bits elementar de vídeo válido para o decodificador de vídeo.

SUMÁRIO DA INVENÇÃO

[0015] Defrontando-se com essas restrições, os inventores fornecem um método e um dispositivo para encapsular dados de mídia temporizada em tiles em um servidor e para fornecer um fluxo de bits de dados de mídia temporizada de dados de mídia temporizada em tiles encapsulados em uma pluralidade de arquivos de segmento de mídia.

[0016] É um amplo objetivo da invenção reparar as deficiências da técnica anterior conforme descrito acima.

[0017] De acordo com um primeiro aspecto da invenção, é fornecido um método de encapsulação de dados de mídia temporizada particionados em um servidor, em que os dados de mídia temporizada particionados compreendem amostras temporizadas, em que cada amostra temporizada compreende uma pluralidade de subamostras, em que o método compreende:

[0018] selecionar pelo menos uma subamostra dentre a pluralidade de subamostras de uma das amostras temporizadas;

[0019] criar, para cada subamostra selecionada, uma trilha de partição que compreende a subamostra selecionada e uma subamostra correspondente de cada uma das outras amostras temporizadas;

[0020] criar pelo menos uma trilha de referência que compreende pelo menos um extrator que identifica pelo menos uma das trilhas de partição criadas; e

[0021] encapsular independentemente cada uma das trilhas criadas em pelo menos um arquivo de segmento de mídia.

[0022] Consequentemente, a invenção permite a combinação de porções

diferentes e, em particular, de diferentes tiles, e a criação de um formato de arquivo válido e fluxo elementar no instante de análise seja qual for o conjunto selecionado de porções ou tiles transferidos por download por um dispositivo de cliente.

[0023] Portanto, o método da invenção é adequado para fluxo contínuo eficaz de porções ou tiles independentes, exige que apenas dados úteis sejam enviados para um dispositivo de cliente, é adequado para mais de uma porção ou seleção de tile (isto é, ROI arbitrária), reduz a sobrecarga de indexação (em comparação a solicitações de trilha de byte) e pode ser integrado a um padrão de MPEG.

[0024] Em uma modalidade, os dados de mídia temporizada particionados são dados de mídia temporizada em tiles, em que as subamostras são subamostras espaciais, em que a pelo menos uma trilha de partição é pelo menos uma trilha de tile e em que a pelo menos uma trilha de referência é pelo menos uma trilha compósita. A pelo menos uma trilha de referência pode conter extratores e, opcionalmente, dados de inicialização.

[0025] Em uma modalidade, o pelo menos um extrator identifica adicionalmente pelo menos uma subamostra espacial da pelo menos uma trilha de tile identificada.

[0026] Em uma modalidade, o método compreende adicionalmente uma etapa de criação de um arquivo de segmento de inicialização que compreende parâmetros que permitem a descrição de uma porção espacial dos dados de mídia temporizada em tiles das trilhas de tile criadas e da pelo menos uma trilha compósita criada. O arquivo de segmento de inicialização pode compreender referências para as trilhas de tile criadas, uma lista de trilhas de tile referidas associadas a pelo menos uma trilha compósita criada, um manipulador de tile associado a cada uma das trilhas de tile criadas indicando que a trilha contém informações de subamostra espacial e/ou informações de apresentação geral associadas com cada uma das trilhas de tile criadas.

[0027] Em uma modalidade, pelo menos duas trilhas de tile são criadas na etapa de criação de uma trilha de tile par cada subamostra espacial selecionada, em que a pelo menos uma trilha compósita compreende pelo menos um item de dados que deve ser usado para decodificar qualquer uma de duas subamostras espaciais

compreendidas dentro da mesma amostra temporizada.

[0028] Em uma modalidade, os dados de mídia temporizada em tiles são escalonáveis dados de mídia temporizada em tiles de acordo com qualquer uma das amostras temporizadas compreendem uma camada de referência e pelo menos uma camada de aprimoramento, em que a camada de referência compreende pelo menos uma subamostra espacial de referência e a pelo menos uma camada de aprimoramento compreende uma pluralidade de subamostras espaciais de aprimoramento e em que as trilhas de tile criadas na etapa de criação de uma trilha de tile para cada subamostra espacial selecionada são trilhas de tile de aprimoramento que compreendem subamostras espaciais de aprimoramento, em que o método compreende adicionalmente uma etapa de criação de pelo menos uma trilha de referência que compreende subamostras espaciais de referência.

[0029] Em uma modalidade, a camada de referência compreende uma pluralidade de subamostras espaciais de referência e em que é criada uma pluralidade de trilha de referências, em que cada trilha de referência da pluralidade de trilha de referências forma uma trilha de tile de referência.

[0030] Em uma modalidade, os arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas compreendem uma estrutura de dados para armazenar metadados incluindo definições da trilha correspondente.

[0031] Em uma modalidade, pelo menos um dos arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas compreende uma estrutura de dados para armazenar metadados incluindo pelo menos uma definição de pelo menos um grupo de subamostras da trilha correspondente.

[0032] Em uma modalidade, pelo menos um dos arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas compreende uma estrutura de dados para armazenar metadados incluindo diferentes definições de diferentes grupos de subamostras da trilha correspondente.

[0033] Em uma modalidade, a estrutura de dados que armazena metadados inclui pelo menos uma definição e um tipo de versão indicando que pelo menos uma definição se aplica a todas as subamostras da trilha correspondente.

[0034] Em uma modalidade, pelo menos um parâmetro da estrutura de dados que armazena metadados incluindo pelo menos uma definição de pelo menos um grupo de subamostras da trilha correspondente, relacionado a dependências de decodificação de subamostra, indica que cada subamostra do grupo pode ser decodificada sem o uso de uma subamostra de outro grupo.

[0035] Em uma modalidade, as subamostras de um grupo são identificadas de acordo com um tipo de agrupamento ao qual as subamostras pertencem.

[0036] Em uma modalidade, as subamostras de um grupo são identificadas de acordo com um tipo de agrupamento ao qual as subamostras pertencem, em que um tipo de agrupamento fornece uma definição para cada subamostra do grupo.

[0037] Em uma modalidade, o arquivo de segmento de inicialização compreende uma definição a ser usada por padrão para processar subamostras de pelo menos uma trilha de tile que não são associadas a nenhuma definição específica armazenada no arquivo de segmento de mídia que resulta da encapsulação da pelo menos uma trilha de tile correspondente.

[0038] Em uma modalidade, o método compreende adicionalmente uma etapa de adição de dados de inicialização em pelo menos um dos arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas para permitir a modificação da definição armazenada no arquivo de segmento de inicialização que deve ser usado por padrão.

[0039] Em uma modalidade, os arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas compreendem uma estrutura de dados para armazenar subamostras espaciais. Os arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas compreendem uma estrutura de dados para armazenar extratores.

[0040] Em uma modalidade, pelo menos um dos arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas compreende dados de inicialização armazenados na trilha de referência de modo que o pelo menos um dos arquivos de segmento de mídia possa ser processado independentemente como um arquivo de segmento de mídia padrão.

[0041] Em uma modalidade, pelo menos um dos arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas compreende uma estrutura de dados para armazenar pelo menos um extrator que identifica dados de inicialização armazenados na trilha de referência.

[0042] Em uma modalidade, pelo menos um dos arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas compreende uma estrutura de dados para armazenar pelo menos um extrator que identifica dados de inicialização da trilha compósita de modo que o pelo menos um dos arquivos de segmento de mídia possa ser processado independentemente como um arquivo de segmento de mídia padrão.

[0043] Em uma modalidade, o método compreende adicionalmente uma etapa de definição de um sinalizador indicando que o pelo menos um dos arquivos de segmento de mídia pode ser processado independentemente como um arquivo de segmento de mídia padrão.

[0044] Em uma modalidade, o método compreende adicionalmente uma etapa de criação de pelo menos uma trilha de dados de inicialização que compreende dados de inicialização, em que a pelo menos uma trilha de dados de inicialização é independentemente encapsulada em pelo menos um arquivo de segmento de mídia.

[0045] Em uma modalidade, pelo menos um dos arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas compreende uma estrutura de dados para armazenar pelo menos um extrator que identifica dados de inicialização armazenados na trilha de dados de inicialização de modo que o pelo menos um dos arquivos de segmento de mídia possa ser processado independentemente como um arquivo de segmento de mídia padrão.

[0046] Em uma modalidade, o método compreende adicionalmente uma etapa de definição de um sinalizador indicando que o pelo menos um dos arquivos de segmento de mídia pode ser processado independentemente como um arquivo de segmento de mídia padrão.

[0047] Em uma modalidade, o servidor é compatível com o Protocolo de Transferência de Hipertexto (HTTP).

[0048] Em uma modalidade, os arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas são compatíveis com o Formato de Arquivo de Mídia Base e com o Fluxo Contínuo Adaptativo Dinâmico sobre formato HTTP conforme definido pela Organização Internacional para Padronização.

[0049] Em uma modalidade, o método de encapsulação de dados de mídia temporizada particionados em um servidor, em que os dados de mídia temporizada particionados compreendem amostras temporizadas, em que cada amostra temporizada compreende uma pluralidade de subamostras espaciais, compreende:

[0050] selecionar pelo menos uma subamostra dentre a pluralidade de subamostras de uma das amostras temporizadas;

[0051] criar, para cada subamostra espacial selecionada, uma trilha de partição que compreende a subamostra selecionada e uma subamostra correspondente de cada uma das outras amostras temporizadas;

[0052] criar pelo menos uma trilha de referência que compreende informações associadas a uma primeira subamostra espacial incluída em uma primeira trilha de partição e uma segunda subamostra espacial incluída em uma segunda trilha de partição; e

[0053] encapsular de modo independente cada trilha criada em pelo menos um arquivo de segmento de mídia.

[0054] Tal modalidade pode ser combinada com uma ou várias das modalidades anteriores reveladas acima a título de referência em relação ao primeiro aspecto da invenção.

[0055] Um segundo aspecto da invenção fornece um método de fornecimento de um fluxo de bits de dados de mídia temporizada de dados de mídia temporizada particionados encapsulados em uma pluralidade de arquivos de segmento de mídia, em um dispositivo de cliente, em que os dados de mídia temporizada particionados compreendem amostras temporizadas, em que cada amostra temporizada compreende uma pluralidade de subamostras, em que a pluralidade de arquivos de segmento de mídia compreende pelo menos uma trilha de referência e pelo menos uma trilha de partição, em que cada uma da pelo menos uma trilha de partição

compreende uma subamostra selecionada dentre a pluralidade de subamostras de uma das amostras temporizadas e compreende uma subamostra espacial correspondente de cada uma das outras amostras temporizadas, em que a pelo menos uma trilha de referência compreende pelo menos um extrator que identifica pelo menos uma trilha de partição e identifica pelo menos uma subamostra da trilha de partição identificada, em que o método compreende:

[0056] selecionar um item de informações representativo de uma subamostra;

[0057] solicitar a pelo menos uma trilha de partição que compreende a subamostra selecionada, em que a subamostra selecionada corresponde ao item de selecionado informações;

[0058] receber a pluralidade de arquivos de segmento de mídia; e

[0059] substituir extratores da pelo menos uma trilha de referência por subamostras identificadas correspondentes para gerar o fluxo de bits de dados de mídia temporizada.

[0060] Consequentemente, a invenção permite a combinação de diferentes porções e, em particular, de diferentes tiles, e a criação de um formato de arquivo válido e o fluxo elementar no instante de análise seja qual for o conjunto selecionado de porções ou tiles transferidos por download por um dispositivo de cliente.

[0061] Portanto, o método da invenção é adequado para fluxo contínuo eficaz de porções ou tiles independentes, exige que apenas dados úteis sejam enviados para um dispositivo de cliente, é adequado para mais de uma porção ou seleção de tile (isto é, ROI arbitrária), reduz a sobrecarga de indexação (em comparação a solicitações de trilha de byte) e pode ser integrado em um padrão de MPEG.

[0062] Em uma modalidade, os dados de mídia temporizada particionados são dados de mídia temporizada em tiles, em que as subamostras são subamostras espaciais, em que a pelo menos uma trilha de partição é pelo menos uma trilha de tile e em que a pelo menos uma trilha de referência é pelo menos uma trilha compósita.

[0063] Em uma modalidade, os dados de mídia temporizada particionados são dados de mídia temporizada em tiles, em que as subamostras são subamostras espaciais, em que a pelo menos uma trilha de partição é pelo menos uma trilha de tile

em que a pelo menos uma trilha de referência é pelo menos uma trilha compósita contendo extratores e, opcionalmente, dados de inicialização.

[0064] Em uma modalidade, o método compreende adicionalmente:

[0065] obter uma referência de trilha de um extrator da pelo menos uma trilha compósita;

[0066] verificar se a trilha que corresponde à referência de trilha obtida foi recebida ou não; e

[0067] remover, se a trilha que corresponde à referência de trilha obtida não foi recebida, o extrator correspondente para gerar o fluxo de bits de dados de mídia temporizada.

[0068] Em uma modalidade, o método compreende adicionalmente uma etapa de verificação se a trilha que corresponde à referência de trilha obtida é uma trilha do tipo tile.

[0069] Em uma modalidade, o método compreende adicionalmente:

[0070] obter uma referência de trilha de um extrator da pelo menos uma trilha compósita;

[0071] verificar se a trilha que corresponde à referência de trilha obtida foi recebida ou não; e

[0072] substituir, se a trilha que corresponde à referência de trilha obtida não foi recebida, o extrator correspondente por meio de preenchimento para gerar o fluxo de bits de dados de mídia temporizada.

[0073] Em uma modalidade, o método compreende adicionalmente uma etapa de verificação se a trilha que corresponde à referência de trilha obtida é uma trilha do tipo tile.

[0074] Em uma modalidade, o método compreende adicionalmente uma etapa de recebimento de um arquivo de segmento de inicialização que compreende parâmetros que permitem a formação de uma porção espacial dos dados de mídia temporizada em tiles das trilhas de tile criadas e a pelo menos uma trilha compósita criada.

[0075] Em uma modalidade, são recebidos pelo menos dois arquivos de segmento de mídia representativos de pelo menos duas trilhas de tile, em que a pelo menos uma

trilha compósita compreende pelo menos um item de dados que deve ser usado para decodificar qualquer uma de duas subamostras espaciais que pertencem ao mesmo conjunto de subamostras espaciais.

[0076] Em uma modalidade, os dados de mídia temporizada em tiles são dados de mídia temporizada em tiles escalonáveis de acordo com cada uma das amostras temporizadas compreendem uma camada de referência e pelo menos uma camada de aprimoramento, em que a camada de referência compreende pelo menos uma subamostra espacial de referência e em que a pelo menos uma camada de aprimoramento compreende uma pluralidade de subamostras espaciais de aprimoramento e em que pelo menos dois arquivos de segmento de mídia representativos de pelo menos duas trilhas são recebidos, em que uma das pelo menos duas trilhas recebidas é uma trilha de referência que compreende subamostras espaciais de referência e a outra das pelo menos duas trilhas recebidas é uma trilha de tile de aprimoramento que compreende subamostras espaciais de aprimoramento.

[0077] Em uma modalidade, a camada de referência compreende uma pluralidade de subamostras espaciais de referência e em que uma pluralidade de trilha de referências, em que cada uma forma uma trilha de tile de referência, é recebida.

[0078] Em uma modalidade, pelo menos um dos arquivos de segmento de mídia compreende uma estrutura de dados para armazenar metadados incluindo pelo menos uma definição de pelo menos um grupo de subamostras da trilha correspondente, em que o método compreende uma etapa de obtenção da pelo menos uma definição para processar o pelo menos um grupo de subamostras da trilha correspondente.

[0079] Em uma modalidade, pelo menos um dos arquivos de segmento de mídia compreende dados de inicialização armazenados na trilha de referência, em que o método compreende adicionalmente uma etapa de processamento independentemente do pelo menos um dos arquivos de segmento de mídia como um arquivo de segmento de mídia padrão.

[0080] Em uma modalidade, pelo menos um dos arquivos de segmento de mídia compreende uma estrutura de dados para armazenar pelo menos um extrator que

identifica dados de inicialização armazenados na trilha de referência, em que o método compreende adicionalmente uma etapa de acesso dos dados armazenados na trilha de referência durante o processamento de pelo menos um dos arquivos de segmento de mídia.

[0081] Em uma modalidade, pelo menos um dos arquivos de segmento de mídia compreende uma estrutura de dados para armazenar pelo menos um extrator que identifica dados de inicialização da trilha compósita, em que o método compreende adicionalmente uma etapa de acesso dos dados armazenados na trilha de referência para processar de modo independente o pelo menos um dos arquivos de segmento de mídia como um arquivo de segmento de mídia padrão.

[0082] Em uma modalidade, o método compreende adicionalmente uma etapa de obtenção de pelo menos uma trilha de dados de inicialização que compreende dados de inicialização, em que pelo menos uma trilha de dados de inicialização é independentemente encapsulada em pelo menos um arquivo de segmento de mídia.

[0083] Em uma modalidade, pelo menos um dos arquivos de segmento de mídia compreende uma estrutura de dados para armazenar pelo menos um extrator que identifica dados de inicialização armazenados na trilha de dados de inicialização, em que o método compreende adicionalmente uma etapa de acesso dos dados armazenados na trilha de referência para processar de modo independente o pelo menos um dos arquivos de segmento de mídia.

[0084] Em uma modalidade, o dispositivo de cliente é compatível com o Protocolo de Transferência de Hipertexto (HTTP).

[0085] Em uma modalidade, os arquivos de segmento de mídia recebidos são compatíveis com o Formato de Arquivo de Mídia Base e com o Fluxo Contínuo Adaptativo Dinâmico sobre formato HTTP conforme definido pela Organização Internacional para Padronização.

[0086] De acordo com um terceiro aspecto da invenção, é fornecido um dispositivo que compreende meios adaptados para executar cada etapa do método descrito acima.

[0087] De acordo com um quarto aspecto da invenção, é fornecido um dispositivo

para encapsular dados de mídia temporizada particionados em um servidor, em que os dados de mídia temporizada particionados compreendem amostras temporizadas, em que cada amostra temporizada compreende uma pluralidade de subamostras, em que o dispositivo compreende pelo menos um microprocessador configurado para executar as etapas a seguir:

[0088] selecionar pelo menos uma subamostra dentre a pluralidade de subamostras de uma das amostras temporizadas;

[0089] criar, para cada subamostra selecionada, uma trilha de partição que compreende a subamostra selecionada e uma subamostra correspondente de cada uma das outras amostras temporizadas;

[0090] criar pelo menos uma trilha de referência que compreende pelo menos um extrator que identifica pelo menos uma das trilhas de partição criadas; e

[0091] encapsular de modo independente cada uma das trilhas criadas em pelo menos um arquivo de segmento de mídia.

[0092] Consequentemente, a invenção permite a combinação de diferentes porções e, em particular, de diferentes tiles, e a criação de um formato de arquivo válido e fluxo elementar em instante de análise seja qual for o conjunto selecionado de porções ou tiles transferidos por download por um dispositivo de cliente.

[0093] Portanto, o método da invenção é adequado para fluxo contínuo eficaz de porções ou tiles independentes, exige que apenas dados úteis sejam enviados para um dispositivo de cliente, é adequado para mais de uma porção ou seleção de tile (isto é, ROI arbitrária), reduz a sobrecarga de indexação (em comparação a solicitações de trilha de byte) e pode ser integrado a um padrão de MPEG.

[0094] Em uma modalidade, os dados de mídia temporizada particionados são dados de mídia temporizada em tiles, em que as subamostras são subamostras espaciais, em que a pelo menos uma trilha de partição é pelo menos uma trilha de tile e em que a pelo menos uma trilha de referência é pelo menos uma trilha compósita.

[0095] Em uma modalidade, os dados de mídia temporizada particionados são dados de mídia temporizada em tiles, em que as subamostras são subamostras espaciais, em que a pelo menos uma trilha de partição é pelo menos uma trilha de tile

e em que a pelo menos uma trilha de referência é pelo menos uma trilha compósita contendo extratores e, opcionalmente, dados de inicialização.

[0096] Em uma modalidade, o microprocessador é adicionalmente configurado para executar a etapa de criação de um arquivo de segmento de inicialização que compreende parâmetros que permitem a formação de uma porção espacial dos dados de mídia temporizada em tiles das trilhas de tile criadas e da pelo menos uma trilha compósita criada, em que o arquivo de segmento de inicialização compreende referências às trilhas de tile criadas.

[0097] Em uma modalidade, o arquivo de segmento de inicialização compreende adicionalmente uma lista de trilhas de tile referidas associada a pelo menos uma trilha compósita criada.

[0098] Em uma modalidade, o arquivo de segmento de inicialização compreende adicionalmente um manipulador de tile associado a cada uma das trilhas de tile criadas indicando que a trilha contém informações de subamostra espacial.

[0099] Em uma modalidade, o arquivo de segmento de inicialização compreende adicionalmente informações de apresentação geral associadas a cada uma das trilhas de tile criadas.

[0100] Em uma modalidade, o microprocessador é adicionalmente configurado de modo que pelo menos duas trilhas de tile sejam criadas na etapa de criação de uma trilha de tile para cada subamostra espacial selecionada, em que a pelo menos uma trilha compósita compreende pelo menos um item de dados que deve ser usado para decodificar qualquer uma de duas subamostras espaciais compreendidas dentro da mesma amostra temporizada.

[0101] Em uma modalidade, os dados de mídia temporizada em tiles são dados de mídia temporizada em tiles escalonáveis de acordo com cada uma das amostras temporizadas compreende uma camada de referência e pelo menos uma camada de aprimoramento, em que a camada de referência compreende pelo menos uma subamostra espacial de referência e em que a pelo menos uma camada de aprimoramento compreende uma pluralidade de subamostras espaciais de aprimoramento, em que o microprocessador é adicionalmente configurado de modo

que as trilhas de tile criadas na etapa de criação de uma trilha de tile para cada subamostra espacial selecionada sejam trilhas de tile de aprimoramento que compreendem subamostras espaciais de aprimoramento, em que o microprocessador é adicionalmente configurado para executar a etapa de criação de pelo menos uma trilha de referência que compreende subamostras espaciais de referência.

[0102] Em uma modalidade, a camada de referência compreende uma pluralidade de subamostras espaciais de referência e em que o microprocessador é adicionalmente configurado de modo que uma pluralidade de trilha de referências seja criada, em que cada trilha de referência da pluralidade de trilha de referências forma uma trilha de tile de referência.

[0103] Em uma modalidade, os arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas compreendem uma estrutura de dados para armazenar metadados incluindo definições da trilha correspondente.

[0104] Em uma modalidade, os arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas compreendem uma estrutura de dados para armazenar subamostras espaciais.

