



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

(11) BR 112016028586-7 B1



(22) Data do Depósito: 10/06/2015

(45) Data de Concessão: 20/12/2022

(54) Título: MEIO DE ARMAZENAMENTO DE COMPUTADOR E MÉTODO IMPLEMENTADO POR COMPUTADOR PARA ANÁLISE DE IMPORTÂNCIA DE VÍDEO BASEADO EM REGRAS

(51) Int.Cl.: G06K 9/00; G06F 17/30.

(30) Prioridade Unionista: 12/06/2014 US 14/303,466.

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

(72) Inventor(es): NITIN SURI; TZONG-JHY WANG; OMKAR MEHENDALE; ANDREW S. IVORY; WILLIAM D. SPROULE.

(86) Pedido PCT: PCT US2015034992 de 10/06/2015

(87) Publicação PCT: WO 2015/191650 de 17/12/2015

(85) Data do Início da Fase Nacional: 06/12/2016

(57) Resumo: ANÁLISE DE IMPORTÂNCIA DE VÍDEO BASEADO EM REGRAS. A importância de seções de vídeo de um arquivo de vídeo pode ser determinada de características do arquivo de vídeo. O arquivo de vídeo pode ser decodificado para obter quadros de vídeo e dados de áudio associados com os quadros de vídeo. Pontuações de característica para cada quadro de vídeo podem ser obtidas analisando as características do quadro de vídeo ou dos dados de áudio associados com o quadro de vídeo com base em uma regra local, uma regra global, ou ambas. As pontuações de característica são ainda combinadas para derivar uma pontuação de importância de quadro para o quadro de vídeo. Com base nas pontuações de característica dos quadros de vídeo no arquivo de vídeo, o arquivo de vídeo pode ser segmentado em seções de vídeo de diferentes valores de importância de seção.

Relatório Descritivo da Patente de Invenção para "MEIO DE ARMAZENAMENTO DE COMPUTADOR, MÉTODO IMPLEMENTADO POR COMPUTADOR E SISTEMA PARA ANÁLISE DE IMPORTÂNCIA DE VÍDEO BASEADO EM REGRAS".

ANTECEDENTES

[001] Os consumidores frequentemente capturam vídeos utilizando seus smartphones e gravadores de vídeo pessoal. No entanto, somente uma pequena percentagem destes consumidores edita e compartilha seus vídeos com outras pessoas. Ainda, um consumidor pode achar a edição de tais vídeos ser um processo tedioso, já que os vídeos são geralmente feitos em um modo casual sem muito planejamento, e pode conter somente poucos momentos interessantes. Conforme mais vídeos são gerados por um consumidor ao longo do tempo, o consumidor pode também ter dificuldade em lembrar o conteúdo dos vídeos. Apesar da maioria dos dispositivos de reprodução de vídeo poderem prover representações de imagem em miniatura dos vídeos, tais representações de imagem em miniatura podem não prover pistas suficientes sobre o conteúdo dos vídeos.

SUMÁRIO

[002] Descritas aqui estão técnicas para executar uma análise baseada em regras de um arquivo de vídeo para classificar seções do arquivo de vídeo com base em sua importância. As técnicas podem também incluir executar uma análise baseada em regras de uma coleção de arquivos de vídeo para classificar múltiplos arquivos de vídeo com base em sua importância. A importância de um arquivo de vídeo ou uma seção de vídeo pode correlacionar com a quantidade de interesse que o arquivo de vídeo ou a seção de vídeo é esperado gerar em um espectador. Em várias modalidades, a análise baseada em regras pode certificar a importância de um arquivo de vídeo ou a

seção de vídeo com base na importância subjetiva e/ou a qualidade técnica objetiva dos quadros de vídeo no arquivo de vídeo ou na seção de vídeo.

[003] Este Sumário está provido para introduzir uma seleção de conceitos em uma forma simplificada que estão adicionalmente abaixo descritos na Descrição Detalhada. Este Sumário não pretende identificar características chave ou características essenciais do assunto reivindicado, nem pretende ser utilizado para limitar o escopo do assunto reivindicado.

BREVE DESCRIÇÃO DOS DESENHOS

[004] A descrição detalhada está descrita com referência às figuras acompanhantes. Nas figuras, o(s) dígito(s) mais à esquerda de um número de referência identifica a figura na qual o número de referência primeiro aparece. A utilização do mesmo número de referência em diferentes figuras indica itens similares ou idênticos.

[005] Figura 1 é um diagrama de blocos que ilustra um esquema exemplar para utilizar uma máquina de análise de vídeo baseada em regras para classificar seções de vídeo de um arquivo de vídeo ou arquivos de vídeo em uma coleção de arquivos de vídeo com base em sua importância.

[006] Figura 2 é um diagrama ilustrativo que mostra componentes exemplares de uma máquina de análise de vídeo baseada em regras para classificar seções de vídeo de um arquivo de vídeo ou arquivos de vídeo em uma coleção de arquivos de vídeo com base em sua importância.

[007] Figura 3 é um diagrama ilustrativo que mostra a utilização de uma transformada de homógrafo para alinhar pontos de característica exemplares em múltiplos quadros de vídeo.

[008] Figura 4 é um fluxograma que ilustra um processo exemplar para utilizar análise de vídeo baseada em regras para analisar

características em um arquivo de vídeo de modo a classificar seções de vídeo do arquivo de vídeo com base em sua importância.

[009] Figura 5 é um fluxograma que ilustra um processo exemplar para utilizar a análise de vídeo baseada em regras para analisar as características de arquivos de vídeo para classificar os arquivos de vídeo com base em importância.

[0010] Figura 6 é um fluxograma que ilustra um processo exemplar para computar uma pontuação de importância de face para um quadro de vídeo.

[0011] Figura 7 é um fluxograma que ilustra um processo exemplar para determinar importantes seções de vídeo dentro de um vídeo analisando o movimento de pontos de característica.

DESCRIÇÃO DETALHADA

[0012] Descritas aqui estão técnicas para executar uma análise baseada em regras de um arquivo de vídeo para classificar as seções do arquivo de vídeo com base em sua importância. As técnicas podem também incluir executar uma análise baseada em regras de uma coleção de arquivos de vídeo para classificar os arquivos de vídeo com base em sua importância. A importância de um arquivo de vídeo ou uma seção de vídeo pode correlacionar com a quantidade de interesse que o arquivo de vídeo ou a seção de vídeo é esperado gerar em um espectador. Em várias modalidades, a análise baseada em regras pode certificar a importância de um arquivo de vídeo ou uma seção de vídeo com base na importância subjetiva e/ou na qualidade técnica objetiva dos quadros de vídeo no arquivo de vídeo ou na seção de vídeo. Um exemplo de importância subjetiva pode ser que um espectador considerará um quadro de vídeo que apresenta uma face de uma pessoa que é conhecida para o espectador como mais importante do que um quadro de vídeo que apresenta a face de uma pessoa que o espectador não conhece. Um exemplo de

qualidade técnica objetiva pode ser a qualidade de exposição do quadro de vídeo.

[0013] A análise baseada em regras compreende analisar o conteúdo de áudio e o conteúdo de vídeo dos arquivos de vídeo por múltiplas características de baixo nível e características de alto nível em uma base quadro por quadro. Por exemplo, as características de baixo nível podem incluir características tais como qualidade de exposição, qualidade de saturação, instabilidade de quadros de vídeo, luminosidade média, entropia de cor, e/ou diferenças de histograma entre quadros de vídeo adjacentes. As características de alto nível podem incluir características tais como as quantidades, posições, e/ou características faciais de faces humanas que são detectadas nos quadros de vídeo. A análise pode incluir a aplicação de regras locais e regras globais. As regras locais podem ser aplicadas durante a geração de resultados de análise de característica para um quadro de vídeo, e as regras globais podem ser aplicadas durante a geração de resultados de análise de característica para um arquivo de vídeo inteiro.

[0014] As regras podem prover e combinar os resultados da análise de característica para gerar pontuações de importância. As pontuações de importância podem ser geradas para quadros de arquivos de vídeo, seções de vídeos, e/ou arquivos de vídeo em suas totalidades. Estas pontuações de importância podem ser utilizadas para classificar seções de arquivos de vídeo e/ou arquivos de vídeo. As pontuações de importância podem ser utilizadas para facilitar a visualização, edição, e compartilhamento de arquivos de vídeo. Por exemplo, um consumidor pode selecionar um conjunto de arquivos de vídeo com as mais altas pontuações de importância para compartilhar em um website. Em outro exemplo, uma aplicação pode costurar juntas seções de um arquivo de vídeo com as mais altas pontuações

de importância para criar um arquivo de vídeo de destaque.

[0015] Em alguns casos, um arquivo de vídeo pode ser segmentado em seções de vídeo com variáveis graus de importância com base na quantidade de movimento de câmera detectada. O movimento de câmera pode ser detectado através do movimento de pontos de característica que são detectadas nos quadros de vídeo no arquivo de vídeo. Em alguns casos, a importância de uma seção de vídeo pode correlacionar com a quantidade de interesse que a seção de vídeo é esperada evocar dos espectadores.

[0016] A utilização das técnicas aqui descritas pode permitir um usuário classificar arquivos de vídeo com base em sua importância para o usuário. Com base em tal classificação, o usuário pode decidir quais arquivos de vídeo manter e quais arquivos de vídeo apagar. Em alguns casos, o usuário pode também utilizar as classificações dos arquivos de vídeo para determinar se postar arquivos de vídeo específicos em um website de compartilhamento online. As técnicas aqui descritas podem também apresentar representações de imagem em miniatura que representam seções de importância de um arquivo de vídeo, de modo que o usuário possa dizer em um olhar as porções interessantes de um arquivo de vídeo. Tais informações podem ajudar o usuário em editar o arquivo de vídeo para aperfeiçoar a qualidade de conteúdo ou destacar seções específicas do arquivo de vídeo. Exemplos de técnicas para executar uma análise baseada em regras de arquivos de vídeo de acordo com várias modalidades estão abaixo descritos com referência às Figuras 1-7.

ESQUEMA EXEMPLAR

[0017] A Figura 1 é um diagrama de blocos que ilustra um esquema exemplar 100 para utilizar uma máquina de análise de vídeo baseada em regras para classificar seções de vídeo de um arquivo de vídeo ou arquivos de vídeo em uma coleção de arquivos de vídeo com

base em importância. O esquema exemplar 100 pode incluir uma máquina de análise de vídeo 102. A máquina de análise de vídeo 102 pode ser executada em um ou mais dispositivos de computação 104. Os um ou mais dispositivos de computação 104 podem incluir computadores de uso geral, tal como computadores desktop, computadores tablet, computadores laptop, servidores, e assim por diante. No entanto, em outras modalidades, os dispositivos de computação 104 podem incluir smartphones, consoles de jogos, ou quaisquer outros dispositivos eletrônicos. Os múltiplos dispositivos de computação 104 podem incluir vários processadores, tal como as unidades de processador central (CPUs), unidades de processador gráfico (GPUs), outros tipos de processadores, ou uma combinação de quaisquer dos acima.

[0018] A máquina de análise de vídeo 102 pode executar uma análise baseada em regras de uma coleção de vídeos 106. A coleção de vídeos 106 pode incluir múltiplos arquivos de vídeo, tal como os arquivos de vídeo 108(1)-108(N). A análise baseada em regras pode compreender analisar o conteúdo de áudio e o conteúdo de vídeo dos arquivos de vídeo 108(1)-108(N) para múltiplas características de baixo nível 110 e múltiplas características de alto nível 112 em uma base quadro por quadro. Por exemplo, as múltiplas características de baixo nível 110 podem incluir características tais como qualidade de exposição, qualidade de saturação, e instabilidade de quadros de vídeo. As múltiplas características de alto nível 112 podem incluir características tais como as quantidades, posições, e características faciais de faces humanas que são detectadas nos quadros de vídeo.