[0105] Em uma modalidade, os arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas compreendem uma estrutura de dados para armazenar extratores.

[0106] Em uma modalidade, o servidor é compatível com o Protocolo de Transferência de Hipertexto (HTTP) e em que os arquivos de segmento de mídia que resultam da encapsulação de cada uma das trilhas criadas são compatíveis com o Formato de Arquivo de Mídia Base e com o Fluxo Contínuo Adaptativo Dinâmico sobre formato HTTP conforme definido pela Organização Internacional para Padronização.

[0107] Em uma modalidade, o dispositivo para encapsular dados de mídia temporizada particionados em um servidor, em que os dados de mídia temporizada particionados compreendem amostras temporizadas, em que cada amostra temporizada compreende uma pluralidade de subamostras espaciais, compreende:

[0108] uma unidade de seleção configurada para selecionar pelo menos uma subamostra dentre a pluralidade de subamostras de uma das amostras temporizadas;

[0109] uma primeira unidade de criação configurada para criar, para cada subamostra espacial selecionada, uma trilha de partição que compreende a subamostra selecionada e uma subamostra correspondente de cada uma das outras amostras temporizadas;

[0110] uma segunda unidade de criação configurada para criar pelo menos uma trilha de referência que compreende informações associadas a uma primeira subamostra espacial incluída em uma primeira trilha de partição e uma segunda subamostra espacial incluída em uma segunda trilha de partição; e

[0111] uma unidade de encapsulação configurada para encapsular de modo independente cada trilha criada em pelo menos um arquivo de segmento de mídia.

[0112] Tal modalidade pode ser combinada a uma ou várias das modalidades anteriores reveladas acima a título de referência em relação ao quarto aspecto da invenção.

[0113] De acordo com um quinto aspecto da invenção, é fornecido um decodificador de vídeo que compreende o dispositivo conforme descrito acima.

[0114] De acordo com um sexto aspecto da invenção, é fornecido um dispositivo para fornecer um fluxo de bits de dados de mídia temporizada de dados de mídia temporizada particionados encapsulados em uma pluralidade de arquivos de segmento de mídia, em um dispositivo de cliente, em que os dados de mídia temporizada particionados compreendem amostras temporizadas, em que cada amostra temporizada compreende uma pluralidade de subamostras, em que a pluralidade de arquivos de segmento de mídia compreende pelo menos uma trilha de referência e pelo menos uma trilha de partição, em que cada uma da pelo menos uma trilha de partição compreende uma subamostra selecionada dentre a pluralidade de subamostras de uma das amostras temporizadas e compreende uma subamostra correspondente de cada uma das outras amostras temporizadas, em que a pelo menos uma trilha de referência compreende pelo menos um extrator que identifica pelo menos uma trilha de partição e identifica pelo menos uma subamostra da trilha de partição identificada, em que o dispositivo compreende pelo menos um microprocessador configurado para executar as etapas de:

- [0115]** selecionar um item de informações representativo de uma subamostra;
- [0116]** solicitar a pelo menos uma trilha de partição que compreende a subamostra selecionada, em que a subamostra selecionada corresponde ao item de selecionado informações;
- [0117]** receber a pluralidade de arquivos de segmento de mídia; e
- [0118]** substituir extratores da pelo menos uma trilha de referência por subamostras identificadas correspondentes para gerar o fluxo de bits de dados de mídia temporizada.
- [0119]** Consequentemente, a invenção permite a combinação de diferentes porções e, em particular, de diferentes tiles, e a criação de um formato de arquivo válido e fluxo elementar em instante de análise seja qual for o conjunto selecionado de porções ou tiles transferidos por download por um dispositivo de cliente.
- [0120]** Portanto, o método da invenção é adequado para fluxo contínuo eficaz de porções ou tiles independentes, exige que apenas dados úteis sejam enviados para um dispositivo de cliente, é adequado para mais de uma porção ou seleção de tile (isto é, ROI arbitrária), reduz a sobrecarga de indexação (em comparação a solicitações de trilha de byte) e pode ser integrado a um padrão de MPEG.
- [0121]** Em uma modalidade, os dados de mídia temporizada particionados são dados de mídia temporizada em tiles, em que as subamostras são subamostras espaciais, em que a pelo menos uma trilha de partição é pelo menos uma trilha de tile e em que a pelo menos uma trilha de referência é pelo menos uma trilha compósita.
- [0122]** Em uma modalidade, os dados de mídia temporizada particionados são dados de mídia temporizada em tiles, em que as subamostras são subamostras espaciais, em que a pelo menos uma trilha de partição é pelo menos uma trilha de tile e em que a pelo menos uma trilha de referência é pelo menos uma trilha compósita contendo extratores e, opcionalmente, dados de inicialização.
- [0123]** Em uma modalidade, o microprocessador é adicionalmente configurado para executar as etapas a seguir:
- [0124]** obter uma referência de trilha de um extrator da pelo menos uma trilha compósita;

[0125] verificar se a trilha que corresponde à referência de trilha obtida foi recebida ou não; e

[0126] remover, se a trilha que corresponde à referência de trilha obtida não foi recebida, o extrator correspondente para gerar o fluxo de bits de dados de mídia temporizada.

[0127] Em uma modalidade, o microprocessador é adicionalmente configurado para executar a etapa de verificação se a trilha que corresponde à referência de trilha obtida é uma trilha do tipo tile.

[0128] Em uma modalidade, o microprocessador é adicionalmente configurado para executar as etapas a seguir:

[0129] obter uma referência de trilha de um extrator da pelo menos uma trilha compósita;

[0130] verificar se a trilha que corresponde à referência de trilha obtida foi recebida ou não; e

[0131] substituir, se a trilha que corresponde à referência de trilha obtida não foi recebida, o extrator correspondente por preenchimento para gerar o fluxo de bits de dados de mídia temporizada.

[0132] Em uma modalidade, o microprocessador é adicionalmente configurado para executar a etapa de verificação se a trilha que corresponde à referência de trilha obtida é uma trilha do tipo tile.

[0133] Em uma modalidade, o microprocessador é adicionalmente configurado para executar a etapa de recebimento de um arquivo de segmento de inicialização que compreende parâmetros que permitem a formação de uma porção espacial dos dados de mídia temporizada em tiles das trilhas de tile criadas e da pelo menos uma trilha compósita criada.

[0134] Em uma modalidade, o microprocessador é adicionalmente configurado de modo que pelo menos dois arquivos de segmento de mídia representativos de pelo menos duas trilhas de tile sejam recebidos, em que a pelo menos uma trilha compósita compreende pelo menos um item de dados que deve ser usado para decodificar qualquer uma de duas subamostras espaciais que pertencem ao mesmo conjunto de

subamostras espaciais.

[0135] Em uma modalidade, os dados de mídia temporizada em tiles são dados de mídia temporizada em tiles escalonáveis de acordo com cada uma das amostras temporizadas compreendem uma camada de referência e pelo menos uma camada de aprimoramento, em que a camada de referência compreende pelo menos uma subamostra espacial de referência e em que a pelo menos uma camada de aprimoramento compreende uma pluralidade de subamostras espaciais de aprimoramento e em que o microprocessador é adicionalmente configurado de modo que pelo menos dois arquivos de segmento de mídia representativos de pelo menos duas trilhas sejam recebidos, em que uma das pelo menos duas trilhas recebidas é uma trilha de referência que compreende subamostras espaciais de referência e a outra das pelo menos duas trilhas recebidas é uma trilha de tile de aprimoramento que compreende subamostras espaciais de aprimoramento.

[0136] Em uma modalidade, o dispositivo de cliente é compatível com o Protocolo de Transferência de Hipertexto (HTTP) e em que os arquivos de segmento de mídia recebidos são compatíveis com o Formato de Arquivo de Mídia Base e com o Fluxo Contínuo Adaptativo Dinâmico sobre formato HTTP conforme definido pela Organização Internacional para Padronização.

[0137] De acordo com um sétimo aspecto da invenção, é fornecido um decodificador de vídeo que compreende o dispositivo conforme descrito acima.

[0138] Visto que a presente invenção pode ser implantada em software, a presente invenção pode ser concretizada como código legível por computador para fornecimento para um aparelho programável em qualquer meio portador adequado. Um meio portador tangível pode compreender um meio de armazenamento como um disquete, um CD-ROM, uma unidade de disco rígido, um dispositivo de fita magnética ou um dispositivo de memória e estado sólido e similares. Um meio portador transiente pode incluir um sinal como um sinal elétrico, um sinal eletrônico, um sinal óptico, um sinal acústico, um sinal magnético ou um sinal eletromagnético, por exemplo, um sinal de micro-ondas ou de RF.

BREVE DESCRIÇÃO DOS DESENHOS

[0139] As vantagens adicionais da presente invenção se tornarão evidentes para o versado na técnica mediante o exame dos desenhos e da descrição detalhada.

Prevê-se que qualquer vantagem adicional seja incorporada ao presente documento.

[0140] As modalidades da invenção serão descritas agora, apenas a título de exemplo, e com referência aos desenhos a seguir, em que:

[0141] A Figura 1, que compreende As Figuras 1a e 1b, ilustra um exemplo de codificação de tiles em um fluxo de bits de vídeo codificado;

[0142] A Figura 2 ilustra uma conexão temporal de tiles selecionados por um usuário a serem exibidos;

[0143] A Figura 3 ilustra um exemplo de diagramas de bloco de um arquivo de segmento de inicialização com o uso de trilhas de dados de mídia escalonáveis de acordo com uma modalidade particular;

[0144] A Figura 4 ilustra um exemplo de diagramas de bloco de arquivos de segmento de mídia que compreendem trilhas de tile e uma trilha compósita, de acordo com uma modalidade particular, que corresponde a trilhas declaradas em um arquivo de segmento de inicialização conforme aquele mostrado na Figura 3;

[0145] A Figura 5 ilustra um exemplo de segmentos de dados de mídia transferidos por download em concatenação para construir um fluxo de bits de dados de mídia temporizada decodificável válido que representa uma parte espacial de quadros de vídeo consecutivos por um dado período de tempo;

[0146] A Figura 6, que compreende as Figuras 6a e 6b, é um fluxograma que ilustra etapas para transmitir dados de mídia temporizada entre um servidor e um dispositivo de cliente de acordo com uma modalidade particular;

[0147] A Figura 7 é um fluxograma que ilustra a geração de um fluxo de bits de dados de mídia temporizada válido de segmentos de mídia concatenados recebidos por um dispositivo de cliente; e

[0148] A Figura 8 representa um diagrama de bloco de um servidor ou um dispositivo de cliente em que as etapas de uma ou mais modalidades podem ser implantadas;

[0149] A Figura 9, que compreende as Figuras 9a, 9b e 9c, ilustra exemplos de

tiles e segmentos de fatia em um fluxo de bits de HEVC;

[0150] A Figura 10 ilustra um exemplo de encapsulação de um fluxo de bits de HEVC como um conjunto de trilhas que compreende uma trilha compósita e trilhas de tile independentes, de acordo com uma modalidade da invenção;

[0151] A Figura 11 ilustra um exemplo de diagramas de bloco de um arquivo de segmento de inicialização com o uso de trilhas de dados de mídia, de acordo com uma primeira modalidade, para encapsular um fluxo de bits de HEVC;

[0152] A Figura 12, que compreende as Figuras 12a e 12b, ilustra a sinalização da posição de um tile em um vídeo completo, do tamanho do tile e da indicação que a trilha de tile pode ser decodificada sem nenhum artefato, em nível de subamostra, adaptado para manipular várias configurações de tile;

[0153] A Figura 13 ilustra um exemplo de encapsulação de um fluxo de bits de HEVC as um conjunto de trilhas que compreende uma trilha compósita e trilhas de tile independentes que são reproduzíveis como trilhas de vídeo padrão, de acordo com uma modalidade da invenção; e

[0154] A Figura 14 ilustra um exemplo de encapsulação de um fluxo de bits de HEVC como um conjunto de trilhas que compreende uma trilha compósita, uma trilha de dados de inicialização e trilhas de tile independentes que são reproduzíveis como trilhas de vídeo padrão, de acordo com outra modalidade da invenção.

DESCRIÇÃO DETALHADA DE MODALIDADES DA INVENÇÃO

[0155] De acordo com uma modalidade particular, os dados de mídia temporizada particionados como os dados de mídia temporizada em tiles (por exemplo, dados de vídeo) que compreendem amostras temporizadas (por exemplo, imagens) são transmitidos como um conjunto de diversas trilhas de dados de mídia temporizada, tipicamente uma trilha de camada de base e diversas trilhas de tile, e uma referência ou trilha compósita que compreende referências em relação a trilhas de dados de mídia temporizada. Cada trilha de tile compreende uma subamostra espacial (por exemplo, diversas unidades de NAL) de diversas amostras temporizadas. Um tipo de extrator estendido é definido para fazer referência às trilhas de dados de mídia temporizada de uma trilha compósita. As trilhas de dados de mídia temporizada são

marcadas como não exibíveis e transportam e descrevem dados de mídia temporizada para tiles. Tal conjunto de trilhas de dados de mídia temporizada e uma trilha compósita permitem a seleção, composição e o fluxo contínuo de tiles de vídeo espacial. Cada trilha pode ser transmitida de um dispositivo de servidor para um dispositivo de cliente como um conjunto de arquivos de segmento de mídia. Um arquivo de segmento de inicialização pode ser usado para transmitir metadados necessários para decodificar arquivos de segmento de mídia.

[0156] A Figura 1, que compreende as Figuras 1a e 1b, ilustra um exemplo de tiles de codificação em um fluxo de bits de vídeo codificado.

[0157] Por uma questão de ilustração, considera-se, na seguinte descrição, que cada quadro de vídeo (amostra temporizada) é composto de tiles decodificáveis de modo independente que correspondem a subpartes espaciais (subamostras espaciais) do quadro de vídeo. O vídeo é, de preferência, escalonável e organizado em diferentes níveis de escalabilidade. Conforme ilustrado na Figura 1 a, um quadro de vídeo 100 pode compreender uma camada de base de HD (102) e uma camada de aprimoramento de 4K2K (104). Ainda por uma questão de ilustração, a camada de aprimoramento 104 pode ser dividida em quatro tiles regulares denotados a, b, c e d. Deve ser observado que tiles de diferentes formatos podem ser manipulados. Da mesma maneira, a camada de base 102 pode ser dividida em diversos tiles. Nesse caso, podem ser usadas diversas trilhas compósitas, por exemplo, uma da camada de base e uma para as camadas de aprimoramento ou para cada uma das camadas de aprimoramento.

[0158] Também deve ser observado que a invenção não se limita a formato de vídeo escalonável. Pode ser aplicada a todos os formatos de vídeo, permitindo que os tiles sejam decodificados de modo independente. Consequentemente, qualquer algoritmo de compactação de vídeo como MPEG4, AVC, HEVC, SVC ou SHVC futuro pode ser usado em conjunto com uma modalidade da invenção.

[0159] A Figura 1b representa um fluxo de bits de vídeo codificado típico em uma ordem de decodificação. Conforme ilustrado, o fluxo de bits de vídeo codificado compreende, no presente documento, três quadros de vídeo (110, 112 e 114)

codificados em uma ordem temporal. Cada quadro de vídeo compreende todas as unidades de camada de abstração de rede (NAL) da camada de base (BL) seguidas pelas unidades de NAL da camada de aprimoramento. Por exemplo, as unidades de NAL (1BL, 116) da camada de base (102-1) do primeiro quadro de vídeo (110) são seguidas pelas unidades de NAL (1 comum, 1a, 1b, 1c, 1d, 118) da camada de aprimoramento (104-1) do primeiro quadro de vídeo.

[0160] A parte do fluxo de bits de vídeo que corresponde a uma camada de aprimoramento com tiles espaciais é composta de unidades de NAL de cada tile. Opcionalmente, também pode conter unidades de NAL que são comuns para todos os tiles e que são exigidas para decodificar qualquer um dos tiles. As unidades de NAL que são comuns para todos os tiles de um dado quadro podem estar situadas em qualquer lugar na parte corresponde do fluxo de bits de vídeo (isto é, antes, entre ou após as unidades de NAL dos tiles do quadro de vídeo).

[0161] Conforme ilustrado, a parte do fluxo de bits de vídeo que corresponde à camada de aprimoramento do primeiro quadro de vídeo (110), que compreende tiles espaciais a, b, c e d, é composta de unidades de NAL para cada tile (1a, 1b, 1c e 1d) e de unidades de NAL (1 comum) que são comuns para todos os tiles a, b, c e d.

[0162] A Figura 2 ilustra uma conexão temporal de tiles selecionados por um usuário para serem exibidos. Mais precisamente, a Figura 2 representa um primeiro quadro de vídeo n e um segundo quadro de vídeo $n+m$ (em que n e m são valores de número inteiro), em que cada um dentre o primeiro e o segundo quadros de vídeo compreende doze tiles numerados 1 a 12. Dentre esses doze tiles, apenas o terceiro e o sétimo devem ser exibidos (conforme denotado com linhas em negrito). Os quadros de vídeo n e $n+m$ pertencem a uma série de quadros consecutivos que correspondem a um dado período de tempo. Portanto, o terceiro e o sétimo tiles de cada quadro do quadro n ao quadro $n+m$ são exibidos consecutivamente.

[0163] Conforme ilustrado na Figura 1, os dados de um fluxo de bits de vídeo são organizados como amostras temporais que correspondem a quadros completos. Consequentemente, é necessário acessar várias trilhas de byte pequenas para cada quadro quando áreas espaciais particulares desses quadros devem ser acessadas

durante um dado período de tempo conforme descrito acima com referência à Figura 2. Isso é ineficaz no fluxo contínuo de HTTP em termos do número de solicitações geradas e em termos de sobrecarga de dados.

[0164] Portanto, para fornecer um acesso mais eficaz em vídeos compactados para fluxo contínuo de ROI, o fluxo de bits de dados de mídia temporizada deve ser reorganizado de modo que os dados de um tile particular sejam organizados como uma trilha de byte contígua (formar uma conexão) para um dado período de tempo (isto é, um conjunto de quadros consecutivos).

[0165] Consequentemente, quando apenas uma subparte espacial de quadros de vídeo deve ser exibida, apenas conexões de tiles que correspondem à área espacial selecionada precisam ser transferidas por download (por exemplo, tiles 3 e 7 na Figura 2) com o uso de uma solicitação de HTTP por conexão e por período de tempo.

[0166] A Figura 3 ilustra um exemplo de diagramas de bloco de um arquivo de segmento de inicialização com o uso de trilhas de dados de mídia de acordo com uma modalidade particular.

[0167] O arquivo de segmento de inicialização ilustrado na Figura 3 e os arquivos de segmento de mídia ilustrados na Figura 4 ilustram uma modalidade da invenção para reorganizar e encapsular um fluxo de bits de vídeo em ISO Formato de Arquivo de Mídia Base enquanto diversas trilhas são transmitidas de modo contínuo e independente. Por uma questão de ilustração, o conteúdo das Figuras 3 e 4 se baseiam no fluxo de bits de vídeo ilustrado na Figura 1.

[0168] Para reorganizar e encapsular um fluxo de bits de vídeo em ISO BMFF enquanto diversas trilhas são transmitidas de modo contínuo e independente, um novo tipo de trilha chamado de trilha de tile é definido. Uma trilha de tile é uma trilha que é, por definição, uma sequência temporizada de amostras relacionadas, em que uma amostra representa todos os dados associados a um único carimbo de data e hora. Ao contrário de uma trilha de mídia de vídeo conhecida, em que uma amostra é tipicamente um quadro de vídeo individual, uma amostra de uma trilha de tile define uma subparte espacialmente definida de um quadro de vídeo completo. Consequentemente, as trilhas de tile contêm apenas unidades de NAL relacionadas a

um dado tile. Dessa forma, é possível criar conexões de tiles com trilhas de byte contíguas mediante o armazenamento de cada trilha em tiles de segmento independentes.

[0169] No entanto, visto que uma trilha de tile não contém todas as unidades de NAL e informações necessárias a ser decodificadas e reproduzidas, uma trilha adicional chamada de “trilha compósita” é usada. De acordo com uma modalidade particular, uma trilha compósita é uma trilha de mídia que representa um quadro em tiles completo (isto é, a composição de todos os tiles). Uma trilha compósita é a representação de um fluxo de bits de dados de mídia temporizada codificado em sua organização típica conforme descrito com referência à Figura 1 (quadro a quadro em ordem de decodificação). A trilha compósita usa objetos de extrator para fazer referência a unidades de NAL em suas respectivas trilhas de tile. Além disso, pode conter unidades de NAL que são comuns para todos os tiles se existirem tais unidades de NAL.

[0170] De acordo com uma modalidade particular, um arquivo de segmento de inicialização é usado para transmitir todos os metadados que são necessários para definir fluxos de bits de dados de mídia temporizada encapsulados em outros arquivos de segmento de mídia. Conforme ilustrado na Figura 3, o arquivo de segmento de inicialização 300 contém uma caixa tipo tile “ftyp” 302 e uma caixa de filme “moov” 304. A caixa de tipo tile 302 identifica, de preferência, quais especificações de ISO BMF os tiles de segmento cumprem e indica um número de versão daquela especificação. A caixa de filme “moov” 304 fornece todos os metadados que descrevem a apresentação armazenada em arquivos de segmento de mídia e, em particular, todas as trilhas disponíveis na apresentação.

[0171] A caixa de filme “moov” 304 contém uma definição para cada uma das trilhas (caixas de “trilha” 306-1 a 306-6) que corresponde ao fluxo de bits de vídeo escalonável fornecido como um exemplo na Figura 1.

[0172] A caixa de trilha 306-1 representa a camada de base (track_ID=1), as quatro caixas de trilha 306-2 a 306-5 (caixas de trilha 306-3 e 306-4 não são mostradas) representam os quatro tiles a, b, c e d da camada de aprimoramento

(track_ID=2 a 5) e a caixa de trilha 306-6 representa uma trilha compósita que descreve a camada de aprimoramento (track_ID=6).

[0173] Cada caixa de trilha contém pelo menos uma caixa de cabeçalho de trilha “tkhd”, genericamente referida 308 e uma caixa de mídia de trilha “mdia” genericamente referida 310. Se uma trilha depender de dados de outras trilhas, há também uma caixa de referência de trilha “tref”. Conforme ilustrado, a trilha compósita que tem o identificador track_ID=6 compreende caixa de referência de trilha “tref” 312 indicando que a trilha depende de dados de trilhas que têm os identificadores track_ID=1 a 6.

[0174] Deve ser observado que outras caixas podem ser mandatórias ou opcionais dependendo de especificações de ISO BMFF usadas para encapsular o fluxo de bits de dados de mídia temporizada. No entanto, visto que as modalidades da invenção não dependem dessas caixas para serem aplicáveis, as mesmas não são apresentadas no presente documento.