[0019] Executando a análise baseada em regras, a máquina de análise de vídeo 102 pode gerar pontuações de importância para seções de um arquivo de vídeo, tal como o vídeo 108(1), e pontuações de importância para arquivos de vídeo, tal como arquivos de vídeo

108(1)-108(N) da coleção de vídeos 106. Consequentemente, a máquina de análise de vídeo 102 pode classificar seções de um vídeo de acordo com suas pontuações de importância. Por exemplo, o arquivo de vídeo 108(1) pode incluir uma seção de vídeo 114 e uma seção de vídeo 116 que são classificadas de acordo com suas pontuações de importância, tal como as pontuações de importância 118 e 120, respectivamente. Uma vez que as seções de vídeo são classificadas, a máquina de análise de vídeo 102 pode exibir representações de imagem em miniatura das seções de vídeo, nas quais uma seleção de uma de representação de imagem em miniatura pode fazer com que um reproduutor de mídia 122 execute a seção de vídeo correspondente.

[0020] A máquina de análise de vídeo 102 pode também classificar os arquivos de vídeo na coleção de vídeos 106 de acordo com suas pontuações de importância. Por exemplo, os arquivos de vídeo 108(1)- 108(N) da coleção de vídeos 106 podem ser classificados de acordo com suas pontuações de importância 124(1)-124(N). Uma vez que os arquivos de vídeo são classificados, a máquina de análise de vídeo 102 pode exibir representações de imagem em miniatura dos arquivos de vídeo, nas quais uma seleção de uma representação de imagem em miniatura pode fazer com que o reproduutor de mídia 122 execute arquivo de vídeo correspondente ou uma seção do arquivo de vídeo correspondente.

COMPONENTES EXEMPLARES

[0021] A Figura 2 é um diagrama ilustrativo que mostra componentes exemplares de uma máquina de análise de vídeo baseada em regras 102 para classificação de seções de vídeo de um arquivo de vídeo ou arquivos de vídeo em uma coleção de arquivos de vídeo com base em sua importância. A máquina de análise de vídeo 102 pode ser implementada por um ou mais dispositivos de

computação 104. Os dispositivos de computação 104 podem incluir um ou mais processadores 202, interfaces 204, e memória 206. Cada um dos processadores 202 pode ser um processador de núcleo único ou um processador de múltiplos núcleos. As interfaces 204 podem incluir interfaces de usuário e interfaces de rede. As interfaces de usuário podem incluir um dispositivo de saída de dados (por exemplo, display visual, alto-falantes de áudio), e um ou mais dispositivos de entrada de dados. Os dispositivos de entrada de dados podem incluir, mas não estão limitados a, combinações de um ou mais de teclados compactos, teclados, dispositivos de mouse, telas de toque que aceitam gestos, microfones, dispositivos de reconhecimento de voz ou fala, e quaisquer outros dispositivos adequadas ou outros métodos de seleção eletrônica / de software.

[0022] A interface de rede pode incluir componentes de interface de comunicação com fio e/ou sem fio que permitem os dispositivos de computação 104 transmitir e receber dados através de uma rede. Em várias modalidades, o componente de interface sem fio pode incluir, mas não está limitado a celular, Wi-Fi, Ultrabanda larga (UWB), redes de área pessoal (por exemplo, Bluetooth), transmissões de satélite, e/ou assim por diante. O componente de interface com fio pode incluir uma interface I/O direta, tal como uma interface Ethernet, uma interface serial, uma interface de Barramento Seria Universal (USB), e /ou assim por diante. Como tal, os dispositivos de computação 104 podem ter capacidades de rede. Por exemplo, os dispositivos de computação 104 podem trocar dados com outros dispositivos eletrônicos (por exemplo, computadores laptop, computadores desktop, servidores de telefones móveis, etc.) através de uma ou mais redes, tal como a Internet, redes móveis, redes de área a ampla, redes de área local, e assim por diante.

[0023] A memória 206 pode ser implementada utilizando um meio

legível por computador, tal como um meio de armazenamento de computador. O meio legível por computador inclui, pelo menos, dois tipos de meios legíveis por computador, a saber um meio de armazenamento de computador e meio de comunicação. O meio de armazenamento de computador inclui um meio volátil e não volátil, removível e não removível implementado em qualquer método ou tecnologia para armazenamento de informações tal como instruções legíveis por computador, estruturas de dados, módulos de programa, ou outros dados. O meio de armazenamento de computador inclui, mas não está limitado a, RAM, ROM, EEPROM, memória instantânea ou outra tecnologia de memória, CD-ROM, discos versáteis digitais (DVD) ou outro armazenamento ótico, cassetes magnéticos, fita magnética, armazenamento de disco magnético, ou outros dispositivos de armazenamento magnético, ou qualquer outro meio não de transmissão que pode ser utilizado para armazenar informações para acesso por um dispositivo de computação. Em contraste, o meio de comunicação pode incorporar instruções legíveis por computador, estruturas de dados, módulos de programa, ou outros dados em um sinal de dados modulado, tal como uma onda portadora, ou outro mecanismo de transmissão. Como aqui definido, meio de armazenamento de computador não inclui meio de comunicação.

[0024] A memória 206 dos dispositivos de computação 104 pode armazenar um sistema de operação 208 e módulos que implementar a máquina de análise de vídeo 102. O sistema de operação 208 pode incluir componentes que permitem que os dispositivos de computação 104 recebam dados através de várias entradas (por exemplo, controles de usuário, interfaces de rede, e/ou dispositivos de memória), e processar os dados utilizando os processadores 202 para gerar uma saída. O sistema de operação 208 pode ainda incluir um ou mais componentes que apresentam a saída (por exemplo, exibam uma

imagem em um display eletrônico, armazenem dados na memória, transmitam dados para outro dispositivo eletrônico, etc.). O sistema de operação 208 pode permitir um usuário interagir com módulos da máquina de análise de vídeo 102 utilizando a interface 204. Além disso, o sistema de operação 208 pode incluir outros componentes que executam várias outras funções geralmente associadas com um sistema de operação.

[0025] Os módulos podem incluir um módulo de decodificador de vídeo 210, um módulo de análise de baixo nível 212, um módulo de análise de alto nível 214, um módulo de análise de movimento 216, um módulo de cálculo de importância 218, um módulo de segmentação de vídeo 220, um módulo de classificação de vídeo 222, e um módulo de interface de usuário 224. Cada um dos módulos pode incluir rotinas, instruções de programas, objetos, e/ou estruturas de dados que executam tarefas específicas ou implementam tipos de dados abstratos específicos. Além disso, um armazenamento de dados 226 pode residir na memória 206. Cada um do módulo de análise de baixo nível 212 e do módulo de análise de alto nível 214 pode aplicar regras locais ou regras globais para analisar a importância de dados de característica, isto é, dados de vídeo e/ou dados de áudio, em arquivos de vídeo. Uma regra local pode afetar a atribuição de importância para um único quadro de vídeo com base nos dados de característica no único quadro de vídeo. Ao contrário, uma regra global pode afetar a atribuição de importância para um arquivo de vídeo inteiro com base nos dados de característica nos múltiplos quadros de vídeo do arquivo de vídeo, ou afetar a atribuição de importância para cada um de um conjunto de quadros de vídeo no arquivo de vídeo com base nas características que são compartilhadas através do conjunto de quadros de vídeo.

[0026] O módulo de decodificador de vídeo 210 pode decodificar

um arquivo de vídeo, tal como o arquivo de vídeo 108(1), para obter quadros de vídeo e/ou dados de áudio associados com cada um dos quadros de vídeo. O módulo de decodificador de vídeo 210 pode utilizar vários codecs para decodificar arquivos de vídeo, tal como H.264, MPEG-2, MPEG-4, etc.

[0027] O módulo de análise de baixo nível 212 pode analisar cada de quadro de vídeo decodificado por características de baixo nível para produzir pontuações de característica. Em várias modalidades, as características de baixo nível podem incluir qualidade de exposição, qualidade de saturação, variedade de tonalidade, instabilidade, luminosidade média, entropia de cor, e/ou diferenças de histograma entre quadros de vídeo adjacentes. O módulo de análise de baixo nível 212 pode utilizar algoritmos para derivar histogramas que mostram a exposição, saturação, e tonalidade de quadros de vídeo. Na análise de qualidade de exposição, o módulo de análise de baixo nível 212 pode analisar um histograma do equilíbrio de exposição de um quadro de vídeo. O módulo de análise de baixo nível 212 pode atribuir uma pontuação de classificação de exposição para o quadro de vídeo com base no equilíbrio de exposição de acordo com uma regra local, na qual um equilíbrio de exposição mais alto pode resultar em uma pontuação de classificação de exposições mais alta. Ao contrário, um equilíbrio de exposição mais baixo do quadro de vídeo pode resultar em uma pontuação de classificação de exposições mais baixa.

[0028] Na análise de qualidade de saturação, o módulo de análise de baixo nível 212 pode analisar o histograma de saturação de um quadro de vídeo, tal como um histograma de saturação para um espaço de cor HSV. Com base na análise, o módulo de análise de baixo nível 212 pode computar uma pontuação de saturação que reflete uma quantidade de saturação em um conjunto de valores médio no histograma de saturação de acordo com uma regra local. Como tal,

mais saturação nesta faixa média resulta em uma pontuação de saturação mais alta para o quadro de vídeo. Ao contrário, menos saturação nesta faixa média resulta em uma menor pontuação de saturação para o quadro de vídeo.

[0029] Na análise de variedade de tonalidade, o módulo de análise de baixo nível 212 pode certificar o equilíbrio de um histograma de tonalidade para um quadro de vídeo. O módulo de análise de baixo nível 212 pode ainda atribuir pontuações de tonalidade com base em uma regra local. Consequentemente, o módulo de análise de baixo nível 212 pode atribuir uma pontuação de tonalidade mais alta quando as tonalidades de um quadro de vídeo estão bem balanceadas, isto é, mostra mais variedade de cores. Ao contrário, o módulo de análise de baixo nível pode atribuir uma pontuação de tonalidade mais baixa quando as tonalidades do quadro de vídeo estão menos balanceadas, isto é, mostra menos variedade de cores.

[0030] Na análise de instabilidade, o módulo de análise de baixo nível 212 pode utilizar o módulo de análise de movimento 216 para rastrear o movimento de pontos de característica entre quadros e gerar uma transformada que modela este movimento. Os pontos de característica são ponto de interesse em um quadro de vídeo que podem ser confiavelmente localizados através de múltiplos quadros de vídeo. Um ponto de características é distinto em que este contém uma estrutura bidimensional (2D), e pode estar localizado nas direções x e y. O módulo de análise de baixo nível 212 pode utilizar a transformada para analisar tendências locais e globais relativas à magnitude e direção do movimento de ponto de característica. As tendências locais e globais podem quantificar a instabilidade em um arquivo de vídeo como um atributo de mudança por quadro local. A instabilidade de um arquivo de vídeo pode ser determinada pelo módulo de análise de movimento 216 como abaixo descrito com

relação à análise de categorização de movimento. Consequentemente, o módulo de análise de baixo nível 212 pode aplicar uma regra global que atribui uma pontuação de instabilidade para o arquivo de vídeo que é inversamente proporcional à quantidade de instabilidade no arquivo de vídeo, de modo que uma maior instabilidade resulte em uma menor pontuação de instabilidade, e vice versa.

[0031] Na análise de luminosidade média, o módulo de análise de baixo nível 212 pode calcular uma média dos componentes de luminância de todos os pixels em um quadro de vídeo. Em várias modalidades, o módulo de análise de baixo nível 212 pode calcular a média dos valores de pixel em uma representação de imagem em escala de cinza do quadro de vídeo. Por exemplo, os valores de pixel podem variar de 0-255, no qual 0 corresponde à cor preta e 255 corresponde á cor branca. Em algumas modalidades, o módulo de análise de baixo nível 212 pode ser ainda otimizado para ler o valor de pixel de cada enésimo pixel. Em outras palavras, o módulo de análise de baixo nível 212 pode pular um predeterminado número de pixels na direção x e/ou na direção y enquanto executando a análise. Com base na média dos valores de pixel dos pixels em um quadro de vídeo, o módulo de análise de baixo nível 212 pode determinar a luminosidade do quadro de vídeo. Consequentemente, o módulo de análise de baixo nível 212 pode aplicar uma regra local para atribuir uma pontuação de luminosidade que é proporcional ao valor de pixel médio do quadro de vídeo quando o valor de pixel médio do quadro de vídeo cai dentro de uma faixa média de luminosidade predeterminada. No entanto, o módulo de análise de baixo nível 212 pode atribuir uma pontuação de luminosidade que é mais baixa do que qualquer pontuação de luminosidade que é atribuído a um valor de pixel médio que cai dentro da faixa média de luminosidade predeterminada quando o valor de pixel médio cai fora, isto é, é mais alto ou mais baixo do que a faixa

média de luminosidade predeterminada. Tal pontuação de luminosidade pode diminuir conforme o valor de pixel médio diminui enquanto sendo mais baixa do que o limite inferior da faixa média de luminosidade predeterminada. Tal pontuação de luminosidade pode também diminuir conforme o valor de pixel médio aumenta enquanto sendo mais alta do que o limite superior da faixa média de luminosidade predeterminada.

[0032] Na análise de entropia de cor, o módulo de análise de baixo nível 212 pode determinar a quantidade de entropia de cor em um quadro de vídeo. A quantidade de entropia de cor é um indicador das diferenças entre as cores no quadro de vídeo. O valor de entropia de cor atribuído pelo módulo de análise de baixo nível 212 pode variar de "0" a "1", dependendo da quantidade real de entropia de cor no quadro de vídeo. Por exemplo, o módulo de análise de baixo nível 212 pode atribuir um valor de entropia de cor de "1" para o quadro de vídeo quando os pixels do quadro de vídeo têm a maior quantidade de diferença. O módulo de análise de baixo nível 212 pode atribuir um valor de entropia de cor de "0" para o quadro de vídeo quando os pixels do quadro de vídeo têm a menor quantidade de diferença. Em várias modalidades, o módulo de análise de baixo nível 212 pode determinar o valor de entropia de cor calculando um histograma de domínio de cor para um espaço de cor (por exemplo, o espaço de cor RGB ou o espaço de cor HSV).

[0033] Em tais modalidades, o módulo de análise de baixo nível 212 pode inicialmente criar um histograma que captura múltiplas dimensões de cor. Por exemplo, no espaço de cor RGB, cada um de R, G, e B pode ter 256 possíveis valores, em cujo caso o histograma pode ter 256 x 256 x 256 baldes. Em pelo menos uma modalidade, os baldes podem ser ainda quantizados para otimizar o tamanho de balde e/ou velocidade de processamento, por exemplo, o tamanho pode ser

25 x 25 x 25, de modo que múltiplos valores de cor cairão dentro do mesmo balde. Assim, em um exemplo, o histograma pode ser expresso no seguinte formato em código: int Histogram[256 * 256 * 256], o que significa que a rede de histograma tem um elemento para todas as possíveis cores no espaço de cor RGB. Consequentemente, quando o módulo de análise de baixo nível 212 lê um pixel, o módulo de análise de baixo nível 212 pode ajustar um valor como segue:

```
int IndexInHistogramForColor = pixelColor.red + (256
*pixelColor.green) + (256*256*pixelColor.blue);
Histogram[IndexInHistogramForColor] =
Histogram[IndexInHistogramForColor]+1; // quando mais um pixel com
esta cor é observado, incrementar sua contagem
```

[0034] Uma vez que as etapas acima são executadas para cada pixel no quadro de vídeo, o módulo de análise de baixo nível 212 pode normalizar o histograma. Em outras palavras, o módulo de análise de baixo nível 212 pode dividir cada valor com um tamanho do histograma, de modo que os valores no histograma estejam entre 0 e 1 e os valores somam 1. Como um resultado, um elemento com o valor mais alto ocorre mais frequentemente no quadro de vídeo.

[0035] A entropia para os valores de pixel no histograma pode ser formulada como a soma de todo ($\text{Histogram}[n] * \log(\text{Histogram}[n])$), como segue:

$$H(X) = \sum_i P(x_i)I(x_i) = -\sum_i P(x_i)\log_b P(x_i)$$

[0036] O módulo de análise de baixo nível 212 pode determinar uma entropia de cor relativa (isto é, a entropia de cor do quadro de vídeo com relação a outros quadros de vídeo) dividindo a entropia do histograma pela entropia máxima possível. A entropia máxima possível pode ser definida quando todo $\text{Histogram}[n]$ tem o mesmo valor, isto é, igual a $(1,0 / \text{número de possíveis valores})$. Uma vez que o valor de entropia de cor relativa para um quadro de vídeo é obtido, o módulo de

análise de baixo nível 212 pode aplicar uma regra local para obter uma pontuação de entropia para o quadro de vídeo da entropia de cor relativa do quadro de vídeo. Em várias modalidades, a pontuação de entropia do quadro de vídeo pode ser diretamente proporcional ao valor de entropia relativa do quadro de vídeo, por exemplo, um valor de entropia relativa mais alto resulta em uma pontuação de entropia mais alta, e vice versa.

[0037] Na análise de diferença de histograma, o módulo de análise de baixo nível 212 pode determinar a diferença de histograma entre dois quadros de vídeo adjacentes. Em várias modalidades, o módulo de análise de baixo nível 212 pode dividir cada quadro de vídeo em múltiplas células, (por exemplo, 10 x 10 células). Para cada célula do quadro de vídeo t e do quadro de vídeo adjacente t + 1, o módulo de análise de baixo nível 212 pode calcular um histograma de color (por exemplo, um histograma RGB). Subsequentemente, para cada célula no quadro de vídeo t, o módulo de análise de baixo nível 212 pode computar uma diferença entre o seu histograma e o histograma de uma célula que tem uma posição correspondente no quadro de vídeo adjacente t + 1. As diferenças entre os histogramas de pares de células nos dois quadros de vídeo adjacentes podem ser adicionalmente padronizadas (por exemplo, calcular o quadrado, normalizado, e/ou calcular a média, etc.) para obter um valor de diferença de histograma final para os dois quadros adjacentes, no qual o valor pode variar entre "0" e "1". Uma vez que o valor de diferença de histograma para os dois quadros de vídeo adjacentes é obtida, o módulo de análise de baixo nível 212 pode aplicar uma regra local para obter uma pontuação de diferença de histograma para o quadro de vídeo t do valor de diferença de histograma. Em várias modalidades a pontuação de diferença de histograma do quadro de vídeo pode ser diretamente proporcional à diferença de histograma, por exemplo, um

valor de diferença de histograma mais alto resulta em uma pontuação de diferença de histograma mais alta, e vice versa.

[0038] Em pelo menos algumas modalidades, o módulo de análise de baixo nível 212 pode otimizar algumas das análises para acelerar o processamento de um arquivo de vídeo. Por exemplo, o módulo de análise de baixo nível 212 pode executar as análises (por exemplo, análise de classificação de exposição, análise de variedade de tonalidade, etc.) para um subconjunto do quadro de vídeo em um arquivo de vídeo ao invés de todos os quadros de vídeo no arquivo de vídeo. O módulo de análise de baixo nível 212 pode também executar uma ou mais das análises em uma versão de escala diminuída de um quadro original para ganhar alguma eficiência. Por exemplo, a análise de variedade de tonalidade e a análise de saturação para um quadro de vídeo podem ser executadas em uma versão de resolução diminuída do quadro de vídeo. Em outro exemplo, a análise de qualidade de exposição pode ser executada em um monocromo e versão de resolução diminuída do quadro de vídeo. O módulo de análise de baixo nível 212 pode também executar múltiplas análises em paralelo ou substancialmente em paralelo. Por exemplo, a análise de qualidade de saturação e a análise de variedade de tonalidade podem ser executadas em paralelo.

[0039] O módulo de análise de alto nível 214 pode analisar cada quadro de vídeo decodificado para características de alto nível. Em pelo menos uma modalidade, as análises de característica de alto nível pode incluir detecção de face, rastreamento de face, reconhecimento de face, análise de saliência, análise de potência de áudio, análise de classificação de áudio, análise de fala, e análise de movimento.

[0040] Na detecção de face, o módulo de análise de alto nível 214 pode analisar um quadro de vídeo decodificado para detectar se faces

humanas estão apresentadas no quadro de vídeo. Uma face detectada pode estar faceando uma câmera que capturou o quadro de vídeo ou lateralmente com relação à câmera. Com base nesta detecção, o módulo de análise de alto nível 214 pode gerar uma lista de faces detectadas com as suas posições no quadro de vídeo, a área do quadro de vídeo coberta por cada face, e uma pontuação de confiança de detecção para cada face que indica uma confiança na detecção.

[0041] Em várias modalidades, o módulo de análise de alto nível 214 pode aplicar uma regra local para calcular uma pontuação de importância de face para o quadro de vídeo com base em um tamanho de uma face detectada como uma percentagem de um tamanho do quadro de vídeo. As faces com o mesmo tamanho como detectado em dois quadros de vídeo podem ser atribuídas a mesma pontuação de importância de face. No entanto, se uma face em um quadro de vídeo t2 for maior do que uma face em um quadro de vídeo t1, então a pontuação de importância de face para o quadro de vídeo t2 será mais alta, porque faces maiores são consideradas mais importantes do que faces menores. Em outras modalidades, o módulo de análise de alto nível 214 pode estar configurado para calcular uma pontuação de importância de face se o tamanho da face detectada for entre um mínimo limite de tamanho e um máximo limite de tamanho. Ao contrário, faces cujos tamanhos são menores do que o mínimo limite de tamanho ou maiores do que o máximo limite de tamanho podem ser consideradas inválidas para o cálculo de pontuação de importância de face pelo módulo de análise de alto nível 214, ou uma tendência de pontuação negativa pode ser atribuída ao quadro de vídeo correspondente para tais ocorrências.

[0042] Alternativamente ou concorrentemente, a regra local para calcular a importância de face para o quadro de vídeo pode levar em consideração as características faciais de cada face apresentada no

quadro de vídeo. Por exemplo, características faciais podem incluir se a face está sorrindo ou não, ou se os olhos estão abertos ou não. etc. Assim, uma tendência de pontuação positiva pode ser atribuída ao quadro de vídeo correspondente para uma face que está sorrindo, enquanto uma tendência de pontuação negativa pode ser atribuída quando a face não está sorrindo. Do mesmo modo, uma tendência de pontuação pode ser atribuída para o quadro de vídeo correspondente para uma face com olhos abertos, enquanto uma tendência de pontuação negativa pode ser atribuída quando os olhos estão fechados.

[0043] Em reconhecimento de face, o módulo de análise de alto nível 214 pode utilizar um algoritmo de reconhecimento facial para corresponder cada face humana que é detectada em um quadro de vídeo a uma identidade conhecida de uma pessoa. Em algumas modalidades, o módulo de análise de alto nível 214 pode utilizar um banco de dados de conhecimento de faces conhecidas para corresponder uma face humana com uma pessoa conhecida. Alternativamente ou concorrentemente, o módulo de análise de alto nível 214 pode utilizar o módulo de interface de usuário 224 para prover controles de interface de usuário que permitem um usuário identificar cada uma das uma ou mais faces reconhecidas com uma identidade, uma classificação de importância da face, e/ou uma relação da pessoa com a face para o espectador. Em pelo menos uma modalidade, as informações providas pelo espectador com relação às faces podem ser adicionadas ao banco de dados de conhecimento.

[0044] Em rastreamento de face, o módulo de análise de alto nível 214 pode rastrear as faces humanas através de múltiplos quadros de vídeo. Deste modo, o módulo de análise de alto nível 214 pode verificar um conjunto de faces que estão presentes em um arquivo de vídeo, assim como rastrear a frequência que cada face aparece no

arquivo de vídeo. Mais ainda, em agrupamento de faces, o módulo de análise de alto nível 214 pode agrupar faces que são rastreadas para determinar se as faces que são detectadas em diferentes quadros de vídeo pertencem à mesma pessoa. Em várias modalidades, o módulo de análise de alto nível 214 pode obter um conjunto de características faciais para cada uma das faces detectadas nos quadros de vídeo. O módulo de análise de alto nível 214 pode comparar os conjuntos de características faciais para as faces detectadas de modo a agrupar as faces detectadas em grupos de acordo com similaridade de características facial. O módulo de análise de alto nível 214 pode certificar a importância de cada grupo de faces de acordo com os números de faces em cada grupo. O número de faces em cada grupo é diretamente proporcional à prevalência da face no arquivo de vídeo. Ainda, uma prevalência mais alta pode indicar uma importância mais alta da pessoa com a face, e vice versa. Consequentemente, uma potência de grupo de importância para um grupo de faces pode ser um somatório das pontuações de importância de face das faces no grupo. Como tal, o módulo de análise de alto nível 214 pode somar as pontuações de importância de face das faces em um grupo, como apresentado em múltiplos quadros de vídeo, para calcular uma pontuação de importância de grupo para o grupo.