[0175] A caixa de cabeçalho de trilha “tkhd” 308 especifica as características da trilha. Dentre diversos itens de informações, a mesma fornece o identificador da trilha (track_ID), a duração da trilha e/ou o tamanho de apresentação visual da trilha isto é, largura e altura da área de exibição). A mesma também compreende um parâmetro sinalizador que indica se a trilha é, ou não, reproduzível.

[0176] De acordo com uma modalidade, o valor padrão do sinalizador de cabeçalho de trilha para trilhas de tile é 0 (track_enabled = 0, track_in_movie = 0, track_in_preview = 0), significando que as trilhas de tile são ignoradas para reprodução e pré-visualização local e por um dispositivo de cliente. Em outra modalidade, pode ser criado um novo sinalizador de cabeçalho de trilha para sinalizar que a trilha é uma trilha de tile.

[0177] A caixa de mídia de trilha “mdia” 310 pode ser vista como um recipiente contendo todos os objetos usados para declarar parâmetros de dados de mídia temporizada dentro de uma trilha. A mesma contém pelo menos uma caixa de cabeçalho de mídia “mdhd” genericamente referida 314, uma caixa de referência de manipulador “hdlr” genericamente referida 316 e uma caixa de informações de mídia

“minf” genericamente referida 318.

[0178] A caixa de referência de manipulador “hdlr” 316 declara o processo por meio do qual os dados de mídia temporizada da trilha devem ser apresentados e, dessa forma, a natureza dos dados de mídia temporizada na trilha. Por exemplo, uma trilha de vídeo seria manipulada por um manipulador de vídeo (registrado com o atributo tipo manipulador igual a “vide”). Uma amostra de vídeo pode ser descrita por meio do uso de um objeto do tipo `VisualSampleEntry()`. De acordo com uma modalidade particular, um novo tipo de manipulador, chamado de manipulador de tile (registrado com o atributo tipo manipulador igual a “tile”), é definido para indicar que a trilha contém informações de subamostra espacial. Dependendo do formato de codificação, se um objeto do tipo `VisualSampleEntry()` não pode descrever uma amostra em uma trilha de tile, é possível definir um objeto específico do tipo `TileSampleEntry()` para descrever a amostra.

[0179] A caixa de informações de mídia “minf” 318 contém todos os objetos que definem informações características dos dados de mídia temporizada na trilha. Por exemplo, para a camada de base e para a camada de aprimoramento definida na trilha compósita, a caixa “minf” pode conter uma caixa de cabeçalho de mídia de vídeo padrão “vmhd” (320).

[0180] Em relação a trilhas de tile, uma caixa específica chamada de caixa de Cabeçalho de Mídia de Tile (tmhd, 322), que corresponde ao novo manipulador de tile, é usada para definir informações de apresentação geral, independente da codificação, para tiles. Em particular, pode conter as informações de geometria da área espacial coberta pelo tile em relação à resolução de vídeo representada pela trilha de vídeo compósita, que pode ser definida da seguinte maneira:

```
aligned(8) class TileMediaHeaderBox
extends FullBox('tmhd', version = 0, 0) {
    unsigned int(16) horizontal_offset;
    unsigned int( 16) vertical_offset;
}
```

[0181] Conforme descrito anteriormente, a trilha compósita compreende a caixa

de referência de trilha específica “tref” 312 que fornece referência tipada para outra trilha na apresentação. De acordo com uma modalidade particular, tais referências tipadas podem compreender uma referência de “tile” (324) que pode ser usada para estabelecer um link da trilha compósita para a trilha de tile a qual se refere e uma referência “scal” (326) que pode ser usada para estabelecer um link da trilha contendo essa referência com a pista de dados de mídia temporizada que a mesma depende (por exemplo, a trilha de camada de base (track_ID=1)).

[0182] A Figura 4 ilustra um exemplo de diagramas de bloco de arquivos de segmento de mídia que compreende trilhas de tile e uma trilha compósita, de acordo com uma modalidade particular, que corresponde a pistas declaradas em um arquivo de segmento de inicialização como aquele mostrado na Figura 3. Conforme descrito anteriormente, o arquivo de segmento de inicialização ilustrado na Figura 3 e os arquivos de segmento de mídia ilustrados na Figura 4 ilustram uma modalidade da invenção para reorganizar e encapsular um fluxo de bits de vídeo em Formato de Arquivo de Mídia Base ISO enquanto diversas trilhas são transmitidas de modo contínuo e independente.

[0183] Conforme ilustrado na Figura 4, cada um dos arquivos de segmento de mídia 400-1 a 400-6 (arquivos de segmento de mídia 400-3 a 400-5 não são mostrados) compreende, conforme especificado no padrão DASH, uma caixa tipo segmento “styp” genericamente referida 402, pelo menos uma caixa de fragmento de filme “moof” genericamente referida 404 e pelo menos uma caixa de dados de mídia “mdat” genericamente referida 406. Um arquivo de segmento de mídia é associado a uma HTTP-URL.

[0184] Deve ser observado que outras caixas podem ser mandatórias ou opcionais dependendo de especificações de ISO BMFF usadas para encapsular o fluxo de bits de dados de mídia temporizada. No entanto, visto que as modalidades da invenção não dependem dessas caixas para serem aplicáveis, as mesmas não são apresentadas no presente documento.

[0185] O formato da caixa tipo segmento “styp” 402 é similar àquele da caixa tipo tile “ftyp” 302 na Figura 3; contudo, sua referência indica que o arquivo é um arquivo

de segmento de mídia.

[0186] A caixa de fragmento de filme 404 fornece as informações que são geralmente armazenadas dentro da caixa de filme “moov”. Seu cabeçalho (“mfhd”) contém um número de sequência (marcado *seq_num* na Figura 4) que aumenta para cada fragmento de filme. Tal número de sequência permite que um dispositivo de cliente concatene tiles de segmento recebidos em ordem crescente e verifique a integridade da sequência (caso seja necessário). A caixa de fragmento de filme 404 contém uma caixa de fragmento de trilha “traf” (genericamente referida 408) para cada trilha que tem dados em uma caixa de dados de mídia associados (“mdat”, 406). A caixa de fragmento de trilha 408 compreende uma caixa de cabeçalho de fragmento de trilha “tfhd”, genericamente referida 410, que é usada para armazenar um identificador (track_ID) do fluxo de bits da trilha presente nos caixas de dados de mídia correspondentes (“mdat”, 406).

[0187] Uma caixa de dados de mídia contém, em geral, dados de mídia temporizada. Em trilhas de vídeo padrão, a mesma contém quadros de vídeo. Em trilhas de tile, a caixa de dados de mídia 406 contém subpartes espacialmente definidas de quadros de vídeo completos. Por uma questão de ilustração, a caixa de dados de mídia associada ao identificador de trilha track_ID=2 contém todas as unidades de NAL que correspondem ao tile da camada de aprimoramento.

[0188] Em uma trilha compósita (track_ID=6 na Figura 4), a caixa de dados de mídia 406 contém extratores (marcados E na Figura 4) para cada tile e para cada camada dependente e contém unidades de NAL comuns para todos os tiles (se houver).

[0189] Conforme ilustrado na Figura 4, a caixa de dados de mídia 406 de arquivo de segmento de mídia 400-6 associado a uma trilha compósita compreende, em particular:

- um primeiro extrator 412-1 que fornece um link para dados de camada de base (unidades de NAL 1 BL) codificados dentro de uma trilha de camada de base armazenada na caixa de dados de mídia 406 do arquivo de segmento de mídia 400-1 associado a uma trilha de camada de base;

- unidades de NAL 412-2 que são comuns para diversos tiles;
- um segundo extrator 412-3 que fornece um link para dados de camada de aprimoramento (unidades de NAL 1a) de um primeiro tile codificado dentro da caixa de dados de mídia 406 do arquivo de segmento de mídia 400-2 associado a uma primeira trilha de tile da camada de aprimoramento;
- um terceiro extrator 412-4 que fornece um link para dados de camada de aprimoramento (unidades de NAL, 1 b) de um segundo tile codificado dentro da caixa de dados de mídia 406 do arquivo de segmento de mídia 400-3 (não mostrado) associado a uma segunda trilha de tile da camada de aprimoramento;
- um quarto extrator 412-5 que fornece um link para dados de camada de aprimoramento (unidades de NAL, 1c) de um terceiro tile codificado dentro da caixa de dados de mídia 406 do arquivo de segmento de mídia 400-4 (não mostrado) associado a uma terceira trilha de tile da camada de aprimoramento; e
- um quinto extrator 412-6 que fornece um link para dados de camada de aprimoramento (unidades de NAL, 1d) de um quarto tile codificado dentro da caixa de dados de mídia 406 do arquivo de segmento de mídia 400-5 (não mostrado) associado a uma quarta trilha de tile da camada de aprimoramento.

[0190] As unidades de NAL que podem ser obtidas graças ao extrator 412-1 permitem a decodificação de uma camada de base de um quadro cuja camada de aprimoramento pode ser totalmente decodificada com o uso de unidades de NAL 412-2 e de unidades de NAL que podem ser obtidas graças aos extratores 412-3 a 412-6. Conforme pode ser observado na Figura 4, se apenas uma parte espacial de um quadro deve ser decodificada, não é necessário transferir por download todos os arquivos de segmento de mídia 400-2 a 400-5 (isto é, fluxos de bits que correspondem a trilhas de tile).

[0191] De acordo com uma modalidade particular, um extrator é uma estrutura interna de formato de arquivo que tem a seguinte sintaxe:

```

class aligned(8) Extractor () {
    NALUnitHeaderQ;
    unsigned int(8) track_ref_index;
    signed int(8) sample_offset;
    unsigned int((lengthSizeMinusOne+ 1 ) *8)
    data_offset;
    unsigned int((lengthSizeMinusOne+ 1 ) *8)
    data_length;
}

```

[0192] em que NALUnitHeader() representa os quatro primeiros bytes de uma unidade de NAL em conformidade com o formato de codificação usado para codificar o fluxo de bits de vídeo. Esses quatro bytes identificam a unidade de NAL como um extrator (por exemplo, em SVC, o atributo attribute nal_unit_type é definido para o tipo de unidade de NAL de extrator (tipo 31)).

[0193] O valor track_ref_index especifica o índice a ser usado, na caixa de referência de trilha “tref” do tipo “scal” ou “tile” de uma trilha compósita, para encontrar a trilha a partir da qual os dados devem ser extraídos. O valor sample_offset proporciona o índice relativo da amostra na trilha ligada por link que deve ser usada como a fonte de informações. Os valores data_offset e data_length são as compensações do primeiro byte dentro da amostra de referência para copiar e o número de bytes para copiar, respectivamente.

[0194] Por questão de ilustração e com referência à Figura 3, se o valor de track_ref_index de um dado extrator for igual a 2, isso significa que o extrator se refere à trilha identificada pela segunda entrada na caixa tref (isto é, a trilha que tem o identificador track_ID=2, ou seja, a trilha de tile para o tile a, sendo que o primeiro índice representa a trilha de referência (por exemplo, camada de base)).

[0195] A Figura 5 ilustra um exemplo de segmentos de dados de mídia transferidos por download em concatenação para construir um fluxo de bits de dados de mídia temporizada decodificável válido que representa uma parte espacial de quadros de vídeo consecutivos para um dado período de tempo (que corresponde aqui a dois

quadros consecutivos). A mesma Figura poderia ser repetida para outros períodos de tempo.

[0196] Conforme descrito com referência às Figuras 3 e 4, um fluxo de bits de dados de mídia de tiles temporizados é preferencialmente transmitido como um conjunto de dados que compreende um arquivo de segmento de inicialização e uma pluralidade de arquivos de segmento de mídia, sendo que o último compreende diversas trilhas de tile e uma trilha compósita.

[0197] O arquivo de segmento de inicialização compreende uma caixa de filme ("moov") que fornece informações gerais em cada trilha, em particular, o tipo de trilha (por exemplo, trilha de mídia (áudio ou vídeo) ou trilha de tile), um formato de codificação, uma resolução de quadro e a dependência dentre as trilhas (dada em uma caixa de referência de trilha "tref"). Esses dados são usados para processar os arquivos de segmento de mídia transferidos por download. Com referência ao exemplo descrito com referência às Figuras 1, 3 e 4, o conteúdo da caixa de filme do arquivo de segmento de inicialização pode compreender, em particular, o seguinte:

MOOV

- *trilha 1: camada de base*
- *trilha 2: tile a*
- *trilha 3: tile b*
- *trilha 4: tile c*
- *trilha 5: tile d*
- *trilha 6: camada de aprimoramento*
 - *tref (scal): track_ID = 1*
 - *tref (tile): track_ID=2*

track_ID=3

track_ID=4

track_ID=5

[0198] A Figura 5 ilustra aproximadamente o formato de arquivo obtido através de segmentos de mídia em concatenação quando apenas os arquivos de segmento de mídia exigidos (que correspondem, no presente documento, à camada de base e aos

tiles a e c na camada de aprimoramento) são transferidos por download de um servidor. Deve ser notado que um tal mecanismo não só permite a transferência por transferir por download apenas dos arquivos de segmento de mídia exigidos, como também impede a transferência por transferir por download de dados duplicados.

[0199] Conforme ilustrado, a trilha compósita 500 permite a construção de um fluxo de bits de dados de mídia temporizada decodificável válido 502 referenciando-se os dados da trilha de camada de base 504 (no caso de escalabilidade) e de trilhas de tile não reproduzíveis (506 e 508) e manuseando-se de modo apropriado os extratores que referenciam os dados ausentes (conforme descrito através de referência à Figura 7).

[0200] O formato de arquivo obtido é com a definição de formato de arquivo escalonável. Por exemplo, um dispositivo de cliente pode decidir reproduzir apenas a camada de base selecionando-se a trilha de camada de base ou a camada de aprimoramento com os tiles a e c escolhidos através da seleção da trilha compósita. O dispositivo de cliente também pode alterar os tiles a serem exibidos transferindo-se por transferir por download diferentes "trilhas de tile" (isto é, arquivos de segmento de mídia) em um último período de tempo enquanto o mesmo continua a reproduzir a trilha compósita.

[0201] A Figura 6, que compreende a Figura 6a e Figura 6b, é um fluxograma que ilustra as etapas para transmitir os dados de mídia temporizados entre um servidor e um dispositivo de cliente de acordo com uma modalidade em particular. As etapas mostradas na Figura 6a são implantadas no servidor para preparar uma apresentação de mídia criando-se os arquivos de segmento adaptados ao fluxo contínuo de ROI a partir de um fluxo de bits de dados de mídia temporizada de tile enquanto as etapas mostradas na Figura 6b são implantadas no dispositivo de cliente.

[0202] Em uma primeira etapa (etapa 600), o servidor identifica todas as unidades de NAL que são associadas aos tiles e, para cada tile, cria uma trilha de tile que contém subamostras compostas por todas as unidades de NAL que correspondem ao dado tile. Por exemplo, o servidor pode contar com as mensagens de SEI de nível de subgravura para identificar a associação de unidades de NAL com regiões diferentes

e em mensagens de SEI de nível de sequência para identificar a posição e o tamanho de cada ROI conforme foi proposto na padronização HEVC (proposta JCTVC-K0128). Dessa maneira, o servidor pode criar conexões de tiles para dados períodos de tempo.

[0203] Em uma etapa a seguir (etapa 602), o servidor cria uma trilha compósita que contém extratores ligados às trilhas de escalabilidade de nível inferior (caso haja), unidades de NAL comuns para todos os tiles e extratores ligados a cada trilha de tile. Os extratores e unidades de NAL comuns são vantajosamente ordenados de tal modo que a substituição dos extratores pelos dados são resultados de referência em um fluxo de bits de dados de mídia temporizada válido composto por amostras completas na ordem de decodificação (conforme descrito com referência à Figura 1).

[0204] Em seguida, na etapa 604, o servidor gera e armazena um arquivo de segmento de inicialização e arquivos de segmento de mídia que contêm o período de tempo de acordo com a representação de ISO BMFF, conforme descrito com referência às Figuras 3 e 4. Todas as trilhas de dados de mídia temporizada (por exemplo, trilhas de vídeo), trilha compósita e trilhas de tile são armazenadas em arquivos de segmento de mídia separados.

[0205] O servidor, então, serve, mediante solicitação, os arquivos de segmento de mídia e de inicialização a um dispositivo de cliente (etapa 606). O servidor pode ser um servidor de HTTP convencional que responde às solicitações HTTP.

[0206] No contexto de fluxo contínuo HTTP e em uma modalidade preferencial, supõe-se que o dispositivo de cliente tenha acesso a um arquivo de manifesto que descreve a apresentação de mídia disponível a partir do servidor. Esse arquivo de manifesto fornece informações suficientes (propriedades de mídia e uma lista de segmentos) para o dispositivo de cliente para transmitir de modo contínuo a apresentação de mídia solicitando-se, primeiro, os segmentos de inicialização e, então, os arquivos de segmento de mídia do servidor.

[0207] Na seleção de um ROI na extremidade do dispositivo de cliente, tipicamente em um visor com meios de seleção como um dispositivo apontador, durante o fluxo contínuo de um vídeo de tile, sendo que os tiles que correspondem ao ROI selecionado são determinados (etapa 608).

[0208] Em seguida, para cada período de tempo, no caso de dados de mídia escalonável, o dispositivo de cliente envia uma solicitação para o servidor para transferir por download os arquivos de segmento que correspondem às camadas dependentes (etapa 610). De acordo com a modalidade em particular, as camadas que são dependentes são transferidas por download antes de as camadas que dependem dessas dependerem das camadas. Por exemplo, os arquivos de segmento de camada de base são transferidos por download antes dos arquivos de segmento de camada de aprimoramento.

[0209] Em uma etapa a seguir, o dispositivo de cliente envia uma solicitação para o servidor para transferir por download os arquivos de segmento de mídia que correspondem aos tiles selecionados (etapa 612) e envia uma solicitação para o servidor para transferir por download os arquivos de segmento de mídia que correspondem à trilha compósita (etapa 614).

[0210] Em seguida, os arquivos de segmento transferidos por download são concatenados pelo dispositivo de cliente para construir um fluxo de bits de dados de mídia temporizada em conformidade com o padrão ISO BMFF (etapa 616), que corresponde ao ROI selecionado.

[0211] A etapa 616 é descrita em detalhes com referência à Figura 7.

[0212] A Figura 7 é um fluxograma que ilustra a geração de um fluxo de bits de dados de mídia temporizada válido a partir de segmentos de mídia concatenados recebidos por um dispositivo de cliente, que é a geração de um fluxo de bits decodificável a partir de tiles selecionados e uma trilha compósita quando a trilha compósita é reproduzida pelo dispositivo de cliente.

[0213] Em uma primeira etapa (etapa 700), o dispositivo de cliente recebe os arquivos de segmento de mídia que foram anteriormente solicitados (por exemplo, etapas 612, 614 e 616 na Figura 6) e um teste é realizado para determinar se pelo menos um arquivo de segmento de mídia foi ou não recebido (etapa 702). Se nenhum arquivo de segmento de mídia tiver sido recebido, o processo acaba.

[0214] Se pelo menos um arquivo de segmento de mídia tiver sido recebido e se os arquivos de segmento de mídia recebidos não contiverem uma trilha compósita

(isto é, se eles contiverem trilhas dependentes, por exemplo, camada de nível inferior ou trilhas de tile), os mesmos são armazenados em buffer para o uso posterior.

[0215] Ao contrário, se pelo menos um segmento de mídia tiver sido recebido e se o arquivo de segmento de mídia recebido contiver uma trilha compósita, o fluxo de bits (tipicamente unidades de NAL) que corresponde à trilha compósita é analisado para extrair o primeiro item de dados (ou o próximo item de dados se pelo menos um item de dados do segmento de mídia recebido tiver sido processado, tipicamente uma unidade de NAL) da caixa de dados de mídia "mdat" do segmento de mídia recebido (etapa 704). Se não houver item de dados adicional para analisar na caixa de dados de mídia, o processo retorna para a etapa 700 para receber novos arquivos de segmento de mídia (etapa 706).

[0216] Em seguida, um teste é realizado para determinar se o item de dados extraído (por exemplo, a unidade de NAL extraída) corresponde ou não a um extrator (etapa 708). Se o item de dados extraído não corresponder a um extrator, é retornado conforme deve ser adicionalmente decodificado por um decodificador de vídeo (etapa 710). Ao contrário, se o item de dados extraído for um extrator, o mesmo deve ser substituído pelo item de dados ao que o mesmo se refere. Para esse fim, os valores dos parâmetros do extrator são obtidos a partir de sua estrutura (etapa 712). Conforme descrito acima, um extrator compreende todos os valores de parâmetros exigidos para extrair dados de outra trilha (por exemplo, *track_ref_index*, *sample_offset*, *data_offset* e *data_length*).

[0217] Uma vez que o identificador da trilha referenciada tiver sido identificado, um teste é realizado (etapa 714) para determinar se a trilha referenciada está ou não disponível no conjunto de arquivos de segmento de mídia armazenados em buffer durante a etapa 700. Deve-se recordar que algumas trilhas de tile estão ausentes uma vez que o dispositivo de cliente transfere por download apenas os arquivos de segmento de mídia que correspondem à Região de Interesse selecionada.

[0218] Se a trilha referenciada estiver disponível no conjunto de arquivos de segmento de mídia armazenados em buffer durante a etapa 700, o extrator é substituído pelos dados que o mesmo referencia (etapa 716) e o fluxo de bits é

enviado para um decodificador de vídeo a ser decodificado (etapa 710).

[0219] Se a trilha referenciada não estiver disponível no conjunto de arquivos de segmento de mídia armazenados em buffer durante a etapa 700, as etapas específicas têm que ser realizadas uma vez que a ausência de dados referenciados em um extrator leva a um erro fatal de acordo com o padrão ISO BMF. Um teste foi realizado para determinar se a trilha referenciada é ou não uma trilha de tile (a trilha referenciada pode corresponder a uma camada de escalabilidade dependente) e se o extrator é do tipo de tile (etapa 718).

[0220] Se a trilha referenciada não for uma trilha de tile ou se o extrator não for do tipo de tile, um erro fatal padrão é detectado. Ao contrário, se a trilha referenciada for uma trilha de tile e se o extrator for do tipo de tile, o extrator é removido (etapa 722) ou o extrator é substituído através de preenchimento a partir de uma "trilha de preenchimento" ou "caixa de preenchimento" alternativa que contém dados "ignorados" para os tiles ausentes (etapa 724), dependendo do formato de codificação usado para codificar o fluxo de bits de dados de mídia temporizada (etapa 720). No presente documento, dados "ignorados" representam dados de pixel ausentes em uma imagem atual que são substituídos por outros dados de pixel obtidos a partir de uma imagem previamente decodificada que pertence a uma mesma camada escalonável ou que pertence a outra camada escalonável. Os dados "ignorados" são geralmente representados por pelo menos um sinalizador. Por exemplo, quando se considera o formato de compactação de vídeo HEVC, os dados de preenchimento podem ser um ou mais NALUs que contêm, exclusivamente, unidades de codificação codificadas com um sinalizador ignorado ajustado para 1.