[0045] Por exemplo, a pessoa cuja face aparece no grupo com a mais alta pontuação de importância de grupo pode ser considerada um caráter principal. Com base em tais informações, o módulo de análise de alto nível 214 pode aplicar uma regra global nos quadros de vídeo, de modo que as pontuações de importância de face dos quadros de vídeo que mostram o caráter principal possam ser adicionalmente positivamente tendenciados, isto é, elevados. As posições de faces que estão incluídas em um grupo como capturadas em múltiplos quadros de vídeo podem também diminuir as pontuações de

importância de face atribuídas a quadros de vídeo. Por exemplo, a pontuação de importância de um quadro de vídeo que mostra uma face de uma pessoa específica pode ter uma tendência de acordo com a distância da face para o centro do quadro de vídeo ao longo de um eixo geométrico (por exemplo, eixo geométrico x ou eixo geométrico y), de modo que a face que está mais próxima do centro resulta em uma importância mais alta para o quadro de vídeo, e vice versa.

[0046] Em análise de saliência de quadro, o módulo de análise de alto nível 214 pode detectar as partes salientes de um quadro de vídeo. Por exemplo, uma parte saliente do quadro de vídeo pode capturar um objeto que está em movimento. Com base na análise de saliência, o módulo de análise de alto nível 214 pode aplicar uma regra local para gerar um mapa de calor que exibe a pontuação de saliência de cada pixel em um quadro de vídeo. Um mapa de calor é uma representação gráfica de dados que estão dispostos em uma matriz na qual valores individuais na matriz são representados utilizando cores. O módulo de análise de alto nível 214 pode ainda gerar uma pontuação de saliência de quadro para o quadro de vídeo que está baseada nas pontuações de saliência dos pixels no quadro de vídeo. Por exemplo, a pontuação de saliência de quadro de vídeo para o quadro de vídeo pode ser uma média das pontuações de saliência de pixel.

[0047] Em análise de potência de áudio, o módulo de análise de alto nível 214 pode acessar os dados de áudio que correspondem em duração de tempo a um quadro de vídeo (por exemplo, 1/30 avos ou 1/60 avos de um segundo) e calcular um valor de raiz média quadrática (RMS) da potência de áudio. O valor de RMS mais alto da potência de áudio pode indicar uma importância mais alta do quadro de vídeo correspondente, e vice versa. Assim, o módulo de análise de alto nível 214 pode atribuir uma pontuação de importância de potência

de áudio ao quadro de vídeo correspondente de acordo com a regra local.

[0048] Em análise de classificação de áudio, o módulo de análise de alto nível 214 pode utilizar um classificador de aprendizado de máquina para determinar se os dados de áudio que correspondem em duração de tempo a um quadro de vídeo contêm diferentes tipos de dados de áudio (por exemplo, ruído, fala, ou música). Os diferentes tipos de dados de áudio podem refletir uma diferente importância do quadro de vídeo correspondente. Com base em uma regra local, o módulo de análise de alto nível 214 pode atribuir uma pontuação de importância de classificação de áudio para o quadro de vídeo correspondente com base no tipo de dados de áudio. Por exemplo, a presença de fala pode fazer com que o módulo de análise de alto nível 214 atribua uma alta importância de classificação de áudio a um quadro de vídeo correspondente. Em contraste, a presença de música pode causar a atribuição de uma pontuação de classificação de áudio média para quadro de vídeo correspondente. Mais ainda, a presença de ruído pode fazer com que o módulo de análise de alto nível 214 atribua uma baixa pontuação de classificação de áudio para o quadro de vídeo correspondente.

[0049] Em análise de categorização de movimento, o módulo de análise de alto nível 214 pode utilizar o módulo de análise de movimento 216 para rastrear o movimento de pontos de característica entre quadros de vídeo e gerar uma transformada que modela este movimento. O módulo de análise de alto nível 214 pode utilizar a transformada para analisar as tenências locais e globais relativas à magnitude e direção de movimento. Por sua vez, o módulo de análise de alto nível 214 pode utilizar as tenências locais e globais para levar em conta a instabilidade capturada nos quadros de vídeo e determinar um movimento intencional da câmera com relação a uma cena, tal

com zoom, panoramização, etc.

[0050] Em várias modalidades, o módulo de análise de movimento 216 pode iniciar a análise de categorização de movimento localizando pontos de característica para dois quadros adjacentes. Um ponto de característica pode ser um ponto em uma imagem que permanece identificável mesmo com transformadas bidimensionais (2D) ou 3D na imagem. Para detectar os pontos de característica, o módulo de análise de movimento 216 pode diminuir a resolução da imagem e criar uma pirâmide de imagens de resolução diminuída de menores dimensões. As imagens de resolução diminuída são então comparadas pelo módulo de análise de movimento 216 para determinar pontos comuns (isto é, pontos de característica) entre as imagens de resolução diminuída. Em várias modalidades, o módulo de análise de movimento 216 pode utilizar um ou mais de diversos algoritmos de detecção para detectar os pontos comuns, tal como um algoritmo de detecção de Laplace, um algoritmo de detecção de Harris, um algoritmo de detecção de Hessian, um algoritmo de detecção de HessianLaplace, um algoritmo de detecção de HessianAffine, um algoritmo de detecção de EdgeFoci, etc.

[0051] Uma vez que os pontos de característica são identificados para dois quadros adjacentes, o módulo de análise de movimento 216 pode determinar uma transformada que alinhe os dois quadros adjacentes de modo que um número máximo de pontos de característica coincidam. A transformada pode ser executada utilizando correspondência geométrica que é uma implementação de estimativa de parâmetro robusto. AS transformada pode prover uma matriz de transformada de homografia que é calculada dos pontos de característica coincididos. Em várias modalidades, o módulo de análise de movimento 216 pode utilizar um algoritmo de Amostragem e Consenso Randômico (RANSAC) para obter estimativas de parâmetro

iniciais e uma lista de assomos estatísticos, nos quais as estimativas de parâmetro inicial são adicionalmente referidas. Os vários movimentos de câmera em quadros de vídeo que estão alinhados pela transformada homografia estão ilustrados na Figura 3.

[0052] A Figura 3 é um diagrama ilustrativo que mostra a utilização de uma transformada de homógrafo para alinhar pontos de característica exemplares em múltiplos quadros de vídeo. Como mostrado, cada um dos quadros de vídeo 302-312 pode respectivamente incluir um grupo de pontos de característica identificados que estão apresentados por círculos. Por exemplo, o grupo de pontos de característica no quadro de vídeo 302 está apresentado pelos círculos 314(1)-315(5). Cada ponto de características em um grupo de pontos de característica pode reter as suas posições relativas uns aos outros através de múltiplos quadros de vídeo, apesar do movimento da câmera que capturou os múltiplos quadros de vídeo. Por exemplo, o quadro de vídeo transformado 316 pode ser um quadro de vídeo subsequente ao quadro de vídeo 302 que é um resultado de um movimento de rastreamento de uma câmera. O quadro de vídeo transformado 318 pode ser um quadro de vídeo subsequente ao quadro de vídeo 304 que é um resultado de um movimento de lança da câmera. O quadro de vídeo transformado 320 pode ser um quadro de vídeo subsequente ao quadro de vídeo 306 que é um resultado de um movimento de zoom / carrinho da câmera. O quadro de vídeo transformado 312 pode ser um quadro de vídeo subsequente ao quadro de vídeo 308 que é um resultado de um movimento de rolagem da câmera. O quadro de vídeo transformado 314 pode ser um quadro de vídeo subsequente ao quadro de vídeo 310 que é um resultado de uma panoramização / oscilação / inclinação vertical da câmera. O quadro de vídeo transformado 316 pode ser um quadro de vídeo subsequente ao quadro de vídeo 312 que é um

resultado de uma panoramização horizontal da câmera.

[0053] No entanto, independentemente do movimento da câmera que produziu os quadros de vídeo transformados 326-336 dos quadros de vídeo 302-312, o módulo de análise de movimento 216 pode utilizar a transformada de homógrafo para alinhar o ponto de característica em um quadro de vídeo e seu quadro de vídeo transformado correspondente.

[0054] Retornando à Figura 2, o algoritmo RANSAC diretamente computa os parâmetros de matriz de transformação de um mínimo subconjunto das coincidências de ponto de característica. Por exemplo, uma transformada de similaridade (por exemplo, translação, rotação ou escala) pode ser computada de dois pontos de característica que estão em correspondência entre dois quadros. Uma vez que uma transformação geométrica candidata foi obtida, o algoritmo RANSAC pode validar a transformação testando a transformação em todas as outras coincidências de ponto de característica no conjunto de dados, e gerando uma contagem do número de incluídos os quais são pontos de característica que projetam espacialmente com suficiente precisão. Em outras palavras, o algoritmo RANSAC pode inicialmente randomicamente escolher um conjunto mínimo de coincidências de ponto, computar os parâmetros de transformação deste conjunto, e então validar estes parâmetros contando o número de coincidências de incluídos. Subsequentemente, o algoritmo RANSAC registra a melhor transformação. O algoritmo RANSAC pode repetir este procedimento um número de vezes até que a probabilidade de encontrar um bom conjunto de parâmetros de transformação atinge um limite de probabilidade predeterminado dada a taxa de divergência de dados.

[0055] Em algumas modalidades, o algoritmo RANSAC pode ser modificado para fazer melhor uso da estatística Bayesiana. Ao invés

de contar os incluídos durante a validação dos parâmetros de transformação, o algoritmo RANSAC pode computar uma pontuação de probabilidade de log para cada transformação randômica de todas as coincidências de ponto de característica. Esta pontuação pode incluir duas partes: (1) uma pontuação de probabilidade anterior a qual depende dos parâmetros e quão distantes os parâmetros estão de valores comumente esperados, e (2) uma pontuação de probabilidade com base em uma função robusta da distância de reprojeção das coincidências de ponto de característica. Tal pontuação favorece os pontos de característica os quais projetam para as localizações corretas, mas permite que os discrepantes coexistam.

[0056] Da transformada de homógrafo, o módulo de análise de movimento 216 pode extrair as magnitudes e direção dos componentes de zoom e translação vertical, enquanto ignorando outros tipos de movimento. Estas magnitudes e direções são um movimento intencional de uma câmera como gravado nos dois quadros adjacentes. Em outras palavras, estas magnitudes e direções são derivadas de primeira ordem da mudança de movimento real de um primeiro quadro para um segundo quadro adjacente. O módulo de análise de movimento 216 pode determinar o movimento de instabilidade da câmera que gravou os quadros de vídeo calculando a média de execução de dados de movimento dos vetores de movimento dos quadros de vídeo, e subtraindo o movimento intencional da câmera da média de execução de dados de movimento. O cálculo da média de execução de dados de movimento suprime a variância local e preserva as tendências de longo prazo que representam o movimento intencional. Em outras palavras, a diferença entre o movimento intencional e a mudança de movimento do primeiro quadro para o segundo quadro é o movimento de instabilidade da câmera que gravou os dois quadros.

[0057] A magnitude de valores de zoom e translação vertical que são gravados em um conjunto de quadros pode prover um indício referente à importância destes quadros. Por exemplo, um valor mais alto pode indicar uma aceleração no movimento de câmera com relação a um ou mais objetos em um quadro de vídeo. Ainda, uma região de um quadro de vídeo com aceleração mais alta pode ser assumida ser mais importante, porque a câmera pode ter feito uma rápida mudança em movimento para capturar alguma ação. Consequentemente, o módulo de análise de alto nível 214 pode atribuir uma pontuação de importância de movimento a cada quadro com base na quantidade de aceleração de movimento.

[0058] O módulo de análise de movimento 216 pode analisar os dados de movimento intencional para determinar os máximos e mínimos locais articulados ao redor de cruzamentos zero para os movimentos tanto de zoom quanto de translação vertical. Em algumas modalidades, o módulo de análise de movimento 216 pode utilizar as localizações de máximos e mínimos locais para segmentar os dados em seções de vídeo. Alternativamente ou concorrentemente, o módulo de análise de movimento 216 pode utilizar as localizações de máximos e mínimos locais para segmentar o arquivo de vídeo em seções de vídeo que têm direções de mudança em par, tal como aumento de zoom em par com diminuição de zoom, panoramização para cima em par com panoramização para baixo , etc.