[0221] Em seguida, o fluxo de bits é transmitido para um decodificador de vídeo a ser decodificado (etapa 710) e exibido e o processo volta para a etapa 704 para lidar com um item de dados a seguir.

[0222] A Figura 8 representa um diagrama de blocos de um servidor ou um dispositivo de cliente 800 em que as etapas de uma ou mais modalidades podem ser implantadas.

[0223] Preferencialmente, o dispositivo 800 compreende um barramento de

comunicação 802, uma unidade de processamento central (CPU) 804 com capacidade de executar instruções da ROM do programa 806 na alimentação do dispositivo e instruções relacionadas a um aplicativo de software da memória principal 808 após a alimentação. A memória principal 808 é, por exemplo, do tipo Memória de Acesso Aleatório (RAM) que funciona como uma área de trabalho da CPU 804 por meio do barramento de comunicação 802 e a capacidade da memória da mesma pode ser expandida por uma RAM opcional conectada a uma porta de expansão (não ilustrada). As instruções relacionadas ao aplicativo de software podem ser carregadas para a memória principal 808 de um disco rígido (HD) 810 ou a ROM do programa 806, por exemplo. Tal aplicativo de software, quando executado pela CPU 804, faz com que as etapas descritas com referência às Figuras 6a sejam realizadas no servidor e que as etapas descritas com referência às Figuras 6b e 7 sejam realizadas no dispositivo de cliente.

[0224] A referência numérica 812 é uma interface de rede que permite a conexão do dispositivo 800 à rede de comunicação 814. O aplicativo de software quando executado pela CPU 804 é adaptado para reagir às solicitações recebidas através da interface de rede e para fornecer fluxos contínuos de dados e solicitação por meio da rede a outros dispositivos.

[0225] A referência numérica 816 representa as interfaces de usuário para exibir informações e/ou receber entradas de um usuário.

[0226] Deve ser indicado no presente documento, como uma variante, que o dispositivo 800 para gerenciar a recepção ou o envio de fluxos de bits de multimídia pode consistir em um ou mais circuitos integrados dedicados (ASIC) que têm a capacidade de implantar o método conforme descrito com referência às Figuras 6a, 6b e 7. Esses circuitos integrados são, por exemplo, e de modo não restritivo, integrados em um aparelho para gerar ou exibir sequências de vídeo e/ou para listar sequências de áudio.

[0227] Conforme descrito acima, uma modalidade da invenção pode se aplicar, em particular, ao formato de vídeo conhecido como HEVC.

[0228] De acordo com o padrão HEVC, as imagens podem ser divididas de modo

espacial em tiles, fatias e segmentos de fatia. Nesse padrão, um tile corresponde a uma região retangular de uma imagem que é definida por limites horizontal e vertical (isto é, fileiras e colunas). O mesmo contém um número inteiro de Unidades de Codificação em Árvore (CTU). Portanto, os tiles podem ser usados de modo eficaz para identificar as regiões de interesse definindo-se, por exemplo, posições e tamanhos para as regiões de interesse. No entanto, a estrutura de um fluxo de bits HEVC assim como sua encapsulação como unidades de Camada Abstrata de Rede (NAL) não são organizadas em vista de tiles, mas são baseadas em fatias.

[0229] No padrão HEVC, as fatias são conjuntos de segmentos de fatia, sendo que o primeiro segmento de fatia de um conjunto de segmentos de fatia é um segmento de fatia independente, ou seja, um segmento de fatia que as informações gerais armazenadas em um cabeçalho não se referem a um de outro segmento de fatia. Os outros segmentos de fatia do conjunto de segmentos de fatia, caso haja, são segmentos de fatia dependentes (isto é, segmentos de fatia que as informações gerais armazenadas em um cabeçalho se referem a um de um segmento de fatia independente).

[0230] Um segmento de fatia contém um número inteiro de Unidades de Codificação em Árvore consecutivas (na ordem de varredura por rastreo). Portanto, um segmento de fatia pode ser de um formato retangular ou não então, não é adequado para representar uma região de interesse. É codificado em um fluxo de bits HEVC sob a forma de um segmento de fatia o cabeçalho seguido pelos dados de segmento de fatia. Os segmentos de fatia independentes e dependentes diferem por seu cabeçalho: uma vez que um segmento de fatia dependente de um segmento de fatia independente, a quantidade de informações de seu cabeçalho é menor do que aquela de um segmento de fatia independente. Tanto os segmentos de fatia independentes quanto os dependentes contêm uma lista de pontos de entrada no fluxo de bits correspondente que são usados para definir tiles ou como pontos de sincronização de decodificação por entropia.

[0231] A Figura 9, que compreende Figuras 9a, 9b e 9c, ilustra exemplos de tiles e segmentos de fatia. Mais precisamente, a Figura 9a ilustra uma imagem (900)

dividida em nove porções por limites verticais 905-1 e 905-2 e limites horizontais 910-1 e 910-2. Cada uma das nove porções referidas 915-1 a 915-9 representa um tile em particular.

[0232] A Figura 9b ilustra uma imagem (900') que contém dois tiles verticais delimitados pelo limite vertical 905'. A imagem 900' compreende uma única fatia (não referida) que contém cinco segmentos de fatia, um segmento de fatia independente 920-1 (representado com linhas hachuradas) e quatro segmentos de fatia dependentes 920-2 a 920-5.

[0233] A Figura 9c ilustra uma imagem (900'') que contém dois tiles verticais delimitados pelo limite vertical 905''. O tile esquerdo compreende duas fatias: uma primeira fatia que contém um segmento de fatia independente (920'-1) e um segmento de fatia dependente (920'-2) e uma segunda fatia que também contém um segmento de fatia independente (920'-3) e um segmento de fatia dependente (920'-4). O tile direito compreende uma fatia que contém um segmento de fatia independente (920'-5) e um segmento de fatia dependente (920'-6).

[0234] De acordo com o padrão HEVC, os segmentos de fatia são ligados aos tiles de acordo com as regras que podem ser resumidas conforme segue (uma ou ambas as condições têm que ser satisfeitas):

- todas as CTUs em um segmento de fatia pertencem ao mesmo tile (isto é, um segmento de fatia não pode pertencer aos diversos tiles); e
- todas as CTUs em um tile pertencem ao mesmo segmento de fatia (isto é, um tile pode ser dividido em diversos segmentos de fatia contanto que cada um desses segmentos de fatia pertença apenas àquele tile).

[0235] Por questão de clareza, considera-se, a seguir, que o um tile contém uma fatia que tem apenas um segmento de fatia independente. No entanto, as modalidades da invenção podem ser realizadas com outras configurações como aquelas ilustradas nas Figuras 9b e 9c.

[0236] Conforme mencionado acima, embora os tiles possam ser considerados como um suporte apropriado para as regiões de interesse, os segmentos de fatia são as entidades que são, de fato, colocadas em unidades de NAL para o transporte por

uma rede de comunicação e agregados para formar as unidades de acesso (isto é, gravura ou amostras codificadas em nível de formato de arquivo).

[0237] Deve-se lembrar que, de acordo com o padrão HEVC, o tipo de uma unidade de NAL é codificado em dois bytes do cabeçalho de unidade de NAL que pode ser definido conforme segue:

```
nal_unit_header () {  
  forbidden_zero_bit  
  nal_unit_type  
  nuh_layer_id  
  nuh_temporal_id_plus_1  
}
```

[0238] As unidades de NAL usadas para codificar os segmentos de fatia compreendem cabeçalhos de segmento de fatia que indicam o endereço da primeira CTU no segmento de fatia graças a um elemento de sintaxe de endereço de segmento de fatia. Tais cabeçalhos de segmento de fatia podem ser definidos conforme segue:

```
slice_segment_header () {  
  first_slice_segment_in_pic_flag  
  se (nal_unit_type >= BLA_W_LP && nal_unit_type <= RSV_IRAP_VCL23)  
no_output_of_prior_pics_flag  
  slice_pic_parameter_set_id  
  se (!first_slice_segment_in_pic_flag){  
    se (dependent_slice_segments_enabled_flag)  
    dependent_slice_segment_flag  
    slice_segment_address  
  }  
  se (!dependent_slice_segment_flag){  
    [...]
```

[0239] As informações de agrupamento em tile são fornecidas em uma unidade de NAL de PPS (Conjunto de Parâmetro de Gravura). A relação entre um segmento de fatia e um tile pode, então, ser deduzida a partir desses parâmetros.

[0240] Muito embora as previsões espaciais sejam reajustadas nas bordas de tile (através da definição), nada impede que um tile use os preditores temporais a partir de um tile diferente no(s) quadro(s) de referência. Dessa maneira, para construir tiles independentes, os vetores de movimento para as unidades de previsão são vantajosamente restritos dentro de um tile, durante a codificação, para permanecer no tile colocalizado no(s) quadro(s) de referência. Além disso, nos filtros em laço (filtros de desbloqueio e de deslocamento adaptativo de amostra (SAO)) são preferencialmente desativados nas bordas de tile de modo que nenhum desvio de erro seja introduzido quando se decodifica apenas um tile. Deve-se notar que tal controle dos filtros em laço esteja disponível no padrão HEVC. O mesmo é definido no cabeçalho de segmento de fatia com um sinalizador conhecido como *loop_filter_across_tiles_enabled_flag*. Ajustando-se, explicitamente, esse sinalizador para zero, os pixels nas bordas de tile não podem depender dos pixels que são abrangidos pela borda dos tiles vizinhos. Quando essas duas condições relacionadas aos vetores de movimento e aos filtros em laço forem satisfeitas, os tiles podem ser considerados como "tiles decodificáveis independentemente" ou "tiles independentes".

[0241] Quando um fluxo de bits de vídeo for codificado como um conjunto de tiles independentes, o mesmo possibilita uma decodificação baseada em tile de um quadro para outro sem qualquer risco para os dados de referência ausentes ou propagação de erros de reconstrução. Essa configuração possibilita, então, reconstruir apenas uma parte espacial do vídeo original que pode corresponder, por exemplo, à região de interesse ilustrada na Figura 2 (que compreende os tiles 3 e 7). Tal configuração pode ser indicada como informações complementares em um fluxo de bits de vídeo de modo a indicar que a decodificação baseada em tile seja confiável.

[0242] De acordo com uma modalidade da invenção, um acesso eficaz aos tiles no contexto de fluxo contínuo HTTP é fornecido com o uso do formato de arquivo ISO BMFF aplicado ao padrão HEVC. Dessa maneira, cada um dos tiles independentes a serem codificados (por exemplo, cada um dos doze tiles representados na Figura 2) é representado por uma trilha específica chamada de uma "trilha de tile" conforme

descrito abaixo com referência à Figura 10.

[0243] Essas trilhas de tile são referidas (por meio de uma caixa de referência de trilha "tref" de uma caixa de filme "moov" que contém uma definição para cada uma das trilhas) em uma trilha compósita que corresponde ao fluxo de bits HEVC em quadro completo conforme ilustrado na Figura 10. Cada trilha de tile contém dados de vídeo compactados, empacotados nas unidades de NAL. A trilha compósita contém vários conjuntos de parâmetro (por exemplo, conjunto de parâmetro de vídeo, conjunto de parâmetro de sequência e/ou conjunto de parâmetro de gravura) que correspondem aos dados de inicialização. A mesma também contém extratores que são unidades de NAL de um tipo específico.

[0244] Conforme descrito acima, um extrator pode ter uma estrutura interna de formato de arquivo que tem a seguinte sintaxe:

```
class aligned(8) Extractor () {
    NALUnitHeader();
    unsigned int(8) track_ref_index;
    signed int(8) sample_offset;
    unsigned int((lengthSizeMinusOne+ 1) *8)
    data_offset;
    unsigned int((lengthSizeMinusOne+ 1) *8)
    data_length;
}
```

[0245] O extrator age como apontador ou referências aos dados de outras trilhas e possibilita construir trilhas compactas com referências às trilhas dependentes em vez de duplicação de dados em ambas as trilhas. Um extrator usa preferencialmente a sintaxe de unidade de NAL. Dessa maneira, o mesmo compreende um cabeçalho que tem a mesma estrutura que o cabeçalho de unidade de NAL, que compreende, em particular, informações que pertencem ao tipo de unidade de NAL. Esse tipo de unidade de NAL é ajustado, por exemplo, para o valor "47" que corresponde atualmente a um tipo de unidade de NAL reservada no HEVC. O cabeçalho é seguido por um índice (denotado *track_ref_index*) na caixa de referência de trilha (tref) que

permite recuperar a entrada da caixa tref que contém o identificador da trilha (*track_id*) que corresponde à trilha que é referida pelo extrator. Um terceiro parâmetro é um deslocamento de tempo da amostra (*sample_offset*) que é referido pelo extrator em comparação à amostra atual. O quarto e o quinto parâmetros (denotados *data_offset* e *data_length*) fornecem, respectivamente, a posição (preferencialmente em bytes) a partir de onde copiar e a quantidade de dados a copiar (o valor 0 é reservado para indicar a cópia de toda a unidade de NAL referida).

[0246] A Figura 10 ilustra um exemplo de encapsulação de um fluxo de bits HEVC como um conjunto de trilhas que compreendem uma trilha compósita e trilhas de tile independentes, de acordo com uma modalidade da invenção. Por questão de ilustração, o fluxo de bits encapsulado corresponde à sequência de vídeo ilustra esquematicamente na Figura 2.

[0247] Conforme ilustrado, o fluxo de bits encapsulado 1000 compreende um arquivo de segmento de inicialização 1005 que contém uma caixa de filme ("moov") fornecendo uma definição para as trilhas e o arquivo de segmento de mídia 1010 que representa uma trilha compósita 1015 e doze trilhas de tile 1020-1 a 1020-12 (sendo que cada uma das trilhas de tile 1020-1 a 1020-12 é associada a um tile da sequência de vídeo).

[0248] A trilha compósita 1015 compreende, conforme especificado no padrão DASH, uma caixa do tipo segmento "styp" (não representada), pelo menos uma caixa de fragmento de filme "moof" 1025 que compreende metadados como um tipo de segmento de trilha e um identificador e pelo menos uma caixa de dados de mídia "mdat" 1030 que compreende, para cada amostra de dados de vídeo, PPS e referências aos dados de vídeo.

[0249] Semelhantemente, cada uma das trilhas de tile 1020-1 a 1020-12 compreende uma caixa do tipo segmento "styp" (não representada), pelo menos uma caixa de fragmento de filme "moof" que compreende metadados como um tipo de segmento de trilha e um identificador e pelo menos uma caixa de dados de mídia "mdat" que compreende, dados de vídeo comprimidos empacotados em unidades de NAL (NALUs).

[0250] As trilhas de tile 1020-1 a 1020-12, que tem o identificador 2 a 13, são referidos na caixa de referência de trilha "tref" 1035 do arquivo de segmento de inicialização 1005 (com mais precisão da caixa de filme "moov" do arquivo de segmento de inicialização 1005, na definição da trilha compósita que tem o identificador id=1).

[0251] Conforme ilustrado, a trilha compósita 1015 compreende extratores que agem como apontadores e referências aos dados a partir de outras trilhas. Por questão de ilustração, diversos parâmetros dentre os quais um índice de uma trilha de tile (*track_ref_index*), um deslocamento de dados (*data_offset*) e um comprimento de dados (*data_length*) que corresponde aos extratores 1035-1 e 1035-p da trilha compósita 1015 são representados.

[0252] Ainda por questão de ilustração, quando a unidade de NAL 1035-1 da trilha compósita 1015 for processada, determina-se que a mesma representa uma unidade de NAL do tipo extrator (NALUnitHeader é igual ao valor hexadecimal 5E00). Dessa maneira, a mesma é processada a fim de recuperar os dados de vídeo compactados correspondentes. Para esse fim, seu índice de trilha de tile (isto é, *track_ref_index*) é obtido. Para esse índice, é possível recuperar um identificador de trilha de tile a partir das definições de trilha de tile armazenadas no arquivo de segmento de inicialização 1005. No exemplo dado, uma vez que índice é igual a um, o primeiro identificador de trilha de tile da caixa tref é selecionado (id=2). Em seguida, esse identificador é usado para acessar a trilha de tile correspondente e, então, usar parâmetros de deslocamento de dados (isto é, um índice relativo de uma amostra na trilha identificada que deve ser usada como a fonte de informações) e de comprimento de dados (isto é, o número de bytes para copiar, por exemplo, toda a NALU quando *data_length*=0) do extrator 1035-1, os dados de vídeo compactados são extraídos da trilha de tile 1020-1 (isto é, segmento de fatia codificados NALU 1040 no exemplo dado).

[0253] Após ter sido processado um extrator é substituído pelos dados que o mesmo referencia. De acordo com o exemplo ilustrado na Figura 10, a análise e o processamento do extrator 1035-1 levam a sua substituição pelo segmento de fatia NALU codificado 1040 formando, assim, um fluxo de bits compatível com HEVC.

[0254] Deve-se notar que a semântica usada para armazenar os parâmetros de um extrator HEVC pode ser próxima de uma definida no padrão SVC. Dessa maneira, para um extrator que referencia as unidades de NAL HEVC, o seguinte se aplica:

- o parâmetro conhecido como `forbidden_zero_bit` é definido conforme especificado em ISO/IEC 23008-2;
- o parâmetro conhecido como `nal_unit_type` é definido como 47 (código reservado em FDIS atual);
- os parâmetros conhecidos como `nuh_layer_id` e `nuh_temporal_id_plus1` são copiados a partir da primeira NALU referida pelo extrator (um extrator em uma trilha HEVC que referencia as unidades de NAL HEVC não se refere a diversas unidades de NAL com diferentes valores de `nuh_layer_id` e `nuh_temporal_id_plus1`); e
- o parâmetro conhecido como `sample_offset` é ajustado para 0.

[0255] De acordo com as modalidades da invenção, as trilhas de tile devem ser consideradas como trilhas em particular usadas para suportar a indexação de dados dispostos de modo espacial (por exemplo, tiles HEVC). Portanto, a sinalização específica é exigida para cada trilha para indicar, em particular, a posição do tile no vídeo completo (coordenadas horizontal e vertical) e o tamanho do tile (largura e altura) e para fornecer informações para indicar se a trilha de tile pode ou não ser decodificada sem qualquer artefato. Tal sinalização pode ser alcançada de acordo com diversas modalidades.

[0256] De acordo com a modalidade em particular ilustrada na Figura 11, semelhante àquela descrita acima, a sinalização pode ser feita no nível da trilha.

[0257] A Figura 11 ilustra um exemplo de diagramas de blocos de um arquivo de segmento de inicialização que usa trilhas de dados de mídia, de acordo com uma primeira modalidade, para encapsular um fluxo de bits HEVC. A mesma se baseia no exemplo descrito por meio de referência à Figura 10.

[0258] O arquivo de segmento de inicialização 1100 e os arquivos de segmento de mídia (não representados) são usados para reconhecer e encapsular um fluxo de bits de vídeo em conformidade com o padrão HEVC no formato de arquivo ISO Base

Media. As trilhas são transmitidas de modo contínuo independentemente.

[0259] Conforme descrita acima, o arquivo de segmento de inicialização é usado para transmitir todos os metadados que são necessários para definir fluxos de bits de dados de mídia temporizada encapsulados em outros arquivos de segmento de mídia. Conforme ilustrado na Figura 11, o arquivo de segmento de inicialização 1100 contém uma caixa do tipo de arquivo "ftyp" 1105 e uma caixa de filme "moov" 1110. A caixa do tipo de arquivo 1105 identifica preferencialmente quais especificações ISO BMF os arquivos de segmento são compatíveis e indicam um número de versão dessa especificação. A caixa de filme "moov" 1110 fornece todos os metadados que descrevem a apresentação armazenada nos arquivos de segmento de mídia e, em particular, todas as trilhas disponíveis na apresentação.

[0260] A caixa de filme 1110 contém uma definição para cada uma das trilhas (caixa de "trilha" 1115-1 a 1115-13) que compreendem, no dado exemplo, uma trilha compósita (1115-1) e doze trilhas de tile (1115-2 a 1115-13).

[0261] Cada caixa de trilha contém pelo menos uma caixa de cabeçalho de trilha "tkhd", referenciada de modo genérico 1120 e uma caixa de mídia de trilha "mdia" referenciada de modo genérico 1125. Se uma trilha depender dos dados de outras trilhas, também há uma caixa de referência de trilha "tref". Conforme ilustrado, a trilha compósita que tem o identificador track_ID=1 compreende a caixa de referência de trilha "tref" 1130 que indica que a trilha depende de dados das trilhas de tile que têm identificadores track_ID=2 a 13.

[0262] Conforme mencionado acima, deve-se notar que outras caixas podem ser obrigatórias ou opcionais dependendo das especificações ISO BMFF usadas para encapsular o fluxo de bits de dados de mídia temporizada. No entanto, uma vez que as modalidades da invenção não contam com o fato de que essas caixas sejam aplicáveis, as mesmas não são representadas no presente.

[0263] De acordo com a modalidade descrita por meio de referência à Figura 11, a sinalização da posição do tile no vídeo completo, o tamanho do tile e da indicação de que a trilha de tile pode ser decodificada sem qualquer artefato é feita uma vez para todo o fluxo de bits HEVC a ser encapsulado, na caixa "moov" (1110), em cada

definição de trilha, com o uso da caixa de cabeçalho de trilha "tkhd" (1120) e caixas da caixa de informações de mídia "mdia" (1125).

[0264] As posições dos tiles são colocadas em novo tipo de caixa de informações de cabeçalho de mídia 1135, referida como caixa *TileMediaHandlerEntry* ou "trmhd" (1140) que define os deslocamentos horizontal e vertical (*horizontal_offset* e um *vertical_offset*).

[0265] O tamanho da apresentação visual do tile considerado é colocado nos parâmetros de largura e altura do cabeçalho da trilha existente. Se necessário, as dimensões de pixel reais do tile considerado podem ser documentadas em descrições de amostra (por meio dos parâmetros conhecidos como *VisualSampleEntry()* na caixa conhecida como *SampleDescriptionBox*)

[0266] Por fim, a sinalização de uma trilha específica (trilha de tile) é explicitamente feita definindo-se um novo tipo de trilha ("tile") na caixa de manipulador "hdlr" (1145).

[0267] A caixa de filme "moov" 1110 do arquivo de segmento de inicialização 1100 compreende adicionalmente a caixa "mvex" 1150. Essa caixa é usada para informar ao cliente que acessa o arquivo encapsulado que os fragmentos de filme estão presentes. A mesma possibilita especificar no arquivo de segmento de inicialização a duração da mais longa trilha na apresentação. Isso torna mais simples a computação da duração da apresentação, evitando a examinação de cada duração de fragmento de filme. Conforme ilustrado, a caixa mvex 1150 contém uma caixa estendida de trilha por trilha a fim de evitar a duplicação de informações que são comuns para todos os fragmentos de cada trilha (isto é, as trilhas de tile e a trilha compósita), por exemplo, os identificadores de trilha e o tamanho de falha das amostras na trilha.

[0268] Deve-se notar que tal modalidade de trilhas de tile de sinalização não introduz sobrecarga significativa em comparação à sinalização de trilha de vídeo usual. Além do mais, a mesma permanece independente da quantidade de fragmentos de filme na qual a apresentação é dividida. No entanto, isso não é adaptado para variar a configuração de agrupamento em tiles ao longo de uma sequência de vídeo.

[0269] De acordo com outra modalidade em particular que é adaptada para manipular a variação na configuração de agrupamento em tiles ao longo de uma sequência de vídeo, a sinalização é feita em um nível de amostra, com o uso de mecanismos de agrupamento de amostra do padrão ISO BMFF.

[0270] Tais mecanismos de agrupamento de amostra são usados para representar as partições de amostras nas trilhas. Os mesmos contam com o uso de duas caixas: uma caixa *SampleToGroup* ("sbgp") que descreve a atribuição de amostras aos grupos de amostra e uma caixa *SampleGroupDescription* ("sgpd") que descreve propriedades comuns de amostras em um grupo de amostra em particular. Um tipo em particular de agrupamento de amostra é definido pela combinação de uma caixa *SampleToGroup* e uma caixa *SampleGroupDescription* por meio de um tipo de campo ("grouping_type"). Diversas ocasiões de agrupamento de amostra (isto é, par de caixas de *SampleToGroup* e *SampleGroupDescription*) podem existir com base nos critérios de agrupamento diferentes.

[0271] De acordo com as modalidades da invenção, um novo critério de agrupamento relacionado ao agrupamento em tiles de amostras é definido. Esse novo *grouping_type*, chamado "tile", descreve as propriedades de um tile e é derivado do padrão *VisualSampleGroupEntry*. O mesmo pode ser referido como *TileRegionSampleGroupEntry* ou *HEVCSPatialEntry* e é definido conforme segue:

```
class HEVCSPatialEntry() extends VisualSampleGroupEntry ("trsg")
{ unsigned int(32) tileID;
  unsigned int(16) horizontal_offset;
  unsigned int(16) vertical_offset;
  unsigned int( 16) region_width;
  unsigned int(16) region height;
  unsigned int(2) independent;
  unsigned int(6) reserved=0;
}
```

[0272] De acordo com esse novo tipo de entrada de grupo, o parâmetro *tileID* é um identificador exclusivo para o tile descrito pelo grupo. Os parâmetros

horizontal_offset e *vertical_offset* são usados para definir um deslocamento horizontal e um vertical, respectivamente, do pixel esquerdo de topo da região retangular representada pelo tile, em relação ao pixel esquerdo de topo do quadro HEVC, em amostras luma da região de base. Os parâmetros *region_width* e *region_height* são usados para definir a largura e a altura, respectivamente, da região retangular representada pelo tile, em amostras luma do quadro HEVC. O parâmetro *independente* é uma palavra de 2 bits que especifica que o tile compreende dependências de decodificação relacionadas às amostras que pertencem apenas ao mesmo tile, conforme descrito acima em referência à definição dos tiles independentes. Por questão de ilustração e com referência a um uso padrão de mensagens SEI para descrever a organização de tile, o sinalizador conhecido como *tile_section_exact_match_flag* pode ser usado para definir o valor do sinalizador independente. O significado do último pode ser definido conforme segue:

- se o parâmetro independente for igual a 0, as dependências de codificação entre esse tile e outros tiles no mesmo quadro ou em quadros anteriores são conhecidas;
- se o parâmetro independente for igual a 1, não há dependências de codificação espacial entre esse tile e outros tiles no mesmo quadro, mas pode haver dependências de codificação entre esse tile e o tile que tem o mesmo tileID nos quadros anteriores e
- se o parâmetro independente for igual a 2, não há as dependências de codificação entre esse tile e outros tiles que têm o mesmo tileD no mesmo quadro ou em quadros anteriores;

[0273] sendo que o valor de parâmetro independente 3 é reservado.

[0274] Opcionalmente, um parâmetro que descreve uma taxa de bits média por tile pode ser definida no descritor de tile de modo a ser fornecida para o cliente por meio de fluxo contínuo para a adaptação com base na largura de banda.

[0275] De acordo com essa modalidade, as propriedades de cada tile são dadas uma vez no cabeçalho de filme (caixa "moov") definindo-se, para cada trilha de tile, uma caixa *SampleGroupDescription* ("sgpd") com o tipo de agrupamento de "tile" e

uma *HEVCSpatialEntry*. Então, de acordo com o padrão ISO BMFF, uma caixa *SampleToGroup* é definida em cada fragmento de trilha de tile para associar cada amostra do fragmento de trilha de tile com suas propriedades uma vez que a quantidade de amostras não é conhecida antecipadamente.

[0276] No caso de a grade de tiles mudar com o tempo, uma nova caixa *SampleGroupDescription* ("sgpd") com uma nova *HEVCSpatialEntry* pode ser definida na caixa de fragmento de caixa ("traf") e referenciada pela caixa *SampleToGroup* ("sbgp"). Portanto, no caso de acordo com o que a grade é estática com o tempo, pelo menos uma caixa *SampleToGroup* é definida por trilha de tile e fragmento de trilha de tile. Essa caixa representa, em termos de descrição, pelo menos 28 bytes. Supondo-se 16 tiles com fragmentos de 2 segundos de duração, isso representaria 1.792 bits por segundo para sinalizar a configuração de agrupamento em tiles com o tempo, apenas para a caixa *SampleToGroup*. No caso de acordo com o qual a grade está mudando com o tempo, o custo (em termos de quantidade de dados) seria maior. Conforme descrito abaixo, essa quantidade de dados de inicialização complementares pode ser reduzida.

[0277] Deve-se notar que as caixas de grupo de amostra fornecem uma ferramenta eficaz e extensível para adicionar metadados às amostras individuais. No entanto, é bastante comum que os dados metadados sejam válidos para todas as amostras de uma dada trilha, conforme esse pode ser o caso com o descritor de agrupamento em tiles com um padrão NALU fixo em cada AU (isto é, com uma configuração de agrupamento em tiles estática ao longo do tempo).

[0278] Isso pode ser indicado através da sinalização, dessa maneira, de cada amostra de um grupo que usa uma caixa *SampleToGroup*.

[0279] Alternativamente, a fim de reduzir a quantidade de dados de inicialização por trilha de tile, alguns grupos de amostra podem ser marcados como "padrão" (isto é, válidos para todas as amostras). Esse agrupamento de padrão limita o custo de descrição em termos de bytes uma vez que, para a configuração de agrupamento em tiles estática, define-se no nível de caixa "moov", uma vez por trilha para toda a sequência.

[0280] Para esse fim, uma segunda versão (versão =2) de um tipo de descrição de grupo pode ser usada na caixa conhecida como *SampleGroupDescriptionBox* (pode haver múltiplas *SampleGroupDescriptionBox* por caixa traf/stbl), que indica (por meio do parâmetro conhecido como tipo de agrupamento) que o grupo de amostra referenciado se aplica a todas as amostras na trilha atual ou nos fragmentos de trilha atual.

[0281] A nova versão da caixa de descrição de grupo de amostra pode ser definida como segue:

```
aligned(8) class SampleGroupDescriptionBox (unsigned int (32) handler
type) extends FullBox("sgpd", versão, 0) {
    unsigned int(32) grouping_type; se
    se (version == 1) // (version == 2) { unsigned int (32) default_length; } unsigned
int (32) entry_count;
    int i;
    para ( i = 1; i <= entry_count; i++ ) {
        se (version != 0) {
            se (default_length == 0) {
                unsigned int(32) description_length;
            }
        }
        switch(handler_type) {
            case 'vide': // para trilhas de vídeo
                VisualSampleGroupEntry(grouping_type);
                break;
            case 'soun': // para trilhas de áudio
                AudioSampleGroupEntry(grouping_type);
                break;
            case 'hint': // para trilhas de dica
                HintSampleGroupEntry(grouping_type);
                break;
```


[0282] De acordo com esse tipo estendido de caixa *SampleGroupDescription*, quando a versão for igual a 2, essa descrição de grupo e qualquer *SampleGroupEntry* se aplica a todas as amostras da trilha atual ou fragmento de trilha. O parâmetro do tipo de agrupamento é um inteiro que identifica a caixa *SampleToGroup* que é associada à descrição de grupo de amostra. Deve-se notar que quando uma trilha ou um fragmento de trilha compreender em seus dados de inicialização uma caixa *SampleGroupDescription* com versão =2, não há necessidade de colocar nesses dados de inicialização uma caixa *SampleToGroup* que descreve o tipo de agrupamento correspondente (uma vez que a descrição se aplica a todas as amostras por definição).

[0283] Nessa definição da nova versão da caixa de descrição de grupo de amostra, *entry_count* é um inteiro que fornece inúmeras entradas na tabela seguinte, *default_length* indica o comprimento de toda entrada de grupo (se o comprimento for constante) ou zero se for variável e *description_length* indica o comprimento de uma entrada de grupo individual, no caso de variar de entrada para entrada e *default_length* ser, portanto, zero.

[0284] De acordo com essa modalidade, *grouping_type* poderia ter um valor específico para indicar um agrupamento espacial/tile: pode ser, por exemplo, o valor hexadecimal que corresponde ao código ASCII para "tile" (0x74696C65). No máximo, uma ocorrência dessa caixa com o mesmo valor para o tipo de agrupamento existe para uma trilha.

[0285] Deve-se notar que no caso de grade adaptativa que se move com o tempo, a amostra para a caixa de grupo permanece a mesma (isto é, "tile" de *grouping_type*) e continua aplicando em todas as amostras. Como tal, apenas a caixa de descrição de grupo de amostra precisa ser atualizada nos fragmentos de trilha das trilhas de tile cuja configuração de agrupamento em tiles mudou da configuração padrão sinalizada na caixa moov/trak/mdia/minf/stbl. Isso reduz o custo de sinalização para tiles adaptativos.

[0286] Alternativamente e ainda a fim de reduzir a quantidade de dados de inicialização por trilha de tile (para evitar repetir uma caixa *SampleToGroup* em cada

fragmento de trilha de tile), uma nova *DefaultSampleToGroups* referida "dsgp" (ou uma outra caixa semelhante que tem a mesma semântica, qualquer que seja o nome) é definida para ser incluída apenas na caixa *SampleTable* ("stbl") de cada uma das caixas moov/trak como parte das informações de sinalização. Essa nova caixa associaria a todas as amostras um conjunto de descrições de grupo de amostra que se aplicaria a todas as amostras em uma trilha.

[0287] A nova caixa *DefaultSampleToGroup* pode ser definida conforme segue:

```
aligned(8) class DefaultSampleToGroups extends FullBox('dsgp', version,
0) {
    unsigned int(32) entry_count;
    para (i=1; i <= entry_count; i++) {
        unsigned int(32) grouping_type;
        se (version == 1) {
            unsigned int(32) grouping_type_parameter;
        }
        unsigned int(32) group_description_index;
    }
}
```

[0288] em que o parâmetro *entry_count* fornece o número de entradas na lista de grupos a serem associados a cada amostra e o parâmetro *grouping_type* é um identificador para o tipo de agrupamento, referido na caixa *SampleGroupDescription*. Por exemplo, em uma modalidade em particular, o tipo de agrupamento pode ter um valor específico que indica um agrupamento espacial/de tile. Esse pode ser, por exemplo, o valor hexadecimal que corresponde ao código ASCII para o "tile" (0x74696C65). O parâmetro *group_description_index* é um inteiro que fornece o índice da entrada de grupo de amostra que descreve as amostras nesse grupo. O índice varia de um até o número de entradas de grupo de amostra na caixa *SampleGroupDescription* ou assume o valor zero para indicar que essa amostra não é um membro de qualquer grupo desse tipo. Por fim, o parâmetro *grouping_type_parameter* é uma indicação para o subtipo do agrupamento (caso

usado pelo tipo de agrupamento).

[0289] Isso possibilita sinalizar que todas as amostras de uma trilha seguem a mesma descrição de grupo para um dado tipo de agrupamento, com o uso de no máximo 32 bytes por tile independente do número de fragmentos de filme se apenas o agrupamento de tile estiver em uso (*entry_count*). No caso de a grade adaptativa se mover com o tempo, uma nova caixa *DefaultSampleToGroups* e uma nova caixa *SampleGroupDescription* poderia ser definida nos fragmentos de trilha. A nova caixa *DefaultSampleToGroups* substituiria a definição anterior e se refere à nova descrição de tile na nova caixa *SampleGroupDescription*. Assim, uma caixa *SampleToGroup* não é definida para cada fragmento de trilha, mas apenas quando a definição de grade de tile mudar.

[0290] Ainda de acordo com a modalidade em particular que é adaptada para manipular a variação na configuração de agrupamento em tiles ao longo de uma sequência de vídeo, a sinalização é feita no nível de subamostra com uma nova caixa de mapa de amostra.

[0291] Essa modalidade para a descrição de amostra com relação à configuração de agrupamento em tiles se aplica a um fluxo de bits HEVC de tile encapsulado de única trilha. Pode-se encontrar aplicação para a encapsulação por MPEG-4 de um fluxo de bits HEVC de tile ou no lado do cliente, após a resolução dos extratores de uma trilha compósita a fim de salvar as partes espaciais transferidas por download em uma única trilha de vídeo em conformidade com o padrão MPEG-4. Uma outra aplicação está direcionada à transformação de uma trilha compósita em um padrão de única trilha em conformidade com MPEG-4, endereçado a um analisador de MPEG-4 que não suporta os extratores.

[0292] A Figura 12, que compreende a Figura 12a e a Figura 12b, ilustram a sinalização da posição de um tile em um vídeo completo, do tamanho do tile e da indicação de que a trilha de tile pode ser decodificada sem qualquer artefato, no nível de subamostra, adaptado para manipular várias configurações de agrupamento em tiles.

[0293] A Figura 12a ilustra etapas realizadas por um dispositivo de cliente (por

exemplo, reprodutor de vídeo). Em uma primeira etapa (etapa 1200), o dispositivo de cliente transfere por download dados de inicialização ou lê dados de inicialização se o arquivo for um arquivo local, por exemplo, dados de inicialização de um fluxo de bits encapsulado em conformidade com o padrão MPEG-4, tipicamente o conteúdo de uma caixa moov.

[0294] Desses dados de inicialização, o dispositivo de cliente pode analisar as informações de cabeçalho de trilha em que as informações de tile são codificadas (etapa 1205). Com essas informações de tile, o usuário pode escolher uma região de interesse por meio de uma interface gráfica do dispositivo de cliente (etapa 1210) que pode corresponder a um ou mais tiles.

[0295] As trilhas de tile correspondentes assim como a trilha compósita são transferidas por download ou lidas pelo dispositivo de cliente (etapas 1215 e 1220). Em seguida, os extratores da trilha compósita são resolvidos com o uso de trilhas de tile de modo a obter uma única trilha de vídeo (etapa 1225). Por fim, o dispositivo de cliente constrói e adiciona a descrição de agrupamento em tiles, por exemplo, na *SampleTableBox*, na trilha de vídeo obtida (etapa 1230).

[0296] Um exemplo de descrição de agrupamento em tiles é ilustrado na Figura 12b. Conforme ilustrado, a descrição de agrupamento em tiles 1250 compreende a caixa de filme "moof" 1255 e caixa de dados "mdat" 1260. A caixa "moof" 1255 contém uma caixa *SampleTable* por trilha que contém uma caixa *SampleToGroup* 1265 que descreve os diferentes grupos de amostra, uma caixa de descrição de grupo de amostra 1270 que descreve o mapeamento entre as unidades de NAL de cada amostra e os tiles e uma caixa de descrição de grupo de amostra 1275 que contém as descrições de tile. A amostra para a caixa de grupo 1265 indica o tipo de agrupamento "tsgm" para a entrada de grupo *TileSampleMapEntry*.

[0297] A entrada de grupo *TileNALUMapEntry*1270 define o mapeamento entre as unidades de NAL de uma amostra e os tiles (essa é a razão pela qual tal modalidade se refere à sinalização de nível de subamostra). Essa caixa que o parâmetro *grouping_type* é igual a "tsgm" contém a quantidade de unidades de NAL por amostra.

[0298] A caixa *TileNALUMapEntry* pode ser definido conforme segue (conforme

ilustrado na Figura 12b):

```
class TileNALUMapEntryO extends VisualSampleGroupEntry ('tsgm')
{ unsigned int(8) reserved = 0;
  unsigned int(8) entry_count;
  para (i=1; i'<= entry_count; i++)
    unsigned int(32) tileID;
```

[0299] em que *entry_count* indica a quantidade de NALUs em uma amostra de trilha e *tileID* fornece o identificador exclusivo para o tile espacial descrito pela trilha atual.

[0300] A caixa *TileNALUMapEntry* também pode ser otimizada, em termos de tamanho, conforme segue:

```
class TileNALUMapEntryO extends VisualSampleGroupEntry ('tsgm')
{ unsigned int(6) reserved = 0;
  unsigned int(1) large_size;
  unsigned int( 1 ) mode;
  se (large_size) {
    unsigned int(16) entry_count;
  } ou {
    unsigned int(8) entry_count;
    para (i=1; i'<= entry_count; i++)
      se (mode) {
        se (large_size) {
          unsigned int(16) NALU_start_number;
        } ou {
          unsigned int(8) NALU_start_number;
        }
      }
    unsigned int(16) tileID;
```

[0301] em que o parâmetro *large_size* indica se a quantidade de entradas de unidade de NAL nas amostras de trilha é representada em 8 ou 16 bits e o parâmetro *mode* indica se cada entrada de unidade de NAL é descrita (quando *mode* não for definido) ou apenas as entradas de unidade de NAL que correspondem ao tile ID

mudam (quando *mode* for definido). No último caso, o número de unidade de NAL é codificado em 16 ou 8 bits dependendo do valor do parâmetro *large_size*.

[0302] Isso especifica o mapeamento entre cada uma dessas unidades de NAL e um específico tile. A descrição do tile é fornecida na caixa de descrição do grupo de amostra 1275, como descrito no presente documento acima, cada tile sendo descrito um após o outro.

[0303] Deve-se observar que o exemplo dado é um caso particular em que uma unidade de NAL contém dados para um tile e isso ao longo de uma duração de sequência. Quando os dados de tile dados são divididos em diversas unidades de NAL, diversos agregadores são usados para descrever o conjunto de unidades de NAL correspondentes ao tile. Quando o número de unidades de NAL varia ao longo do tempo, diversas entradas *tileSampleMap* podem ser definidas e as amostras para a referência de grupo alternativamente, a partir de um fragmento para outro, no cabeçalho de fragmento de trilha, o mapa do tile apropriado entra por meio do tipo de agrupamento.

[0304] No esquema de encapsulamento de dados como descrito pela referência à Figura 10, o fluxo de bits HEVC é encapsulado como uma trilha de composição 1015 que aponta para as trilhas em tiles 1020-1 a 1020-2 que contém realmente dados de vídeo compactados. A trilha de composição contém os dados de configuração originários a partir das unidades de NAL de conjunto de parâmetros de HEVC diferentes (chamado de PS na Figura 10). Os outros elementos da trilha de composição consistem principalmente em uma lista de extratores, um por amostra e por trilha de tile, apontando (por meio da referência de trilha caixa (“tref”) contido em caixa moov do arquivo de segmento de inicialização 1005) para dados de vídeo compactados encapsulados nas trilhas em tiles.

[0305] A atual dependência sinala meios em padrão ISO BMFF (Parte 15 do padrão) são localizados na referência de trilha de caixa “tref” que é parte de trilha de caixas na caixa moov do arquivo de segmento de inicialização 1005. A caixa “tref” fornece uma referência a partir da trilha de contenção para outra trilha na apresentação. A que contém trilha pode se referir a diversas outras trilhas na

apresentação. O tipo de dependência entre trilhas é especificado por um parâmetro *reference_type* que pode ter dois valores, “scal” ou “sbas”, no atual padrão, o valor “sbas” representa a *base escalonável*. Isso indica que trilha mencionada é a base escalonável de trilha da atual trilha na apresentação escalonável, o valor “scal” representa *escalabilidade*. Isso indica uma relação entre as trilhas que representam as diferentes camadas de uma representação escalonável. Isso significa que a caixa de contenção é dependente da trilha mencionada.

[0306] Na modalidade descrita por referência à Figura 10, não há dependências relacionadas de escalabilidade específica. Mesmo se vídeos escalonáveis forem levados em consideração, o foco aqui é feito em dependências espaciais entre a trilha de composição e as trilhas em tiles. Essas dependências podem ser indicadas explicitamente, por exemplo, com um novo valor de “tile” como feito na caixa tref da caixa moov do arquivo de segmento de inicialização 1005, que corresponde à trilha de composição1015 (id=1).

[0307] Embora o exemplo ilustrado de dependências a partir de uma trilha para outras trilhas seja direcionado a tiles (isto é, dependências de uma trilha de composição para trilhas em tiles), outros tipos de dependências a partir de uma trilha para uma ou para diversas trilhas podem ser lidados de modo similar. Portanto, o parâmetro *reference_type* pode indicar, por exemplo, uma dependência de subcamada (por exemplo, valor “subl”) para indicar que uma trilha é dependente de uma ou mais trilhas.

[0308] Consequentemente, trilhas de subcamadas podem ser definidas como trilhas que contém partes de um fluxo elementar de bits HEVC, que pode ser descartado sem danificar o processo de decodificação de outras unidades de NAL HEVC. Tal definição se aplica, em particular, às camadas temporárias no fluxo escalonável de bits HEVCs como também às trilhas em tiles, como descrito acima. Cada trilha que corresponde à uma trilha de subcamada pode ser marcada na gravação *HEVCConfiguration* (isto é, no *SampleTableBox*) que usa um bit (ou um sinalizador) que quando ajustado a um valor predeterminado indica que essa trilha HEVC é uma trilha de subcamada e só contém unidades de NAL que são relacionados

a partir de outra(s) trilha(s) (isto é, essa trilha HEVC não é exibível), por exemplo, a partir de uma trilha de composição. Quando o valor desse bit ou sinalizador tem um valor oposto, isso indica que essa trilha HEVC é uma trilha de subcamada que também contém dados de inicialização (isto é, essa trilha HEVC é exibível). Por exemplo, é possível usar bits reservados na atual caixa *HEVCDecoderConfigurationRecord*.