[0059] As seções de vídeo segmentadas de um arquivo de vídeo podem ser consistentes com mudanças em cenas do arquivo de vídeo. Consequentemente, os limites de seção no arquivo de vídeo podem ser utilizados como orientações para dividir o arquivo de vídeo em seções de vídeo de diferente importância. Os limites podem alinhar o início e o final de seções importantes / não importantes com os pontos no tempo no qual existe um deslocamento no movimento da câmera

ou uma mudança na natureza da atividade na cena. Mais ainda, o módulo de análise de movimento 216 pode combinar e calcular a média de magnitudes para zoom e movimento de panoramização para uma seção. A quantidade de aceleração representada pela média das magnitudes de zoom e movimento de panoramização para uma seção de vídeo pode ser utilizada pelo módulo de análise de movimento 216 para atribuir uma pontuação de importância de movimento para a seção de vídeo no mesmo modo como acima descrito com relação a quadros.

[0060] Em pelo menos algumas modalidades, o módulo de análise de alto nível 214 pode otimizar algumas das análises para acelerar o processamento de um arquivo de vídeo. Por exemplo, o módulo de análise de alto nível 214 pode executar a detecção de face, o rastreamento de face, e/ou o reconhecimento de face para uma ou mais faces em cada quadro de vídeo utilizando um monocromo e uma versão de resolução diminuída do quadro de vídeo. O módulo de análise de alto nível 214 pode também executar múltiplas análises em paralelo ou substancialmente em paralelo. Por exemplo, o rastreamento de face e o reconhecimento de face podem ser executados em paralelo.

[0061] O módulo de cálculo de importância 218 pode normalizar as várias pontuações de característica que são geradas para os quadros de vídeo de um arquivo de vídeo e calcular um valor de importância de vídeo para o arquivo de vídeo. Por exemplo, O módulo de cálculo de importância 218 pode calcular a média de um conjunto de pontuações de característica normalizadas (por exemplo, pontuação de importância de face, pontuação de importância de movimento, pontuação de classificação de exposição, pontuação de saturação, etc.) para cada quadro de vídeo para obter uma pontuação de importância de quadro para cada quadro de vídeo. As pontuações de

importância de quadro de vídeo podem ter ainda a média calculada para derivar o valor de importância de vídeo para o arquivo de vídeo. Em algumas modalidades, o cálculo do valor de importância de vídeo para o arquivo de vídeo pode também incluir a tendência de uma ou mais pontuações de característica que estão associadas com os quadros de vídeo. Por exemplo, o módulo de cálculo de importância 218 pode estar configurado para aplicar uma tendência positiva de modo que a presença de uma face em um quadro de vídeo afete uma pontuação de importância de quadro daquele quadro por um grau mais alto do que a pontuação de variedade de tonalidade do quadro de vídeo.

[0062] Em outro exemplo, o módulo de cálculo de importância 218 pode gerar um valor de importância de vídeo para um arquivo de vídeo como segue:

$$\text{frame_score} = w_1 * \text{Face Importance} + w_2 * F_2 + w_3 * F_3 + \dots w_n * F_n \quad (1)$$

$$\text{video_score} = \frac{\sum(\text{frame_score})}{\text{number of frames}} \quad (2)$$

[0063] na qual w_i são pesos e F_i são características. Os pesos podem ditar a importância de características. Por exemplo, se um espectador preferir um vídeo que é luminoso, e F_2 é a característica que correlaciona a esta propriedade, então o módulo de cálculo de importância 218 pode estar configurado para atribuir um valor mais alto para w_2 do que os pesos para outras características. Esta tendência pode ser aplicada em outros cenários. Em um exemplo adicional, se um espectador deseja de preferência selecionar arquivos de vídeo que mostram uma pessoa específica, o espectador pode configurar o módulo de cálculo de importância 218 para induzir quadros que mostram a face da pessoa específica para intensificar a pontuação de importância de quadro de vídeo de tais quadros de vídeo. Em várias modalidades, o módulo de cálculo de importância 218 pode armazenar os valores de característica $\{F1, F2 \dots Fn\}$ para

um arquivo de vídeo no armazenamento de dados 226. O armazenamento dos valores de características para um arquivo de vídeo pode eliminar uma futura análise duplicada do arquivo de vídeo em cenários nos quais diferentes características devem ser induzidas.

[0064] Em pelo menos uma outra modalidade, o módulo de cálculo de importância 218 pode estar configurado para aplicar uma tendência negativa a uma característica mostrada em um quadro de vídeo. Por exemplo, uma tendência negativa que é proporcional à quantidade de instabilidade pode ser implementada pelo módulo de cálculo de importância 218 para diminuir a pontuação de importância de quadro de vídeo do quadro de vídeo proporcionalmente à instabilidade.

[0065] O módulo de segmentação de vídeo 220 pode segmentar um arquivo de vídeo em múltiplas seções de vídeo com base em importância. Em algumas modalidades, o módulo de segmentação de vídeo 220 pode encontrar uma seção de vídeo com uma duração t que é mais curta do que a duração do arquivo de vídeo. Em tais modalidades, o módulo de segmentação de vídeo 220 pode calcular uma massa de janela que a soma das pontuações de importância de quadro em quadros de vídeo em uma janela que tem ($t * \text{taxa de quadro}$) quadros de vídeo de um arquivo de vídeo. Tal massa de janela pode ser calculada sucessivamente em um modo de deslocamento para todos os quadros de vídeo do arquivo de vídeo. Consequentemente, o módulo de segmentação de vídeo 220 pode selecionar um quadro de vídeo com uma mais alta massa de janela como o centro da t seção de vídeo importante longa. Em outras modalidades, o módulo de segmentação de vídeo 220 pode basear-se no módulo de análise de movimento 216 para segmentar um arquivo de vídeo em seções de vídeo com base em dados de movimento. Uma vez que o módulo de segmentação de vídeo 220 segmentou um arquivo de vídeo em seções de vídeo, o módulo de cálculo de

importância 218 pode gerar um valor de importação de seção para cada seção de vídeo em um modo similar como com relação aos arquivos de vídeo inteiros. Em outras palavras, o módulo de cálculo de importância 218 pode gerar um valor de importação de seção com base nas pontuações de importância de quadro normalizada dos quadros de vídeo na seção de vídeo. Em alguns casos, o módulo de cálculo de importância 218 pode também aplicar tendência a uma ou mais pontuações de característica durante a geração de valor de importância de seção de seções de vídeo.

[0066] Em várias modalidades, cada um do módulo de análise de baixo nível 212, o módulo de análise de alto nível 214, e o módulo de cálculo de importância 218 pode armazenar as pontuações, valores, e outras informações que são obtidas para as seções de vídeo e/ou arquivos de vídeo como metadados associados no armazenamento de dados 226. Tais metadados pode ser combinados com outros metadados que estão associados com os arquivos de vídeo, tal como data, localização, número de compartilhamentos online, etc.

[0067] O módulo de classificação de vídeo 222 pode classificar as seções de vídeo de um arquivo de vídeo com base em seus valores de importância de seção. Alternativamente ou concorrentemente, o módulo de classificação de vídeo 222 pode classificar os arquivos de vídeo de acordo com seus valores de importância de vídeo. A classificação pode ser do mais importante para o menos importante, ou vice versa. Para cada arquivo de vídeo classificado, o módulo de classificação de vídeo 222 pode também armazenar metadados referentes às seções de vídeo no arquivo de vídeo classificado. Tais metadados podem incluir a classificação de cada seção de vídeo, o tempo de início e fim de cada seção de vídeo, a duração de cada seção de vídeo, e o valor de importância de seção de cada seção de vídeo. Em algumas modalidades, o módulo de classificação de vídeo

222 pode também calcular valores adicionais para um arquivo de vídeo ou uma seção de vídeo. Estes valores podem incluir uma densidade de importância, a qual pode refletir uma percentagem dos quadros de vídeo em um arquivo de vídeo ou uma seção de vídeo com uma pontuação de importância que excede um limite de pontuação de importância. Estes valores podem também incluir uma densidade de qualidade, a qual pode refletir uma percentagem de quadros em um arquivo de vídeo ou uma seção de vídeo com características negativas ou positivas que excedem um limite correspondente. Tais características negativas ou positivas podem incluir instabilidade, superexposição, subexposição, etc. O módulo de classificação de vídeo 222 pode armazenar os múltiplos tipos de pontuações e outros valores computados que são utilizados para gerar classificações for arquivos de vídeo e/ou seções de vídeo individuais como metadados associados no armazenamento de dados 226.

[0068] Uma vez que um conjunto de arquivos de vídeo ou seções de vídeo são classificados, o módulo de classificação de vídeo 222 pode exibir representações de imagem em miniatura dos arquivos de vídeo ou seções de vídeo classificados. Consequentemente, a seleção de uma representação de imagem em miniatura pode fazer cm que um reproduutor de mídia 122 execute a seção de vídeo correspondente, ou outras aplicações para prover a seção de vídeo correspondente para compartilhar e/ou editar.

[0069] O módulo de interface de usuário 224 pode permitir um usuário interagir com os módulos da máquina de análise de vídeo 102 através das interfaces 204. Por exemplo, o módulo de interface de usuário 224 pode permitir o usuário selecionar arquivos de vídeo para análise de importância, sinalizar faces humanas que estão identificadas em arquivos de vídeo com informações, destacar faces ou pessoas para tendência de pontuação característica positiva, e/ou

selecionar arquivos de vídeo e seções de vídeo para reprodução pelo reprodutor de mídia 122 através de representações de imagem em miniatura. Em algumas modalidades, o usuário pode também utilizar o módulo de interface de usuário 224 para selecionar uma ou mais das características de baixo nível ou uma ou mais das características de alto nível de um arquivo de vídeo para análise pela máquina de análise de vídeo 102.

[0070] O armazenamento de dados 226 pode armazenar os dados que estão utilizados pelos vários módulos. Em pelo menos algumas modalidades, o armazenamento de dados 226 pode armazenar arquivos de vídeo 228, arquivos de vídeo classificados 230, seções de vídeo classificadas 232, e/ou metadados 234 associados com os arquivos de vídeo classificados 230 e as seções de vídeo classificadas 232. Em outras modalidades, o armazenamento de dados 226 pode armazenar dados (por exemplo, pontuações de importância) associados com os arquivos de vídeo ou seções de vídeo que são utilizados para classificar os arquivos de vídeo e as seções de vídeo. O armazenamento de dados 226 pode ainda armazenar produtos e valores adicionais que são gerados pelos módulos, tal como matrizes de transformada de homógrafo, pontuações de característica, valores de importância de vídeo, valores de importância de seção, etc.

[0071] Em algumas modalidades, um ou mais aplicações adicionais podem ser instaladas nos dispositivos de computação 104. Tais aplicações podem incluir uma aplicação de edição de vídeo que é utilizada para compilar um novo arquivo de vídeo de seções de vídeo seletivas de um arquivo de vídeo original. Por exemplo, tal aplicação de edição de vídeo pode permitir um usuário selecionar seções de vídeo com valores de importância de seção que excedem um limite de pontuação específico para serem digitalmente combinadas juntas para criar um arquivo de vídeo de destaque. As aplicações podem também

incluir uma aplicação de compartilhamento online que permite um usuário postar um arquivo de vídeo, seção de vídeo, ou um vídeo de destaque online. Em modalidades adicionais, uma ou mais outras aplicações podem ser instaladas nos dispositivos de computação 104 para acessar os dados armazenados no armazenamento de dados 226 para os arquivos de vídeo e as seções de vídeo através de uma interface de aplicação. Tal aplicação pode acessar os dados de modo a utilizar os resultados de análise em outros modos. Em outras palavras, a máquina de análise de vídeo 102 pode funcionar como um serviço de nível mais baixo para prover dados para estas aplicações.