[0309] De acordo com a modalidade particular descrita por referência na Figura 12, cada trilha de tile é reproduzível como um padrão de trilha de vídeo.

[0310] A Figura 13 ilustra um exemplo de encapsular um fluxo de bits HEVC como um conjunto de trilhas que compreende uma trilha de composição e trilhas independentes em tiles que são reproduzíveis como padrão de trilhas de vídeo, de acordo com uma modalidade da invenção. Para fim de ilustração, o fluxo encapsulado de bits corresponde à sequência de vídeo esquematizada ilustrada na Figura 2.

[0311] O fluxo de encapsulamento de bits HEVC ilustrado na Figura 13 difere principalmente daquele ilustrado na Figura 10 nessa cada trilha de tile compreende um extrator particular que permite recuperação de dados de inicialização e configuração.

[0312] Como ilustrado, cada trilha em tiles de 1300-1 a 1300-12 compreende um extrator de 1305-1 a 1305-12 que aponta unidades de NAL de conjunto de parâmetros de HEVC (chamado de PS) de trilha de composição 1310, que representa os dados de inicialização e configuração, sendo chamados assim, de acordo com padrão HEVC, esses dados de inicialização e configuração correspondente tipicamente à diversos conjuntos de parâmetros do fluxo de bits HEVC. Consequentemente, tais dados de inicialização e configuração faz com que cada trilha de tile seja reproduzível como uma trilha normal de vídeo

[0313] O extrator adicionado à cada trilha de tile está vantajosamente localizado no início da caixa de dados de mídia “mdat” de cada trilha de tile, antes das amostras de dados de vídeo.

[0314] Essas dependências a partir das fitas de tiles (1300-1 a 1300-12) para a trilha de composição (1310), denotado de 1315-1 a 1315-12, tem que ser sinalizadas, por exemplo, no parâmetro *reference_type* de caixas “tref” 1320-1 a 1320-12

associadas às trilhas em tiles (na caixa de vídeo “moov” de arquivo de segmento de inicialização 1325). De acordo com essa modalidade, a trilha que contém o parâmetro ajustado é considerada como a trilha base HEVC “hbas” (isso é próximo ao caso SVC em que a trilha que contém a operação inferior aponta em uma apresentação escalonável é considerado como a “trilha base escalonável” “sbas”). Como ilustrado, as trilhas que dependem da trilha base (isto é, trilhas em lado de 1300-1 a 1300-12, que tem identificadores id=2 a 12, que dependem da trilha de composição 1310, que tem identificador id=1) tem o valor “hbas” na mesma caixa de referência de trilha (1320-1 a 1320-12).

[0315] Novamente, a sinalização do agrupamento em tiles pode ser no nível de trilha, no nível de amostral, ou nos níveis de trilha e amostras.

[0316] Deve-se observar que por padrão as trilhas em tiles são consideradas como não exibíveis. Entretanto, um analisador avançado conforme o padrão MPEG-4 pode detectar trilhas em tiles não exibíveis e expor as mesmas, por exemplo, em um arquivo de manifesto de fluxo contínuo olhando-se para a caixa “tref” (se a trilha de tile contém uma referência do tipo “hbas”, isso pode ser considerado como não reproduzível). Esses meios que essa trilha de tile pode ser considerada como um padrão de trilha de vídeo mesmo que marcada como um valor de “tile” na caixa manipuladora. Quando a sinalização do agrupamento em tiles é baseada em amostras, as trilhas em tiles ou trilhas de subcamadas podem ser etiquetadas como “vide” na mesma caixa manipuladora visto que as informações do agrupamento em tiles são postos na caixa conhecida como *SampleTableBox*.

[0317] A Figura 14 ilustra um exemplo de fluxo encapsular a fluxo de bits HEVC como um conjunto de trilhas que compreende uma trilha de composição, uma trilha de dados de inicialização, e trilhas independentes em tiles que são reproduzíveis com o padrão de trilhas de vídeo, de acordo com outra modalidade da invenção. Para fim de ilustração, o fluxo encapsulado de bits corresponde à sequência de vídeo esquematizada ilustrada na Figura 2.

[0318] A encapsulação de fluxo de bits de HEVC ilustrada na Figura 14 difere principalmente daquela ilustrada na Figura 13 de modo que os dados de inicialização

sejam colocados em uma trilha de dados de inicialização dedicada 1400 (e não na trilha de composição 1310).

[0319] Uma das vantagens fornecidas por tal modalidade quando comparada àquela descrita pela referência à Figura 13 é direcionada à quantidade de dados a ser transmitida quando as trilhas em tiles devem ser reproduzidas de modo independente. Visto que os dados de inicialização são transmitidos em uma trilha dedicada, não é pedido para transmitir a trilha de composição.

[0320] Deve ser lembrado que de acordo com o relatório descritivo atual do formato de arquivo de HEVC, há duas possibilidades para transportar conjuntos de parâmetros (PS) no formato de arquivo: na caixa conhecida como *Entrada de Amostra* apenas ou na caixa de *Entrada de Amostra* e nas amostras de dados. Essas duas configurações são respectivamente sinalizadas com as caixas “hvc1” e “hev1” em uma caixa conhecida como *Tabela de Amostras*. Embora armazenar os parâmetros em amostras seja mais complexo, isso permite mais dinamismo no caso de atualizações de conjunto de parâmetros. Portanto, em uma modalidade preferencial, os conjuntos de parâmetros são transportados na caixa de *Entrada de Amostra* e nos dados amostras (com valor “hev1” no parâmetro de *HEVC Sample Entries* na caixa de *Tabela de Amostras*) a fim de ter capacidade para manusear alterações de conjuntos de parâmetros de gravura (PPS), em particular, para agrupar em tiles as alterações de configuração.

[0321] Em conformidade, a trilha de dados de inicialização dedicada 1400 contém como dados apenas unidades de HEVC de não VCL de unidades similares a NAL de NAL que o tipo é igual a 32, 33 ou 34, que corresponde ao conjunto de parâmetro de vídeo, conjunto de parâmetro de sequência ou conjunto de parâmetro de gravura, respectivamente.

[0322] Conforme ilustrado na Figura 14, os extratores 1415-1 a 1415-12 localizados no início da caixa de dados de mídia “mdat” das trilhas em tiles 1410-1 a 1410-12 apontadas para os dados de trilha de dados de inicialização dedicada 1400. Do mesmo modo, o primeiro extrator (1420) da trilha de composição 1405 aponta para os dados de trilha de dados de inicialização dedicada 1400. Portanto, a trilha de dados

de inicialização 1400 é apenas a trilha do fluxo de bits de HEVC encapsulado que não faz referência a qualquer outra trilha. Como tal, visto que não há é qualquer dependência indicada na caixa tref (nenhuma dependência de “hbas” na caixa tref) associada à trilha de dados de inicialização 1400 (id=2), sendo que a última citada é considerada como não exibível de modo independente.

[0323] Quando alguns dados de inicialização são modificados no fluxo de bits de vídeo (isto é, quando Conjuntos de Parâmetro de Gravura ocorrem no fluxo de bits HEVC), os mesmos são colocados nos dados de amostra conforme ilustrado com referência 1425, na localização temporal em que alterações ocorrem. Os extratores correspondentes referenciados como 1430 e 1435-1 a 1435-12 são inseridos na trilha de composição 1405 e em cada uma das trilhas em tiles 1410-1 a 1410-12, respectivamente, ou seja, em cada trilha de tile que se refere a esses PPS novos.

[0324] Em cada trilha do fluxo de bits HEVC encapsulado, as amostras (e NALUs associados) são organizadas em ordem temporal. Do mesmo modo, os Conjuntos de Parâmetro de Gravura são organizados em ordem temporal na trilha de dados de inicialização dedicada 1400. Uma caixa “trun” (não representada na Figura 14) permite fornecer o tempo de codificação correto para cada amostra.

[0325] Naturalmente, a fim de satisfazer exigências locais e específicas, uma pessoa versada na técnica pode aplicar à solução descrita acima muitas modificações e alterações, todas as quais, no entanto, estão incluídas dentro do escopo de proteção da invenção conforme definido pelas reivindicações a seguir.

REIVINDICAÇÕES

1. Método para encapsular, usando um Formato de Arquivo de Mídia Base ISO, dados de mídia temporizada em tiles em um servidor, em que os dados de mídia temporizada em tile compreendem amostras temporizadas (100, 110, 112, 114), em que cada amostra temporizada compreende uma pluralidade de subamostras espaciais (1a, 1b, 1c, 1d), o método caracterizado por compreender:

para cada uma de uma ou mais subamostras espaciais, criar (600) uma trilha de tile que compreende a subamostra espacial, e uma subamostra espacial correspondente de cada uma dentre as outras amostras temporizadas;

criar (602) pelo menos uma trilha de composição que compreende informação de identificação (412-1, 412-3, 412-4, 412-5, 412-6) para identificar pelo menos uma das trilhas de tile criadas; e

encapsular de modo independente (604) cada uma das trilhas criadas em pelo menos um arquivo de segmento mídia (400-1, 400-2, 400-6).

2. Método, de acordo com a reivindicação 1, caracterizado pelo fato de que a informação de identificação (412-1, 412-3, 412-4, 412-5, 412-6) identifica adicionalmente pelo menos uma subamostra espacial da pelo menos uma trilha de tile identificada.

3. Método, de acordo com a reivindicação 1, caracterizado pelo fato de que compreende adicionalmente uma etapa de criar um arquivo de inicialização (300) que compreende parâmetros que permitem a descrição de uma porção espacial dos dados de mídia temporizada em tile a partir das trilhas de tile criadas e a pelo menos uma trilha de composição criada.

4. Método, de acordo com a reivindicação 3, caracterizado pelo fato de que o arquivo de inicialização (300) compreende adicionalmente uma lista de trilhas de tile referenciadas associadas a pelo menos uma trilha de composição criada.

5. Método, de acordo com a reivindicação 1, caracterizado pelo fato de que pelo menos um dos arquivos de segmento de mídia que resultam do encapsulamento de cada uma das trilhas criadas compreende uma estrutura de dados para armazenar

metadados incluindo pelo menos uma definição de pelo menos um grupo de subamostras da trilha correspondente.

6. Método, de acordo com a reivindicação 1, caracterizado pelo fato de que pelo menos um dos arquivos de segmento de mídia que resulta do encapsulamento de cada uma das trilhas criadas compreende uma estrutura de dados para armazenar metadados incluindo definições diferentes de grupos diferentes de subamostras da trilha correspondente.

7. Método, de acordo com a reivindicação 6, caracterizado pelo fato de que pelo menos um parâmetro da estrutura de dados que armazena metadados inclui pelo menos uma definição de pelo menos um grupo de subamostras da trilha correspondente, relacionadas a dependências de decodificação de subamostra, indica que cada subamostra do grupo pode ser decodificada sem usar uma subamostra de um outro grupo.

8. Método, de acordo com a reivindicação 3, caracterizado pelo fato de que o arquivo de inicialização (300) compreende uma definição a ser usada por padrão para processar subamostras de pelo menos uma trilha de tile que não são associadas a qualquer definição específica armazenada no arquivo de segmento de mídia que resulta do encapsulamento da pelo menos uma trilha de tile correspondente.

9. Método para fornecer um fluxo de bits de dados de mídia temporizada em um Formato de Arquivo de Mídia Base ISO a partir de dados de mídia temporizada em tiles encapsulados em uma pluralidade de arquivos de segmento de mídia (400-1, 400-2, 400-6), em um dispositivo de cliente, em que os dados de mídia temporizada em tiles compreendem amostras temporizadas (100, 110, 112, 114), em que cada amostra temporizada compreende uma pluralidade de subamostras espaciais (1a, 1b, 1c, 1d), em que a pluralidade de arquivos de segmento de mídia compreende pelo menos uma trilha de composição e pelo menos uma trilha de tile, cada uma da pelo menos uma trilha de tile compreende uma subamostra espacial selecionada dentre a pluralidade de subamostras espaciais de uma das subamostras temporizadas e que compreende uma subamostra espacial correspondente de cada uma dentre as outras amostras temporizadas, a pelo menos uma trilha de composição compreendendo

informação de identificação (412-1, 412-3, 412-4, 412-5, 412-6) que identifica pelo menos uma trilha de tile e que identifica pelo menos uma subamostra espacial da trilha de tile identificada, o método caracterizado por compreender:

selecionar (608) um item de informações representativo de uma subamostra espacial;

solicitar a pelo menos uma trilha de tile que compreende a subamostra espacial selecionada, a subamostra espacial selecionada correspondendo ao item de informações selecionado;

receber (610, 614) a pluralidade de arquivos de segmento de mídia; e

substituir (616, 716) a informação de identificação (412-1, 412-3, 412-4, 412-5, 412-6) da pelo menos uma trilha de composição por subamostras espaciais identificadas correspondentes para gerar o fluxo de bits de dados de mídia temporizada.

10. Método, de acordo com a reivindicação 9, caracterizado pelo fato de que compreende adicionalmente:

obter uma referência de trilha a partir de um extrator (412-1, 412-3, 412-4, 412-5, 412-6) da pelo menos uma trilha de composição;

inspecionar se a trilha corresponde ou não à referência de trilha obtida que foi recebida; e

se a trilha que corresponde à referência de trilha obtida não for recebida, remover o extrator correspondente para gerar o fluxo de bits de dados de mídia temporizada.

11. Método, de acordo com a reivindicação 10, caracterizado pelo fato de que compreende adicionalmente uma etapa de verificar que a trilha que corresponde à referência de trilha obtida é uma trilha de tipo de tile.

12. Método, de acordo com a reivindicação 9, caracterizado pelo fato de que compreende adicionalmente uma etapa de receber um arquivo de inicialização (300) que compreende parâmetros que permitem formação de uma porção espacial dos dados de mídia temporizada em tiles a partir das trilhas de tile criadas e a pelo menos uma trilha de composição criada.

13. Método, de acordo com a reivindicação 9, caracterizado pelo fato de que pelo menos dois arquivos de segmento de mídia representativos de pelo menos duas trilhas de tile são recebidos, em que a pelo menos uma trilha de composição compreende pelo menos um item de dados que deve ser usado para decodificar qualquer uma dentre duas subamostras espaciais que pertencem ao mesmo conjunto de subamostras espaciais.

14. Método, de acordo com a reivindicação 9, caracterizado pelo fato de que pelo menos um dos arquivos de segmento de mídia compreende uma estrutura de dados para armazenar metadados incluindo pelo menos uma definição de pelo menos um grupo de subamostras da trilha correspondente, o método compreende uma etapa de obter a pelo menos uma definição para processar o pelo menos um grupo de subamostras da trilha correspondente.

15. Método, de acordo com a reivindicação 9, caracterizado pelo fato de que pelo menos um dos arquivos de segmento de mídia compreende dados de inicialização armazenados na trilha de composição, o método compreende adicionalmente uma etapa de processar de modo independente o pelo menos um dos arquivos de segmento de mídia como um arquivo de segmento de mídia padrão.

16. Método, de acordo com a reivindicação 9, caracterizado pelo fato de que pelo menos um dos arquivos de segmento de mídia compreende uma estrutura de dados para armazenar informação de identificação para identificar dados de inicialização armazenados na trilha de composição, o método compreende adicionalmente uma etapa de acessar os dados armazenados na trilha de composição quando processa pelo menos um dos arquivos de segmento de mídia.

17. Meio de armazenamento legível por computador não transitório caracterizado pelo fato de ser codificado com instruções executáveis em computador que, quando executadas pelo computador, resultam no método definido na reivindicação 1.

18. Meio de gravação não transitório que contém instruções executáveis por computador para encapsular, usando um Formato de Arquivo de Mídia Base ISO, dados de mídia temporizada em tiles em um servidor, em que os dados de mídia

temporizada em tiles compreendem amostras temporizadas (100, 110, 112, 114), em que cada amostra temporizada compreende uma pluralidade de subamostras espaciais (1a, 1b, 1c, 1d), o meio caracterizado por compreender:

instruções executáveis por computador para criar (600), para cada uma das uma ou mais subamostras espaciais, uma trilha de tile que compreende a subamostra espacial, e uma subamostra espacial correspondente de cada uma dentre as outras amostras temporizadas;

instruções executáveis por computador para criar (602) pelo menos uma trilha de composição que compreende informação de identificação (412-1, 412-3, 412-4, 412-5, 412-6) para associar uma primeira subamostra espacial incluída em uma primeira trilha de tile e uma segunda subamostra espacial incluída em uma segunda trilha de tile; e

instruções executáveis por computador para encapsular (604) de modo independente cada trilha criada em pelo menos um arquivo de segmento de mídia (400-1, 400-2, 400-6).

19. Dispositivo (800) para encapsular, usando um Formato de Arquivo de Mídia Base ISO, dados de mídia temporizada em tiles em um servidor, em que os dados de mídia temporizada em tiles compreendem amostras temporizadas (100, 110, 112, 114), em que cada amostra temporizada compreende uma pluralidade de subamostras espaciais (1a, 1b, 1c, 1d), o dispositivo caracterizado por compreender pelo menos um microprocessador (804) configurado para realizar as etapas de:

para cada uma de uma ou mais subamostras espaciais, criar (600) uma trilha de tile que compreende a subamostra espacial, e uma subamostra espacial correspondente de cada uma dentre as outras amostras temporizadas;

criar (602) pelo menos uma trilha de composição que compreende informação de identificação (412-1, 412-3, 412-4, 412-5, 412-6) para identificar pelo menos uma das trilhas de tile criadas; e

encapsular de modo independente (604) cada uma das trilhas criadas em pelo menos um arquivo de segmento de mídia (400-1, 400-2, 400-6).

20. Dispositivo (800) para fornecer um fluxo de bits de dados de mídia temporizada em um Formato de Arquivo de Mídia Base ISO a partir de dados de mídia temporizada em tiles encapsulados em uma pluralidade de arquivos de segmento de mídia (400-1, 400-2, 400-6), em um dispositivo de cliente, em que os dados de mídia temporizada em tiles compreendem amostras temporizadas (100, 110, 112, 114), em que cada amostra temporizada compreende uma pluralidade de subamostras espaciais (1a, 1b, 1c, 1d), a pluralidade de arquivos de segmento de mídia compreende pelo menos uma trilha de composição e pelo menos uma trilha de tile, em que cada uma dentre a pelo menos uma trilha de tile compreende uma subamostra espacial selecionada dentre a pluralidade de subamostras espaciais de uma das subamostras temporizadas e compreende uma subamostra espacial correspondente de cada uma dentre as outras amostras temporizadas, em que a pelo menos uma trilha de composição compreende pelo menos informação de identificação (412-1, 412-3, 412-4, 412-5, 412-6) para identificar pelo menos uma trilha de tile e que identifica pelo menos uma subamostra espacial da trilha de tile identificada, o dispositivo caracterizado por compreender pelo menos um microprocessador configurado para realizar as etapas de:

selecionar (608) um item de informações representativo de uma subamostra espacial;

solicitar a pelo menos uma trilha de tile que compreende a subamostra espacial selecionada, em que a subamostra espacial selecionada corresponde ao item de informações selecionado;

receber (610, 614) a pluralidade de arquivos de segmento de mídia; e

substituir (616, 716) informação de identificação (412-1, 412-3, 412-4, 412-5, 412-6) da pelo menos uma trilha de composição por subamostras espaciais identificadas correspondentes para gerar o fluxo de bits de dados de mídia temporizada.

21. Método, de acordo com a reivindicação 1, caracterizado pelo fato de que a trilha de composição compreende informação de identificação (412-1, 412-3, 412-4, 412-5, 412-6) para associar uma primeira subamostra espacial incluída em

uma primeira trilha de tile e uma segunda subamostra espacial incluída em uma segunda trilha de tile.

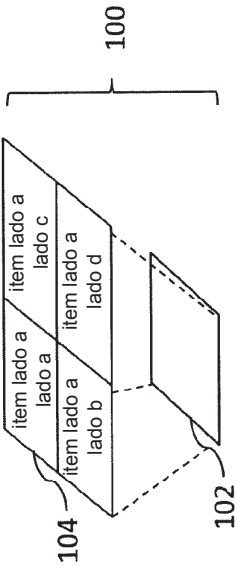
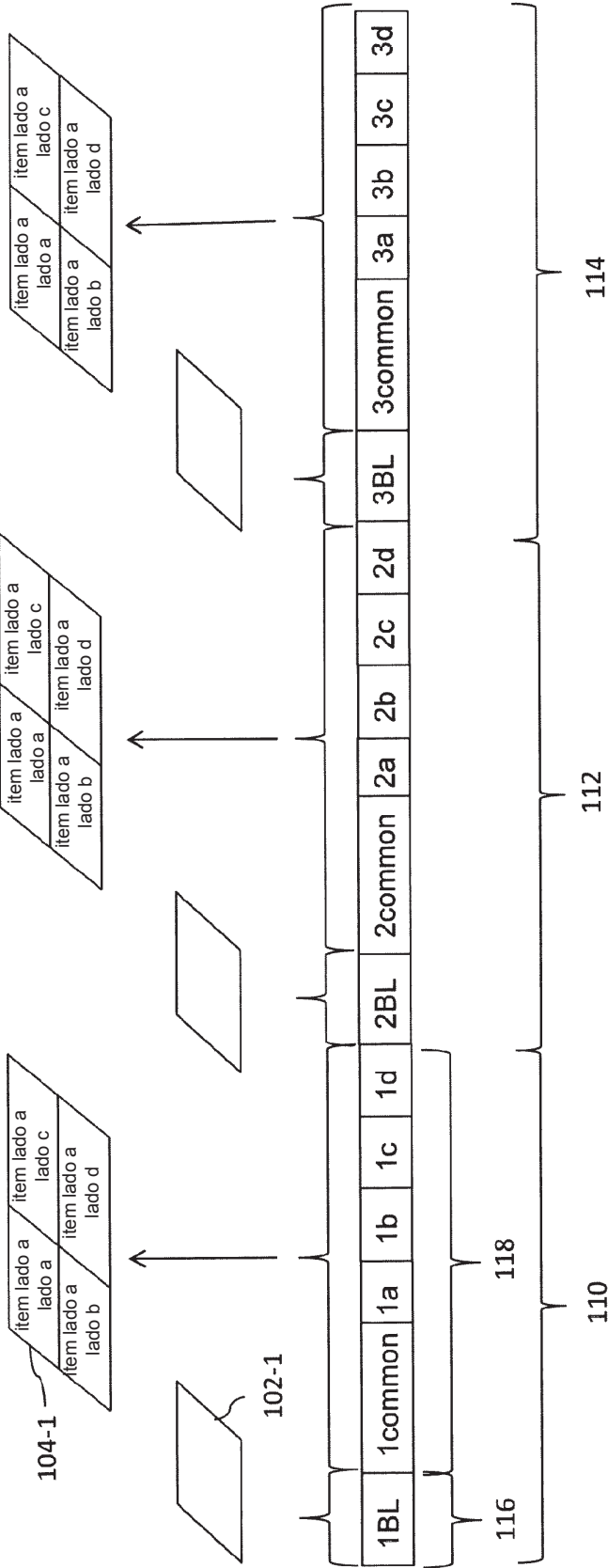


Fig. 1a

Fig. 1b



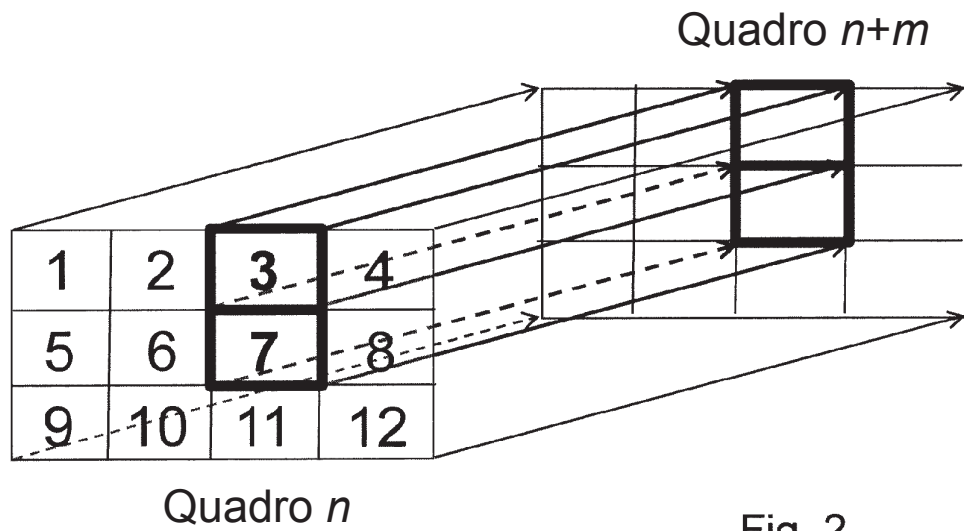


Fig. 2

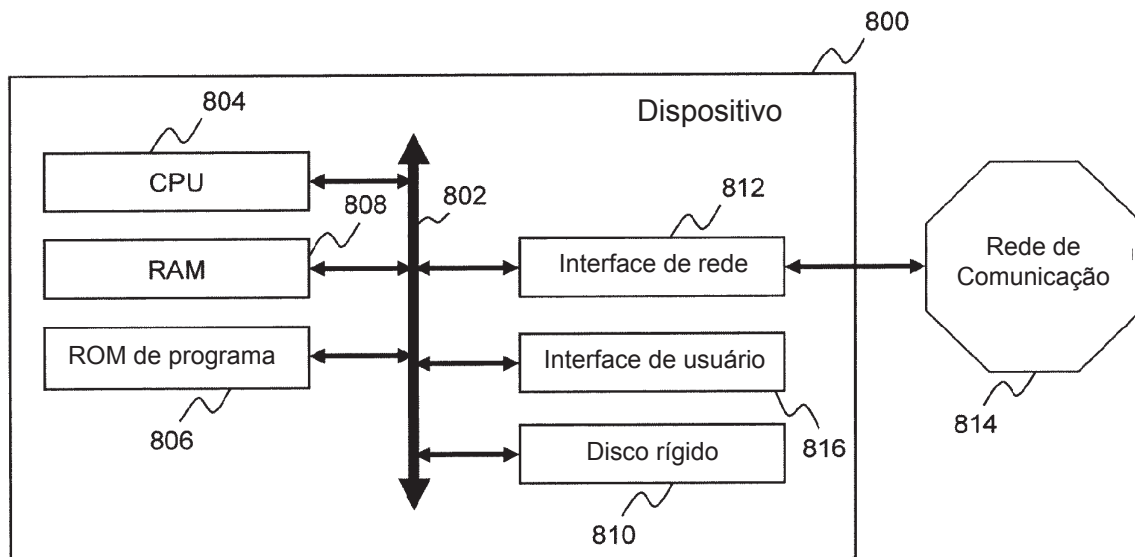


Fig. 8

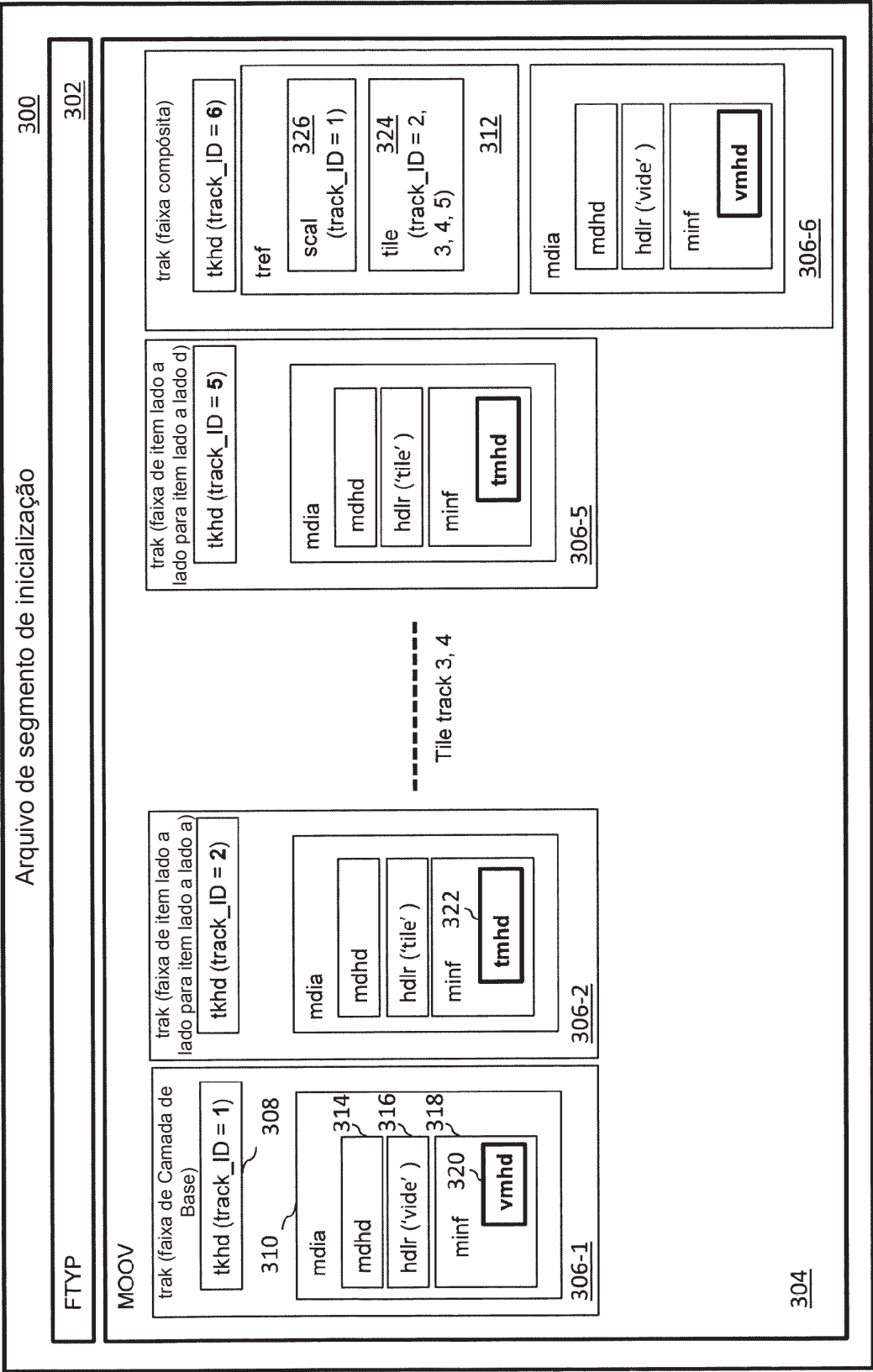


Fig. 3

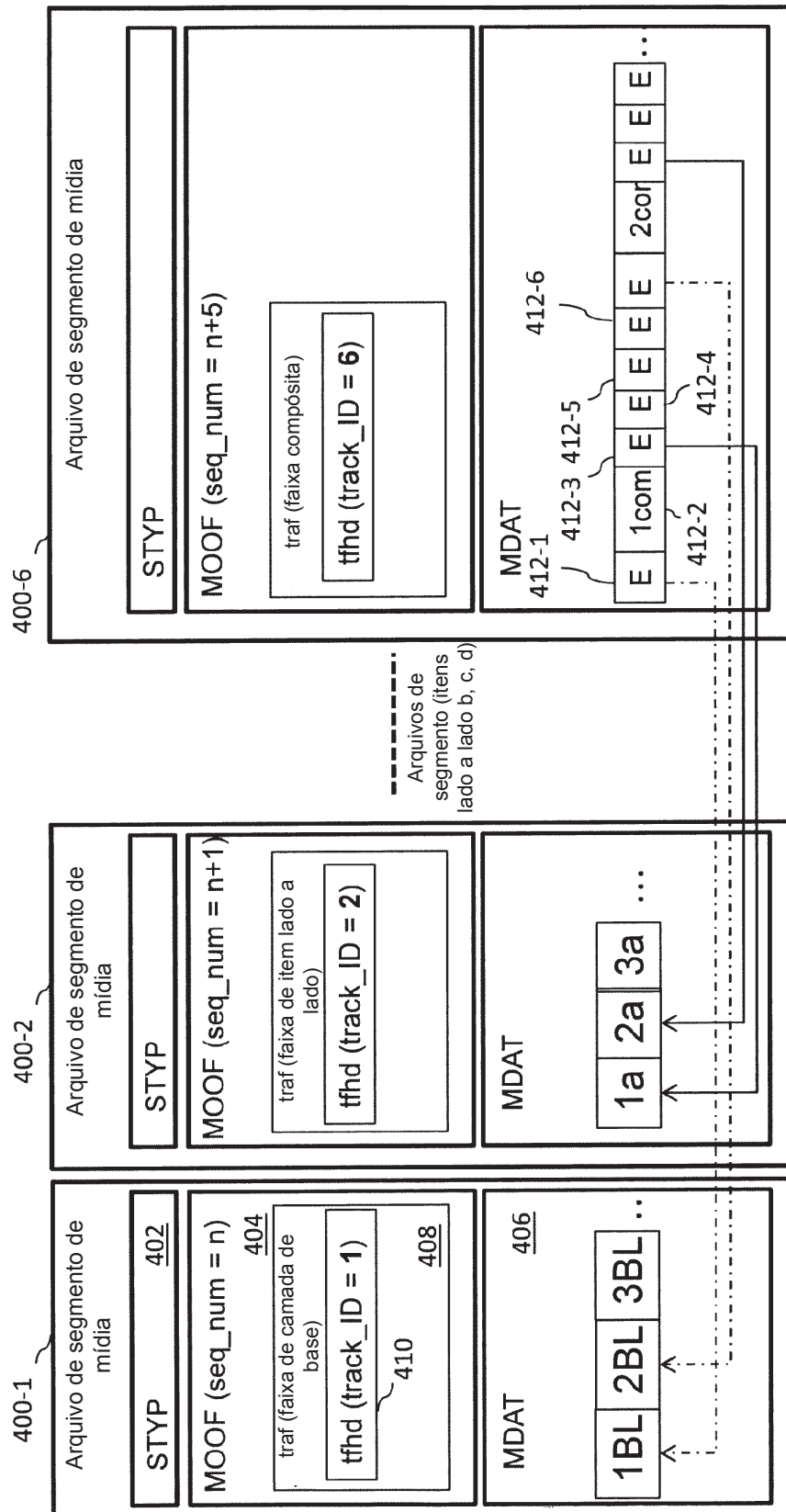
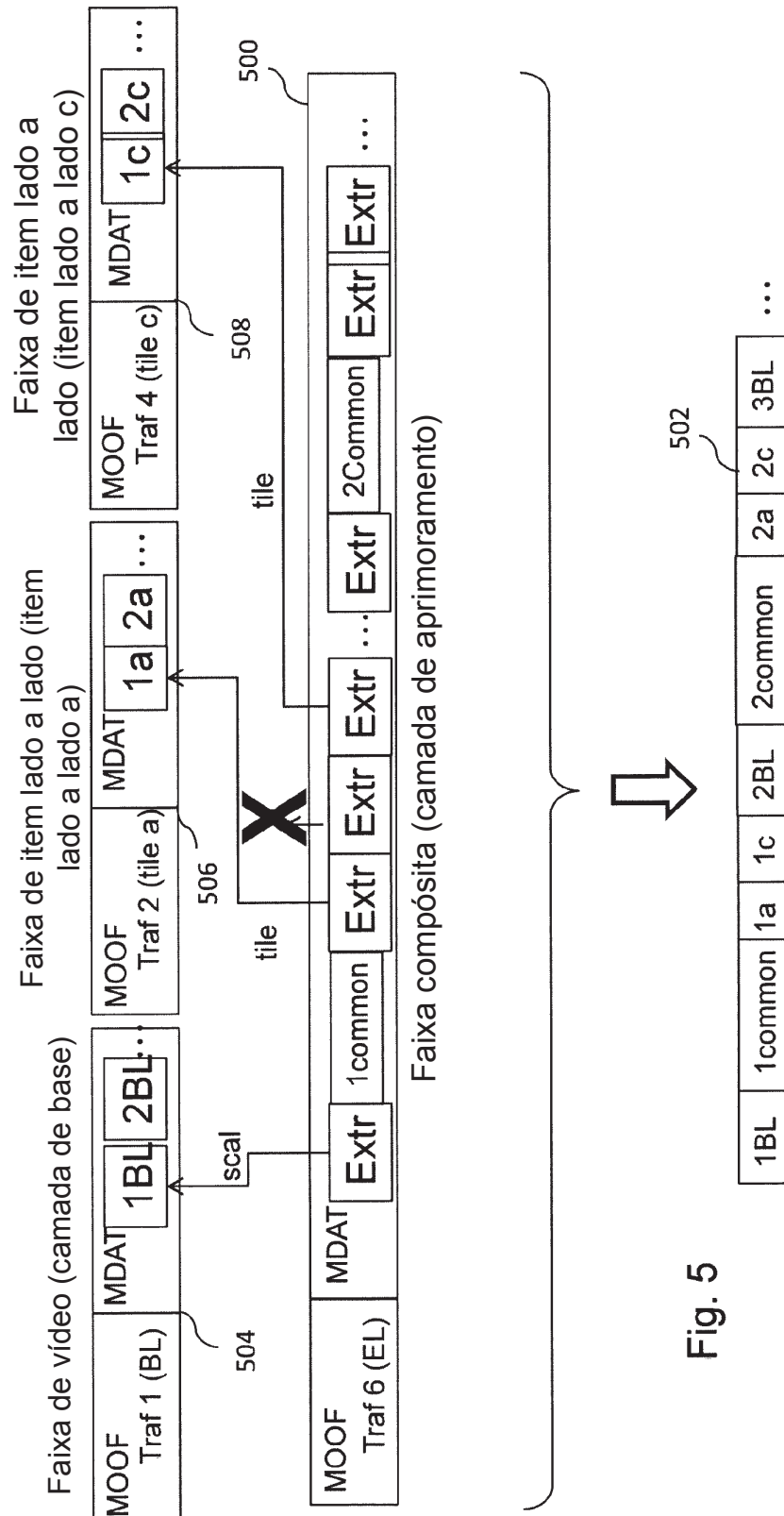