PROCESSOS EXEMPLARES

[0072] As Figuras 4-7 descrevem vários processos exemplares para executar uma análise de importância baseada em regras de arquivos de vídeo. A ordem na qual as operações estão descritas em cada processo exemplar não pretende ser considerada como uma limitação, e qualquer número das operações descritas podem ser combinadas em qualquer ordem e/ou em paralelo para implementar cada processo. Mais ainda, as operações em cada uma das Figuras 4-7 podem ser implementadas em hardware, software, e uma sua combinação. No contexto de software, as operações representam instruções executáveis por computador que, quando executadas por um ou mais processadores, fazem com que um ou mais processadores executem as operações recitadas. Geralmente, as instruções executáveis por computador incluem rotinas, programas, objetos, componentes, estruturas de dados, e assim por diante que fazem com que as funções específicas sejam executadas ou tipos de dados específicos sejam implementados.

[0073] A Figura 4 é um fluxograma que ilustra um processo exemplar 400 para utilizar uma análise de vídeo baseada em regras para analisar características em um arquivo de vídeo de modo a

classificar seções de vídeo do arquivo de vídeo com base em sua importância. No bloco 402, a máquina de análise de vídeo 102 pode decodificar um arquivo de vídeo, tal como o arquivo de vídeo 108(1), para obter um quadro de vídeo e dados de áudio associados com o quadro de vídeo. O módulo de decodificador de vídeo 210 pode utilizar vários codecs para decodificar os arquivos de vídeo, tal como H.264, MPEG-2, MPEG-4, etc. Os dados de áudio associados podem ter a mesma duração de tempo que o quadro de vídeo, por exemplo, 1/30th de um segundo ou 1/60th de um segundo. No entanto, em algumas modalidades, o módulo de decodificador de vídeo 210 pode estar configurado para obter um quadro de vídeo sem obter os dados de áudio associados, ou vice versa.

[0074] No bloco 404, a máquina de análise de vídeo 102 pode analisar pelo menos uma das características de vídeo do quadro de vídeo ou características de áudio dos dados de áudio para obter as pontuações de característica. A máquina de análise de vídeo 102 pode executar tal análise com base em uma ou mais regras locais. Em várias modalidades, as uma ou mais características que são analisadas podem incluir características de alto nível e/ou características de baixo nível. Por exemplo, as características de baixo nível podem incluir características tais como qualidade de exposição, qualidade de saturação, instabilidade de quadros de vídeo, luminosidade média, entropia de cor, e/ou diferenças de histograma entre quadros de vídeo. As características de alto nível podem incluir características tais como as quantidades, posições, e/ou características faciais de faces que são detectadas nos quadros de vídeo.

[0075] No bloco 406, a máquina de análise de vídeo 102 pode armazenar as pontuações de característica para o quadro de vídeo como metadados para o quadro de vídeo. Em várias modalidades, a

máquina de análise de vídeo 102 pode armazenar os metadados no armazenamento de dados 226. Tais metadados podem reduzir ou eliminar análises recorrentes de quadros de vídeo durante uma futura determinação da importância de arquivos de vídeo correspondentes ou seções de vídeo que envolvem o mesmo quadros de vídeo.

[0076] No bloco de decisão 408, a máquina de análise de vídeo 102 pode determinar se existem quadros adicionais do arquivo de vídeo para analisar. Em outras palavras, a máquina de análise de vídeo 102 pode determinar se todos os quadros de vídeo e dados de áudio associados do arquivo de vídeo foram decodificados. Se a máquina de análise de vídeo 102 determinar que existem quadros adicionais para analisar ("sim" no bloco de decisão 408), o processo 400 pode retornar para o bloco 402. No bloco 402, a máquina de análise de vídeo 102 pode obter um quadro de vídeo subsequente e dados associados do arquivo de vídeo. No entanto, se a máquina de análise de vídeo 102 determinar no bloco de decisão 408 que nenhum quadro de vídeo adicional do arquivo de vídeo deve ser analisado ("não" no bloco de decisão 408), o processo 400 pode prosseguir para o bloco 410.

[0077] No bloco 410, a máquina de análise de vídeo 102 pode aplicar pelo menos uma regra global para uma ou mais resultados de característica. Por exemplo, a pessoa cuja face aparece em um grupo com a pontuação de importância de grupo mais alta pode ser considerada um personagem principal. Com base em tais informações, o módulo de análise de alto nível 214 pode aplicar uma regra global nos quadros de vídeo, de modo que as pontuações de importância de face dos quadros de vídeo que mostram o personagem principal podem ser adicionadamente avaliadas.

[0078] No bloco 412, a máquina de análise de vídeo 102 pode combinar todas as pontuações de característica para cada quadro de

vídeo do arquivo de vídeo para derivar uma pontuação de importância de quadro correspondente para cada quadro de vídeo. Por exemplo, a máquina de análise de vídeo 102 pode calcular a média de um conjunto de pontuações de característica normalizadas para cada quadro de vídeo para obter uma pontuação de importância de quadro para cada quadro de vídeo.

[0079] No bloco 414, a máquina de análise de vídeo 102 pode armazenar metadados para o arquivo de vídeo. Os metadados podem incluir as pontuações de importância de quadro de vídeo dos quadros de vídeo do arquivo de vídeo e/ou as pontuações de característica para cada quadro de vídeo.

[0080] No bloco 416, a máquina de análise de vídeo 102 pode segmentar o arquivo de vídeo em seções de vídeo com base nas pontuações de importância de quadro de vídeo dos quadros de vídeo. Em algumas modalidades, a máquina de análise de vídeo 102 pode utilizar o cálculo de massa de janela para segmentar o arquivo de vídeo em seção de vídeo. Em outras modalidades, a máquina de análise de vídeo 102 pode utilizar os cruzamentos zero para o zoom e movimentos de translação vertical dos dados de movimento capturados no arquivo de vídeo para segmentar o arquivo de vídeo em seções de vídeo.

[0081] No bloco 418, a máquina de análise de vídeo 102 pode calcular o valor de importância de seção para cada seção de vídeo do arquivo de vídeo. Em várias modalidades, a máquina de análise de vídeo 102 pode gerar o valor de importância de seção com base nas pontuações de importância de quadro normalizadas dos quadros de vídeo na seção de vídeo. Em alguns casos, o módulo de cálculo de importância 218 pode também aplicar uma tendência a uma ou mais pontuações de característica durante a geração de valores de importância de seção de seções de vídeo.

[0082] No bloco 420, a máquina de análise de vídeo 102 pode classificar as seções de vídeo de acordo com os seus valores de importância de seção. Por exemplo, as seções de vídeo podem ser classificadas da maior importância para menor importância, ou vice versa. Uma vez que as seções de vídeo são classificadas, a máquina de análise de vídeo 102 pode exibir representações de imagem em miniatura das seções de vídeo. A seleção de representação de imagem em miniatura pode fazer com que um reproduutor de mídia 122 execute a seção de vídeo correspondente, ou outras aplicações para prover a seção de vídeo correspondente para compartilhar e/ou editar. Alternativamente, a máquina de análise de vídeo 102 pode prover os dados de classificação para outra aplicação, de modo que a aplicação possa exibir as representações em miniatura das seções de vídeo classificadas.

[0083] A Figura 5 é um fluxograma que ilustra um processo exemplar 500 para utilizar uma análise de vídeo baseada em regras para analisar as características de arquivos de vídeo para classificar arquivos de vídeo com base em sua importância. No bloco 502, a máquina de análise de vídeo 102 pode obter um arquivo de vídeo de uma coleção de arquivos de vídeo, tal como a coleção de vídeos 106. Em várias modalidades, a máquina de análise de vídeo 102 pode obter o arquivo de vídeo com base em uma seleção inserida de um usuário.

[0084] No bloco 504, a máquina de análise de vídeo 102 pode calcular um valor de importância de vídeo para o arquivo de vídeo com base nas pontuações de importância de quadro de vídeo dos quadros de vídeo no arquivo de vídeo. A máquina de análise de vídeo 102 pode computar cada pontuação de importância de quadro para o arquivo de vídeo como descritos nos blocos 402-412 do processo 400 mostrado na Figura 4. Em várias modalidades, a máquina de análise de vídeo 102 pode calcular a média das pontuações de importância de

quadro de vídeo para derivar o valor de importância de vídeo para o arquivo de vídeo. Em modalidades alternativas, a máquina de análise de vídeo 102 pode estar configurada para calcular a média das pontuações de importância de quadro de vídeo de uma ou mais seções de vídeo de um arquivo de vídeo com a importância mais alta para derivar o valor de importância para o arquivo de vídeo. Em algumas modalidades, o cálculo do valor de importância de vídeo para o arquivo de vídeo pode também incluir tendência de uma ou mais pontuações de importância que estão associadas com os quadros.

[0085] No bloco de decisão 506, a máquina de análise de vídeo 102 pode determinar se arquivos de vídeo devem ser analisados. A máquina de análise de vídeo 102 pode fazer tal determinação com base em uma entrada de seleção recebida de um usuário. Se a máquina de análise de vídeo 102 determinar que existem arquivos de vídeo adicionais a serem analisados ("sim" no bloco de decisão 506), o processo 500 pode retornar para o bloco 502. No bloco 502, a máquina de análise de vídeo 102 pode obter outro arquivo de vídeo da coleção de arquivos de vídeo para análise adicional.

[0086] No entanto, se a máquina de análise de vídeo 102 determinar que não existem arquivos de vídeo adicionais a serem analisados ("não" no bloco de decisão 506), o processo 500 pode prosseguir para o bloco 508. No bloco 508, a máquina de análise de vídeo 102 pode classificar os arquivos de vídeo que são analisados com base em valores de importância de vídeo correspondentes. Por exemplo, os arquivos de vídeo podem ser classificados da maior importância para menor importância, ou vice versa. Uma vez que os arquivos de vídeo são classificados, a máquina de análise de vídeo 102 pode exibir representações de imagem em miniatura dos arquivos de vídeo, nas quais uma seleção de uma representação de imagem em miniatura pode fazer com que um reproduutor de mídia 122 execute

o arquivo de vídeo correspondente, ou outras aplicações para prover seção de vídeo correspondente para compartilhar e/ou editar. Alternativamente, a máquina de análise de vídeo 102 pode prover os dados de classificação para outra aplicação de modo que a aplicação possa exibir as representações em miniatura dos arquivos de vídeo classificados.

[0087] Em algumas modalidades, a máquina de análise de vídeo 102 pode inicialmente tentar classificar as seções de vídeo de um arquivo de vídeo e/ou os arquivos de vídeo com base em pontuações de característica que são obtidas para as características de alto nível dos quadros de vídeo. Em tais modalidades, a máquina de análise de vídeo 102 pode recorrer para obter as pontuações de característica tanto das características de alto nível quanto das características de baixo nível dos quadros de vídeo para produzir classificações quando a tentativa inicial falha devido a uma presença insuficiente de características de alto nível nos quadros de vídeo.

[0088] A Figura 6 é um fluxograma que ilustra um processo exemplar 600 para computar uma pontuação de importância de face para um quadro de vídeo. No bloco 602, o módulo de análise de alto nível 214 pode executar uma detecção de face em um quadro de vídeo para detectar uma ou mais faces. Uma face detectada pode estar faceando uma câmera que capturou o quadro de vídeo ou lateralmente com relação à câmera. Com base nesta detecção, o módulo de análise de alto nível 214 pode gerar uma lista de faces detectadas com as suas posições no quadro de vídeo, a área do quadro de vídeo coberta por cada face, e uma pontuação de confiança de detecção para cada face que indica uma confiança na detecção.

[0089] No bloco 604, o módulo de análise de alto nível 214 pode executar um rastreamento de face para rastrear as uma ou mais faces. Em várias modalidades, o módulo de análise de alto nível 214 pode

rastrear as faces humanas através de múltiplos quadros de vídeo. Deste modo, o módulo de análise de alto nível 214 pode verificar um conjunto de faces que estão presentes em um arquivo de vídeo, assim como rastrear a frequência que cada face aparece no arquivo de vídeo.

[0090] No bloco 606, o módulo de análise de alto nível 214 pode determinar se a tendência de pontuação com base em característica facial deve ser atribuída a um quadro de vídeo. Em várias modalidades, a determinação pode ser feita com base em se uma ou mais características faciais estão presentes ou ausentes no quadro de vídeo. Assim, se o módulo de análise de alto nível 214 determinar que a tendência de pontuação com base em característica facial deve ser atribuída ("sim" no bloco de decisão 606), o processo 600 pode prosseguir para o bloco 608. No entanto, se o módulo de análise de alto nível 214 determinar que a tendência de pontuação com base em característica facial não deve ser atribuída ("não" no bloco de decisão 606), o processo 600 pode prosseguir diretamente para o bloco 610.