Fig. 4



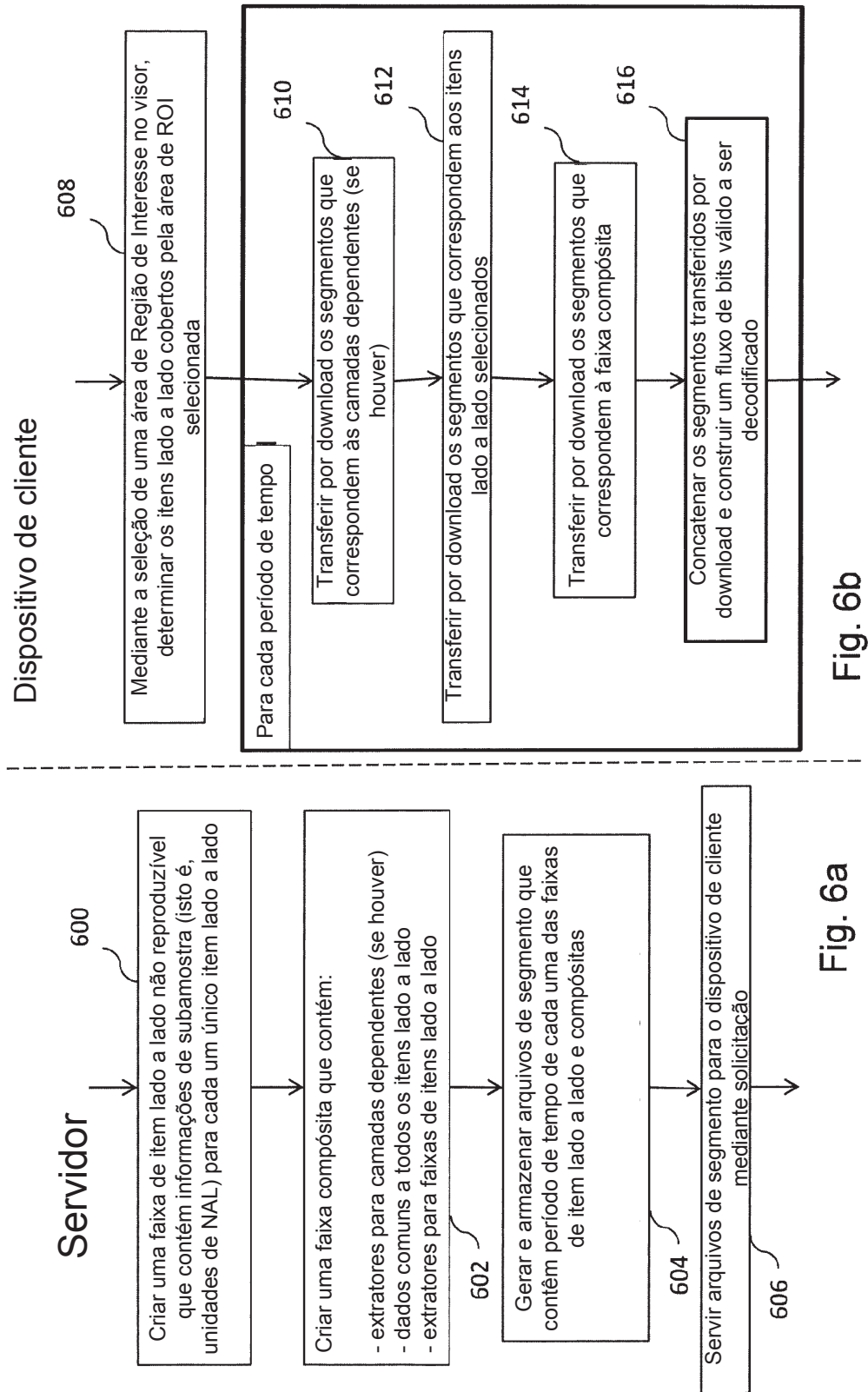


Fig. 6b

Fig. 6a

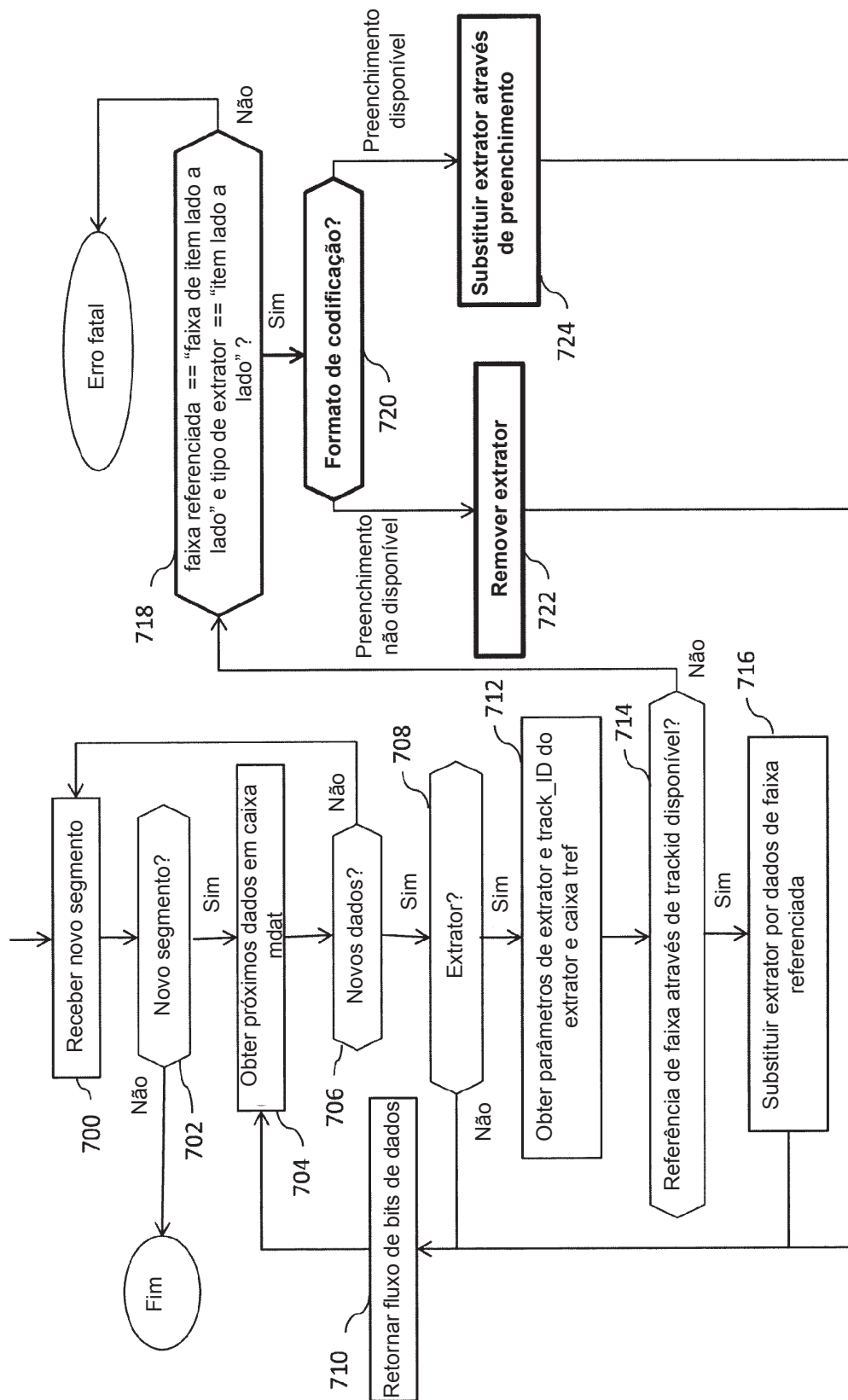


Fig. 7

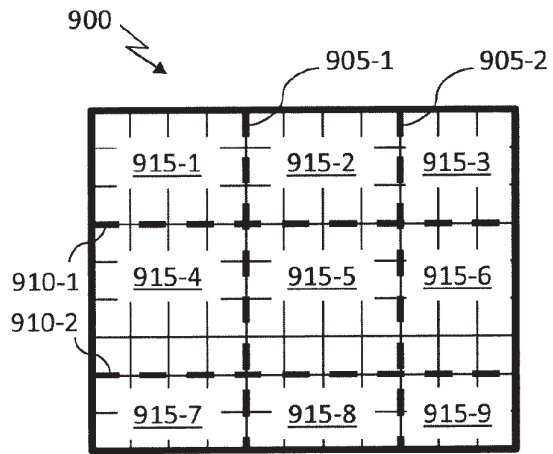


Fig. 9a

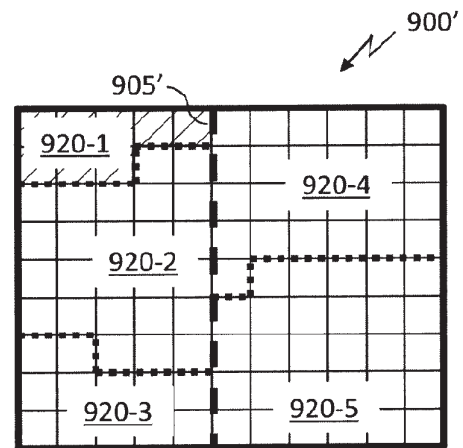


Fig. 9b

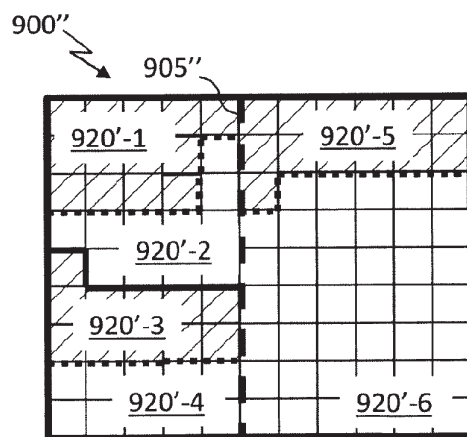


Fig. 9c

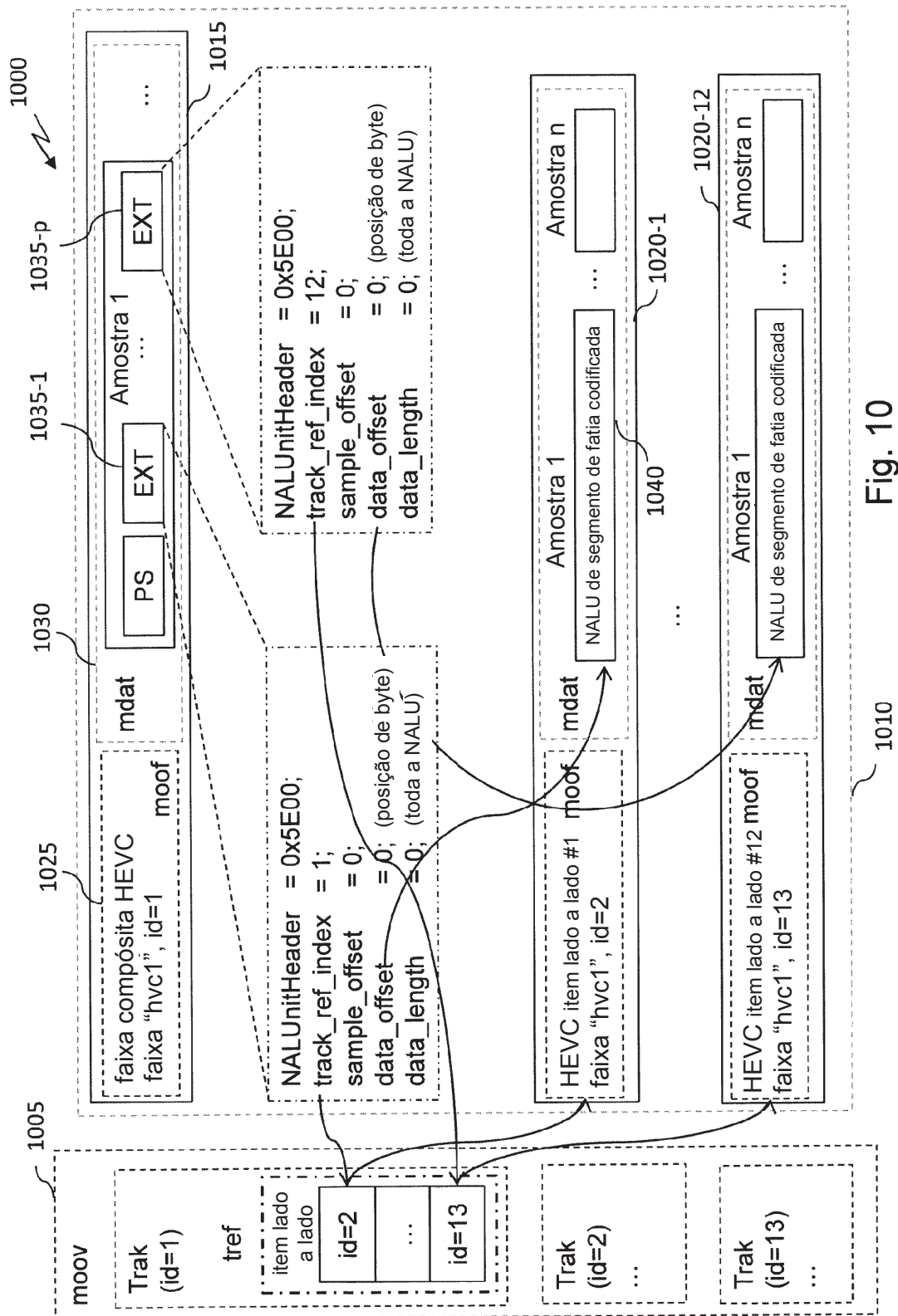


Fig. 10

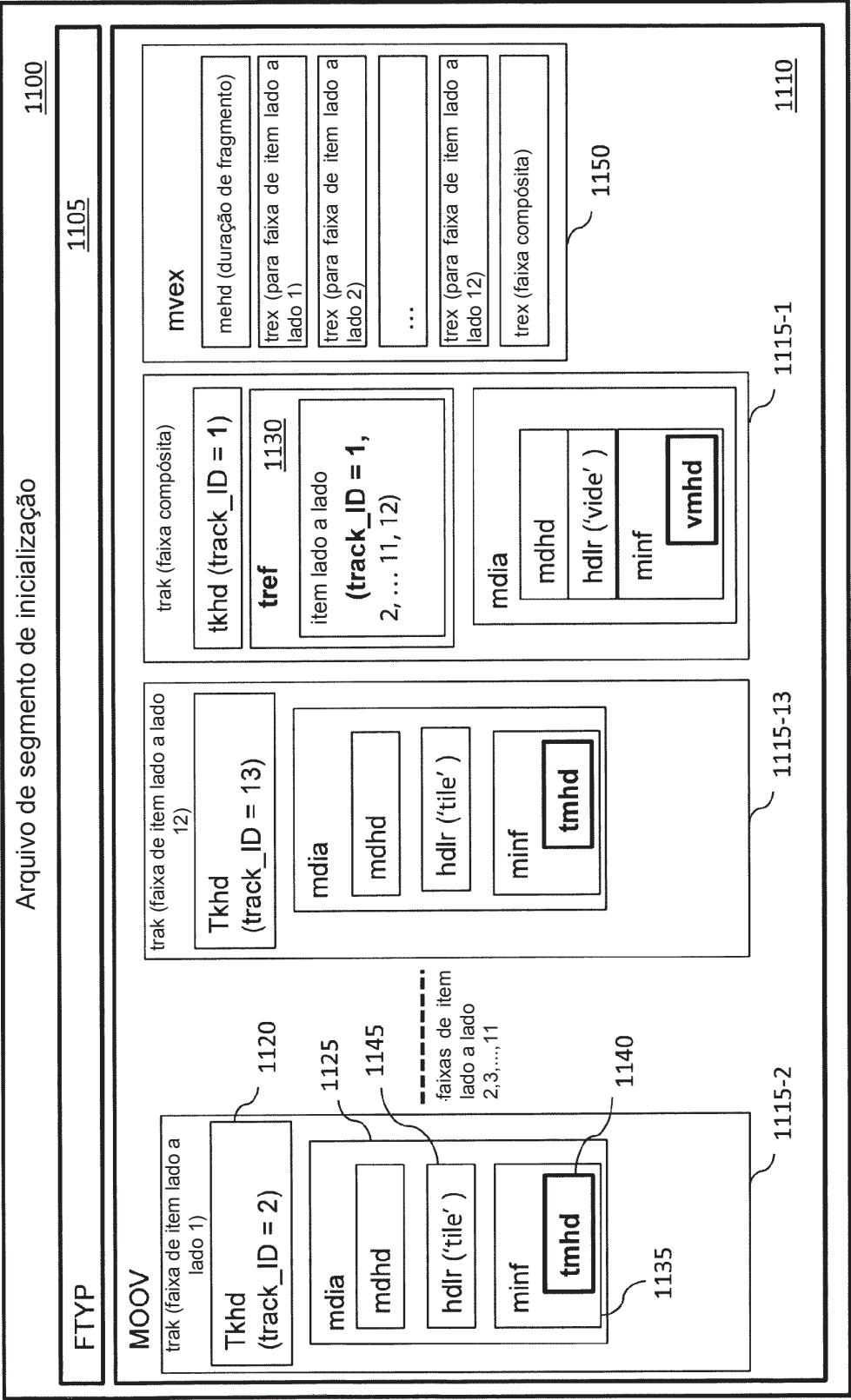


Fig. 11

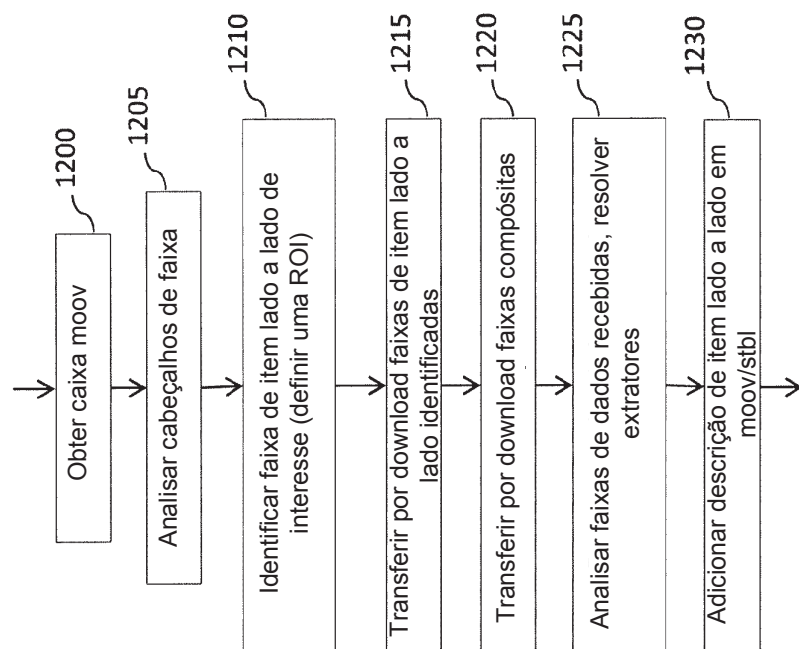


Fig. 12a

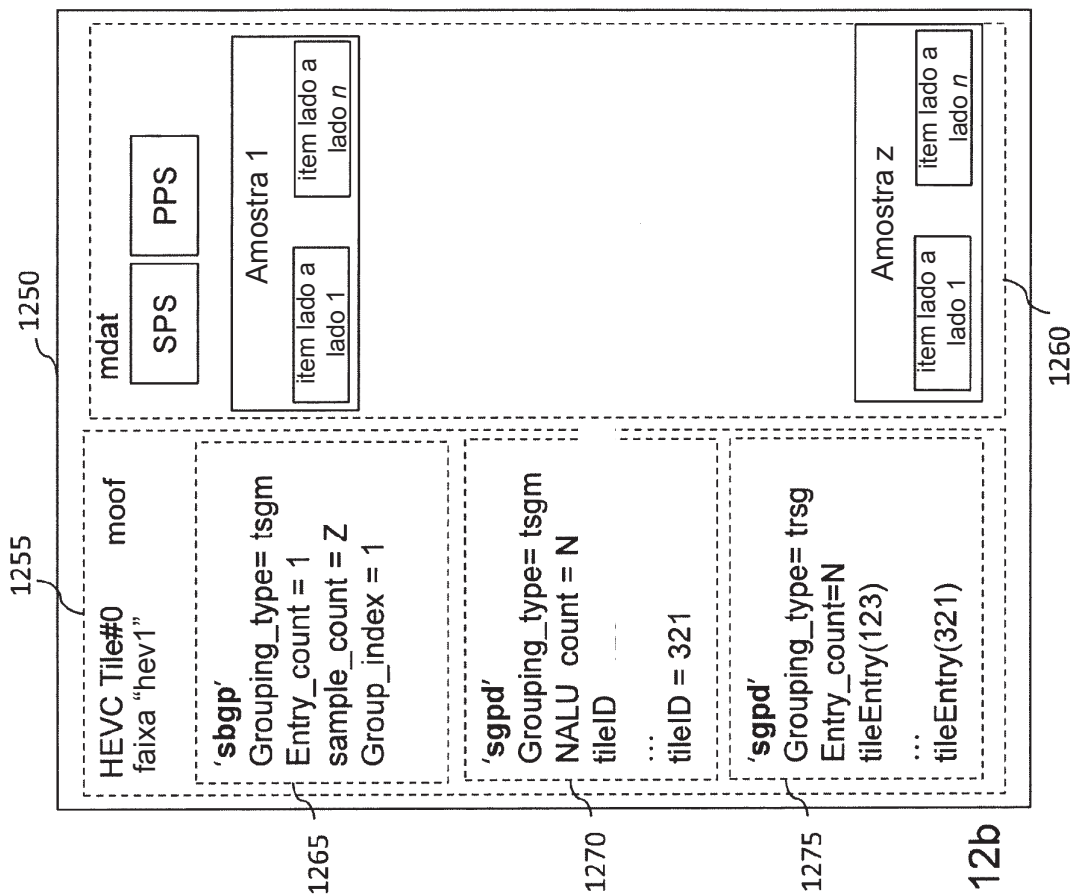


Fig. 12b

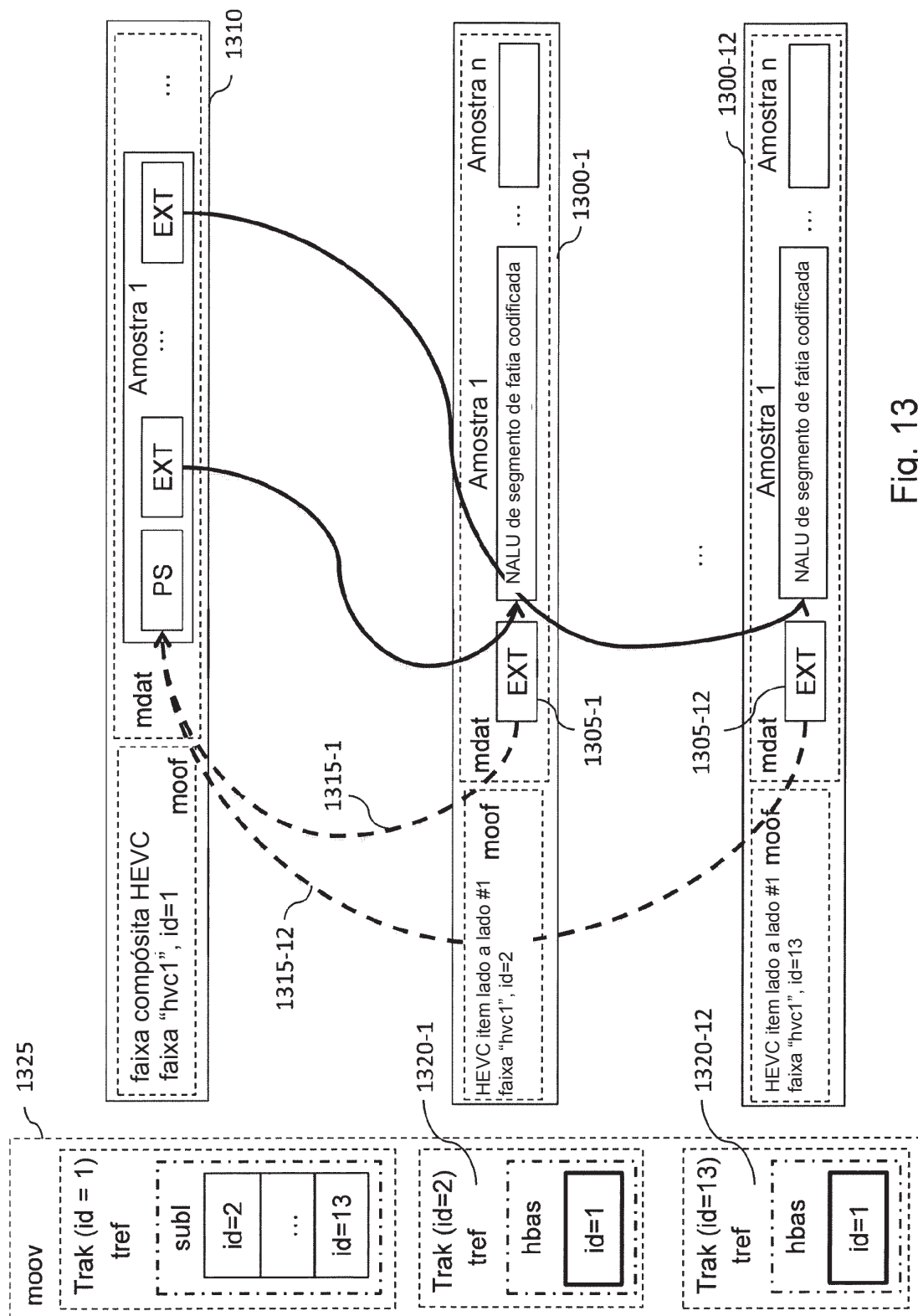


Fig. 13

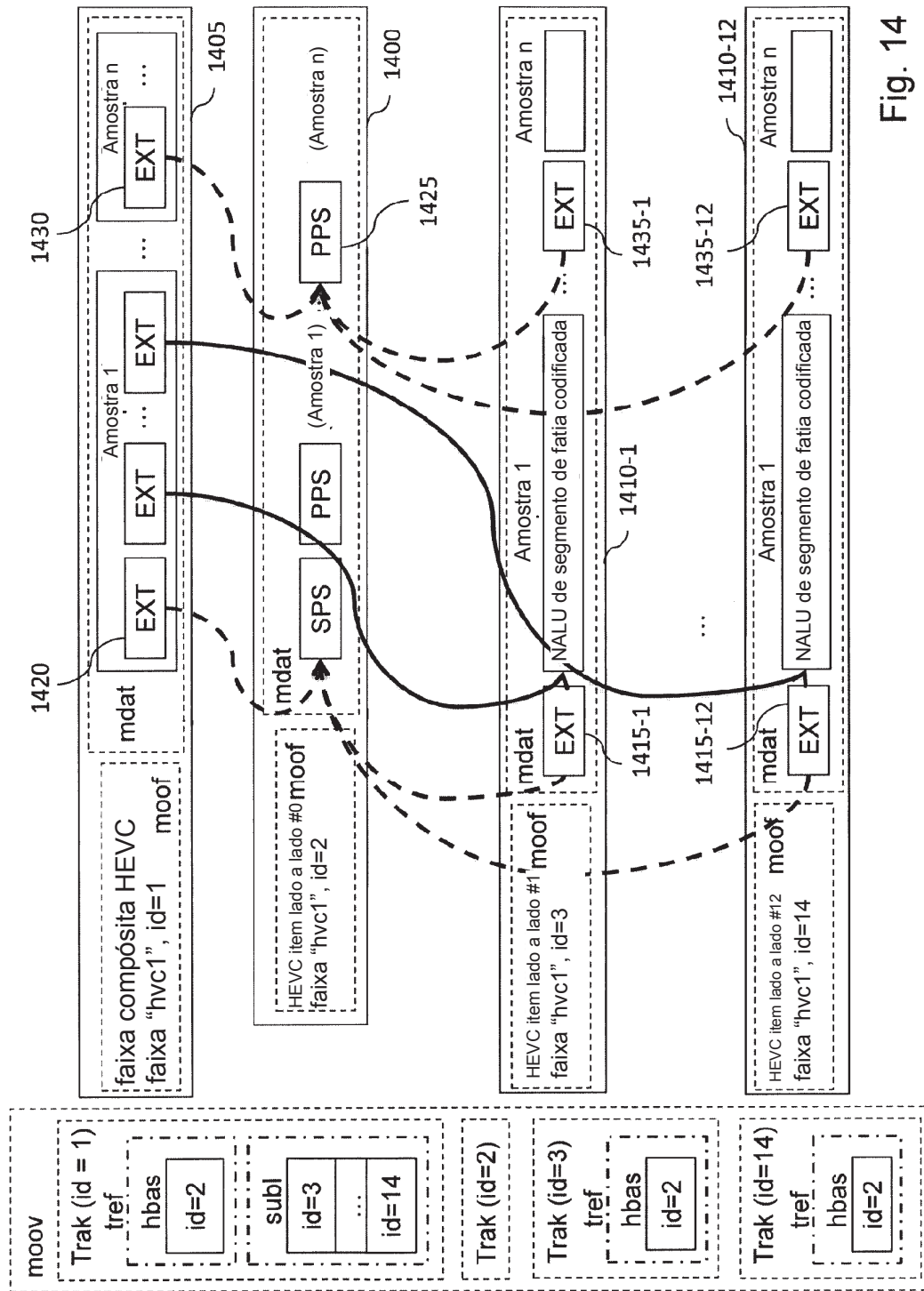


Fig. 14