[0091] No bloco 608, o módulo de análise de alto nível 214 pode atribuir uma tendência de pontuação para o quadro de vídeo com base em características faciais de pelo menos uma face no quadro de vídeo. As características faciais podem incluir um tamanho de face, movimento de face, e/ou a presença ou ausência de certas características faciais (por exemplo, sorrindo ou não, olhos fechados / abertos, etc.). Por exemplo, uma face no quadro de vídeo cujo tamanho é menor do que o mínimo limite de tamanho ou maior do que um máximo limite de tamanho pode resultar na atribuição de uma pontuação tendência negativa pelo módulo de análise de alto nível 214. Em outro exemplo, a pontuação de importância de um quadro de vídeo que mostra uma face de uma pessoa específica pode ser positivamente ou negativamente tendente de acordo com a distância

da face para o centro do quadro de vídeo ao longo de um eixo geométrico (por exemplo, eixo geométrico x ou eixo geométrico y), de modo que a face que está mais próxima do centro resulta em uma importância mais alta para o quadro de vídeo, e vice. A tendência de pontuação negativa pode ser um fator de ponderação que diminui a pontuação de importância de face para o quadro de vídeo. Em um exemplo adicional, o módulo de análise de alto nível 214 pode atribuir uma tendência de pontuação positiva para cada face que está sorrindo e/ou tem olhos que estão abertos.

[0092] No bloco 610, o módulo de análise de alto nível 214 pode executar pelo menos um de reconhecimento de face ou grupo de faces em pelo menos uma face. Em grupamento de faces, o módulo de análise de alto nível 214 pode agrupar faces que são rastreadas para determinar se as faces que são detectadas em diferentes quadros de vídeo pertencem à mesma pessoa. Em reconhecimento de face, o módulo de análise de alto nível 214 pode utilizar um algoritmo de reconhecimento facial para corresponde cada face humana que é detectada em um quadro de vídeo a uma identidade conhecida de uma pessoa.

[0093] No bloco 612, o módulo de análise de alto nível 214 pode identificar um ou mais personagens principais com base em pelo menos um de dados de rastreamento de face ou dados de reconhecimento de face. Por exemplo, o número de faces em cada grupo é diretamente proporcional à prevalência da face no arquivo de vídeo. Ainda, uma prevalência mais alta indica uma importância mais alta da pessoa com a face, e vice versa. Consequentemente, a face que pertence a um grupo com um mais alto número de faces pode ser determinada pelo módulo de análise de alto nível 214 como pertencendo ao personagem principal. Em outro exemplo, o personagem principal pode ser identificado quando uma face

designada como pertencendo a um personagem principal é detectada como estando presente no quadro de vídeo pelo algoritmo de reconhecimento facial.

[0094] No bloco 614, o módulo de análise de alto nível 214 pode atribuir uma tendência de pontuação positiva para o quadro de vídeo para cada presença de um personagem principal no quadro de vídeo. A pontuação de característica positiva pode elevar uma pontuação de importância de face que é calculada para o quadro de vídeo. A tendência de pontuação positiva pode ser um fator de ponderação que aumenta a pontuação de importância de face para o quadro de vídeo.

[0095] No bloco 616, o módulo de análise de alto nível 214 pode computar uma pontuação de importância de face para o quadro de vídeo. A pontuação de importância de face pode ser calculada proporcional ao tamanho e/ou movimento de cada face no quadro de vídeo. A computação da pontuação de importância de face pode ser calculada com base nas tendências de pontuação positiva e/ou negativa.

[0096] A Figura 7 é um fluxograma que ilustra um processo exemplar 700 para determinar seções de importância dentro de um vídeo analisando o movimento de pontos de característica. No bloco 702, o módulo de análise de movimento 216 pode obter um quadro de vídeo de um arquivo de vídeo, tal como o arquivo de vídeo 108(1). O módulo de decodificador de vídeo 210 pode ter decodificado o quadro de vídeo do arquivo de vídeo para análise pelo módulo de análise de movimento 216.

[0097] No bloco de decisão 704, o módulo de análise de movimento 216 pode determinar se o final do arquivo de vídeo foi alcançado. Se o módulo de análise de movimento 216 determinar que o final do arquivo de vídeo não foi alcançado ("não" no bloco de decisão 704), o processo 700 pode continuar para o bloco 706.

[0098] No bloco 706, o módulo de análise de movimento 216 pode detectar pontos de característica no quadro de vídeo. Em várias modalidades, o módulo de análise de movimento 216 pode diminuir a resolução do quadro de vídeo e criar uma pirâmide de imagens de resolução diminuída de menores dimensões. As imagens de resolução diminuída são então comparada pelo módulo de análise de movimento 216 para determinar pontos comuns (isto é, pontos de característica) entre as imagens de resolução diminuída.

[0099] No bloco de decisão 708, o módulo de análise de movimento 216 pode determinar se o quadro de vídeo é o primeiro quadro de vídeo do arquivo de vídeo. Consequentemente, se o módulo de análise de movimento 216 determinar que o quadro de vídeo é o primeiro quadro de vídeo ("sim" no bloco de decisão 708), o processo 700 pode retornar para o bloco 702. Quando retornando para o bloco 702, o módulo de análise de movimento 216 pode obter outro quadro de vídeo do arquivo de vídeo. No entanto, se o módulo de análise de movimento 216 determinar que o quadro de vídeo não é o primeiro quadro de vídeo do vídeo ("não" no bloco de decisão 708), o processo 700 pode prosseguir para o bloco 710.

[00100] No bloco 710, o módulo de análise de movimento 216 pode coincidir os pontos de característica no quadro de vídeo a um conjunto adicional de pontos de característica em um quadro de vídeo precedente do arquivo de vídeo. Em várias modalidades, o módulo de análise de movimento 216 pode executar a coincidência determinando uma transformada que alinha os dois quadros adjacentes de modo que um máximo número de pontos de característica coincidam. Em pelo menos uma modalidade, a transformada por ser executada utilizando coincidência geométrica que uma implementação de estimativa de parâmetro robusto.

[00101] No bloco 712, o módulo de análise de movimento 216 pode

calcular uma matriz de transformada de homografia que descreve o movimento dos pontos de característica entre os quadros de vídeo. Em pelo menos algumas modalidades, o módulo de análise de movimento 216 pode utilizar um algoritmo de Amostragem Randômica e Consenso (RANSAC) para obter a matriz de transformada de homografia.

[00102] No bloco 714, o módulo de análise de movimento 216 pode computar dados de movimento para o quadro de vídeo da matriz de transformada de homógrafo. Em várias modalidades, o módulo de análise de movimento 216 pode extrair a magnitude e direção do zoom e componentes de translação vertical utilizando a matriz transformada de homógrafo. Estas magnitudes e direções representam um movimento intencional de uma câmera que gravou os dois quadros adjacentes. Subsequentemente, o processo 700 pode retornar para o bloco 702, de modo que o módulo de análise de movimento 216 possa obter outro quadro de vídeo do arquivo de vídeo para processamento.

[00103] Retornando ao bloco de decisão 704, se o módulo de análise de movimento 216 determinar que o final do arquivo de vídeo foi alcançado ("sim" no bloco de decisão 704), o processo 700 pode continuar para o bloco 716. No bloco 716, o módulo de análise de movimento 216 pode calcular uma média de execução para os dados de movimento dos quadros de vídeo e instabilidade de movimento de quadro para cada quadro de vídeo. Os dados de movimento dos um ou mais quadros de vídeo do arquivo de vídeo podem ser combinados antes do cálculo. Em várias modalidades, o módulo de análise de movimento 216 pode determinar o movimento de instabilidade da câmera que gravou os quadros de vídeo calculando a média de execução de dados de movimento dos vetores de movimento dos quadros de vídeo, e subtrair o movimento intencional da câmera da média de execução de dados de movimento. O cálculo da média de

execução de dados de movimento suprime a variância local e preserva as tendências de longo prazo que representam o movimento intencional. Em outras palavras.

[00104] No bloco 718, o módulo de análise de movimento 216 pode verificar as zonas de cruzamento zero para encontrar informações de pico e vale locais, isto é, máximos e mínimos locais, para os dados de movimento. As informações de pico e vale locais podem indicar pontos de transição de cena para o arquivo de vídeo.

[00105] No bloco 720, o módulo de análise de movimento 216 pode segmentar o arquivo de vídeo em seções de vídeo com base nas informações de pico e vale locais. Mais ainda, o módulo de análise de movimento 216 pode combinar e calcular a média das magnitudes para zoom e movimento de panoramização para cada seção de vídeo. A quantidade de aceleração representada pela média das magnitudes de zoom e movimento de panoramização para uma seção de vídeo pode ser utilizada pelo módulo de análise de movimento 216 para atribuir uma pontuação de importância de movimento para a seção de vídeo. Subsequentemente, o módulo de análise de movimento 216 pode designar uma ou mais seções de vídeo que têm as mais altas pontuações de importância de movimento como seções de importância do arquivo de vídeo.

[00106] A utilização das técnicas aqui descritas podem permitir um usuário classificar arquivos de vídeo com base em sua importância para o usuário. Com base em tal classificação, o usuário pode decidir quais arquivos de vídeo guardar e quais arquivos de vídeo apagar. Em alguns casos, o usuário pode também utilizar as classificações dos arquivos de vídeo para determinar se postar arquivos de vídeo específicos em um website de compartilhamento online. As técnicas aqui descritas podem também apresentar representações de imagem em miniatura que representam seções de importância de um arquivo

de vídeo, de modo que o usuário possa dizer em um olhar as porções interessantes de um arquivo de vídeo. Tais informações podem ajudar o usuário em editar o arquivo de vídeo para aperfeiçoar a qualidade de conteúdo ou destacar seções específicas do arquivo de vídeo.

CONCLUSÃO

[00107] No fechamento, apesar das várias modalidades terem sido descritas em uma linguagem específica a características estruturais e/ou atos metodológicos, deve ser compreendido que o assunto definido nas representações anexas não está necessariamente limitado às características ou atos específicos descritos. Ao invés as características e atos específicos estão descritos como formas exemplares de implementar o assunto reivindicado.

REIVINDICAÇÕES

1. Meio de armazenamento de computador que armazena um método **caracterizado pelo fato de que** compreende:

decodificar (402) um arquivo de vídeo para obter pelo menos um do um quadro de vídeo ou dados de áudio associados com o quadro de vídeo;

analisar (404) uma ou mais características de alto nível do pelo menos um do quadro de vídeo ou dos dados de áudio associados com o quadro de vídeo com base pelo menos em uma regra local ou uma regra global para obter uma ou mais pontuações de característica para o quadro de vídeo, a uma ou mais características de alto nível incluindo pelo menos uma característica que é relacionada a um movimento ou uma posição de uma face humana no quadro de vídeo;

combinar (412) a uma ou mais pontuações de característica do quadro de vídeo para derivar uma pontuação de importância de quadro para o quadro de vídeo; e

segmentar (416) o arquivo de vídeo em seções de vídeo de diferentes valores de importância de seção com base pelo menos nas pontuações de importância de quadro do quadro de vídeo e outras pontuações de importância de quadro de quadros de vídeo adicionais do arquivo de vídeo.

2. Meio de armazenamento de computador, de acordo com a reivindicação 1, **caracterizado pelo fato de que** ainda compreende classificar (420) as seções de vídeo de acordo com os valores de importância de seção.

3. Meio de armazenamento de computador, de acordo com a reivindicação 1, **caracterizado pelo fato de que** ainda compreende:

calcular (504) um valor de importância de vídeo para o arquivo de vídeo com base pelo menos nas pontuações de

importância de quadro de todos os quadros de vídeo no arquivo de vídeo; e

classificar (508) o arquivo de vídeo em relação a arquivos de vídeo adicionais com base pelo menos no valor de importância de vídeo do arquivo de vídeo e outros valores de importância de vídeo dos arquivos de vídeo adicionais.

4. Meio de armazenamento de computador, de acordo com a reivindicação 3, **caracterizado pelo fato de que** ainda compreende armazenar metadados referentes a um ou mais quadros de vídeo no arquivo de vídeo, os metadados incluindo uma ou mais pontuações de característica de um quadro de vídeo.

5. Meio de armazenamento de computador, de acordo com a reivindicação 1, **caracterizado pelo fato de que** a regra local afeta uma atribuição de importância de característica para o quadro de vídeo e a regra global afeta a atribuição de importância de característica para múltiplos quadros de vídeo no arquivo de vídeo que inclui o quadro de vídeo.

6. Meio de armazenamento de computador, de acordo com a reivindicação 1, **caracterizado pelo fato de que** a uma ou mais características de alto nível incluem pelo menos um de uma posição de uma face humana no quadro de vídeo, uma frequência de ocorrência da face humana sobre múltiplos quadros de vídeo, saliência de quadro do quadro de vídeo, uma potência de áudio dos dados de áudio associados com o quadro de vídeo, ou um tipo de áudio dos dados de áudio associados com o quadro de vídeo.

7. Meio de armazenamento de computador, de acordo com a reivindicação 1, **caracterizado pelo fato de que** a análise ainda inclui analisar uma ou mais características de baixo nível do quadro de vídeo, e em que a combinação ainda inclui combinar as pontuações de característica para as características de alto nível e as características

de baixo nível do quadro de vídeo para derivar a pontuação de importância de quadro para o quadro de vídeo.

8. Meio de armazenamento de computador, de acordo com a reivindicação 6, **caracterizado pelo fato de que** a uma ou mais características de baixo nível incluem pelo menos um dentre uma qualidade de exposição de vídeo do quadro de vídeo, uma qualidade de saturação de quadro do quadro de vídeo, uma variedade de tonalidade de quadro do quadro de vídeo, uma instabilidade de quadro do quadro de vídeo, uma luminosidade média do quadro de vídeo, uma entropia de cor do quadro de vídeo, ou uma diferença de histograma entre o quadro de vídeo e um quadro de vídeo adjacente.

9. Meio de armazenamento de computador, de acordo com a reivindicação 1, **caracterizado pelo fato de que** a análise inclui:

executar (602) detecção de face no quadro de vídeo para detectar uma ou mais faces humanas;

atribuir (606) uma tendência de pontuação negativa ou positiva para o quadro de vídeo com base pelo menos em um tamanho de face, uma posição de face, ou uma ou mais características faciais de pelo menos uma face humana no quadro de vídeo, a uma ou mais características faciais incluindo uma ausência ou presença de pelo menos um sorriso ou olhos abertos na pelo menos uma face humana;

executar (610) pelo menos um dentre reconhecimento de face ou grupamento de face em pelo menos uma face humana da uma ou mais faces humanas;

identificar (612) um ou mais personagens principais com base em pelo menos um dentre dados de grupamento de face ou dados de reconhecimento de face;

atribuir (614) uma tendência de pontuação positiva adicional ao quadro de vídeo para uma presença de um personagem principal no quadro de vídeo; e

computar (616) uma pontuação de importância de face para o quadro de vídeo que leva em conta a tendência de pontuação negativa e a tendência de pontuação positiva.

10. Meio de armazenamento de computador, de acordo com a reivindicação 6, **caracterizado pelo fato de que** a análise da saliência do quadro de vídeo inclui:

calcular uma média de execução de dados de movimento dos quadros de vídeo;

determinar um movimento de instabilidade dos quadros de vídeo;

incluir o movimento de instabilidade na média de execução;

determinar um movimento intencional dos quadros de vídeo; e

subtrair o movimento intencional da média de execução.

11. Método implementado por computador, **caracterizado pelo fato de que** compreende:

decodificar (402) um arquivo de vídeo para obter pelo menos um de um quadro de vídeo ou dados de áudio associados com o quadro de vídeo;

analisar (404) uma ou mais características de pelo menos um dentre o quadro de vídeo ou os dados de áudio associados com o quadro de vídeo com base em pelo menos uma de uma regra local ou uma regra global para obter uma ou mais pontuações de característica para o quadro de vídeo;

combinar (412) a uma ou mais pontuações de característica do quadro de vídeo para derivar uma pontuação de importância de quadro para o quadro de vídeo;

calcular (504) um valor de importância de vídeo para o arquivo de vídeo com base pelo menos nas pontuações de importância de quadro do quadro de vídeo e outras pontuações de

importância de quadro de quadros de vídeo adicionais do arquivo de vídeo;

segmentar (416) o arquivo de vídeo em seções de vídeo de diferentes importâncias de seção com base pelo menos nas pontuações de importância de quadro de quadros de vídeo nos dados de arquivo de vídeo associados com uma pluralidade de quadros de vídeo no arquivo de vídeo; e

classificar (508) o arquivo de vídeo em relação a arquivos de vídeo adicionais com base pelo menos em valores de importância de vídeo do arquivo de vídeo e outros valores de importância de vídeo dos arquivos de vídeo adicionais.

12. Método implementado por computador, de acordo com a reivindicação 11, **caracterizado pelo fato de que** a regra local afeta uma atribuição de importância de característica para o quadro de vídeo e a regra global afeta a atribuição de importância de característica para múltiplos quadros de vídeo no arquivo de vídeo que inclui o quadro de vídeo.

13. Método implementado por computador, de acordo com a reivindicação 11, **caracterizado pelo fato de que** a segmentação do arquivo de vídeo com base pelo menos nos dados de movimento associados com uma pluralidade de quadros de vídeo no arquivo de vídeo inclui:

detectar (706) pontos de característica no quadro de vídeo do arquivo de vídeo detectando pontos comuns entre múltiplas imagens de resolução diminuída do quadro de vídeo;

coincidir (710) os pontos de característica no quadro de vídeo com um conjunto adicional de pontos de característica em um quadro de vídeo precedente do arquivo de vídeo;

calcular (712) uma matriz de transformada de homografia

que descreve movimento dos pontos de característica entre os quadros de vídeo;

computar (714) dados de movimento para o quadro de vídeo da matriz de transformada de homografia, os dados de movimento incluindo magnitude e direção de zoom e componentes de translação vertical do movimento dos pontos de característica;

combinar os dados de movimento para o quadro de vídeo com dados de movimento para um ou mais quadros restantes no arquivo de vídeo;

calcular (716) médias de execução para os dados de movimento dos quadros de vídeo no arquivo de vídeo;

certificar (718) zonas de cruzamento zero para encontrar informações de pico e vale locais para os dados de movimento com base pelo menos nas médias de execução; e

determinar (720) seções de vídeo do arquivo de vídeo com base pelo menos nas informações de pico e vale locais dos dados de movimento.

14. Método implementado por computador, de acordo com a reivindicação 11, **caracterizado pelo fato de que** ainda compreende:

calcular valores de importância de vídeo das seções de vídeo com base pelo menos em pontuações de importância de quadro de quadros de vídeo correspondentes nas seções de vídeo; e

classificar as seções de vídeo de acordo com os valores de importância de seção.

15. Método implementado por computador, de acordo com a reivindicação 11, **caracterizado pelo fato de que** a uma ou mais características incluem saliência de quadro do quadro de vídeo, uma potência de áudio dos dados de áudio associados com o quadro de vídeo, um tipo de áudio dos dados de áudio associados com o quadro

de vídeo, uma qualidade de exposição de quadro do quadro de vídeo, uma qualidade de saturação de quadro do quadro de vídeo, uma variedade de tonalidade de quadro do quadro de vídeo, uma instabilidade de quadro do quadro de vídeo, uma luminosidade média do quadro de vídeo, uma entropia de cor do quadro de vídeo, ou uma diferença de histograma entre o quadro de vídeo e um quadro de vídeo adjacente.

16. Método implementado por computador, de acordo com a reivindicação 11, **caracterizado pelo fato de que** a análise inclui:

executar (602) detecção de face em um quadro de vídeo para detectar uma ou mais faces humanas;

executar (604) rastreamento de face para rastrear a uma ou mais faces humanas;

atribuir (608) uma tendência de pontuação negativa ou tendência de pontuação positiva para o quadro de vídeo com base pelo menos em um tamanho de face, uma posição de face, ou uma ou mais características faciais de pelo menos uma face humana no quadro de vídeo, a uma ou mais características faciais incluindo uma ausência ou presença de pelo menos um sorriso ou olhos abertos na pelo menos uma face humana;

executar (610) pelo menos um reconhecimento de face ou grupamento de face em pelo menos uma face humana da uma ou mais faces humanas;

identificar (612) um ou mais personagens principais com base pelo menos em pelo menos um dentre dados de grupamento de face ou dados de reconhecimento de face;

atribuir (614) uma tendência de pontuação positiva para o quadro de vídeo para uma presença de um personagem principal no quadro de vídeo; e

computar (616) uma pontuação de importância de face para

o quadro de vídeo que leva em conta a tendência de pontuação negativa e a tendência de pontuação positiva.

17. Sistema **caracterizado pelo fato de que** comprehende:

um ou mais processadores;

uma memória que inclui uma pluralidade de componentes executáveis por computador que são executáveis pelos um ou mais processadores para executar uma pluralidade de ações, a pluralidade de ações compreendendo:

decodificar um arquivo de vídeo para obter pelo menos um de um quadro de vídeo ou dados de áudio associados com o quadro de vídeo;

analisar uma ou mais características de pelo menos um dentre o quadro de vídeo ou os dados de áudio associados com o quadro de vídeo com base em pelo menos uma de uma regra local ou uma regra global para obter uma ou mais pontuações de característica para o quadro de vídeo, a regra local afeta uma atribuição de importância do recurso para o quadro de vídeo e a regra global afeta a atribuição de importância do recurso para vários quadros de vídeo no arquivo de vídeo que inclui o quadro de vídeo;

combinar a uma ou mais pontuações de característica do quadro de vídeo para derivar uma pontuação de importância de quadro para o quadro de vídeo;

segmentar o arquivo de vídeo em seções de vídeo de diferentes importâncias de seção com base pelo menos nas pontuações de importância de quadro de quadros de vídeo nos dados de arquivo de vídeo associados com uma pluralidade de quadros de vídeo no arquivo de vídeo; e

classificar as seções de vídeo de acordo com os valores de importância das seções que são calculadas com base, pelo menos, nas pontuações de importância de quadro de vídeo e nos quadros de

vídeo adicionais no arquivo de vídeo.

18. Sistema, de acordo com a reivindicação 17, **caracterizado pelo fato de que** ainda compreende:

calcular (504) um valor de importância de vídeo para o arquivo de vídeo com base pelo menos nas pontuações de importância de quadro de todos os quadros de vídeo no arquivo de vídeo; e

classificar (508) o arquivo de vídeo em relação a arquivos de vídeo adicionais com base pelo menos no valor de importância de vídeo do arquivo de vídeo e outros valores de importância de vídeo dos arquivos de vídeo adicionais.

19. Sistema, de acordo com a reivindicação 17, **caracterizado pelo fato de que** segmentar o arquivo de vídeo inclui:

detectar (706) pontos de característica no quadro de vídeo do arquivo de vídeo detectando pontos comuns entre múltiplas imagens de resolução diminuída do quadro de vídeo;

coincidir (710) os pontos de característica no quadro de vídeo com um conjunto adicional de pontos de característica em um quadro de vídeo precedente do arquivo de vídeo;

calcular (712) uma matriz de transformada de homografia que descreve movimento dos pontos de característica entre os quadros de vídeo;

computar (714) dados de movimento para o quadro de vídeo da matriz de transformada de homografia, os dados de movimento incluindo magnitude e direção de zoom e componentes de translação vertical do movimento dos pontos de característica;

combinar os dados de movimento para o quadro de vídeo com dados de movimento para um ou mais quadros restantes no arquivo de vídeo;

calcular (716) médias de execução para os dados de

movimento dos quadros de vídeo no arquivo de vídeo;

certificar (718) zonas de cruzamento zero para encontrar informações de pico e vale locais para os dados de movimento com base pelo menos nas médias de execução; e

determinar (720) seções de vídeo do arquivo de vídeo com base pelo menos nas informações de pico e vale locais dos dados de movimento.

20. Sistema, de acordo com a reivindicação 17, **caracterizado pelo fato de que** ainda compreende combinar várias seções de vídeo em um arquivo de vídeo destacado do arquivo de vídeo.

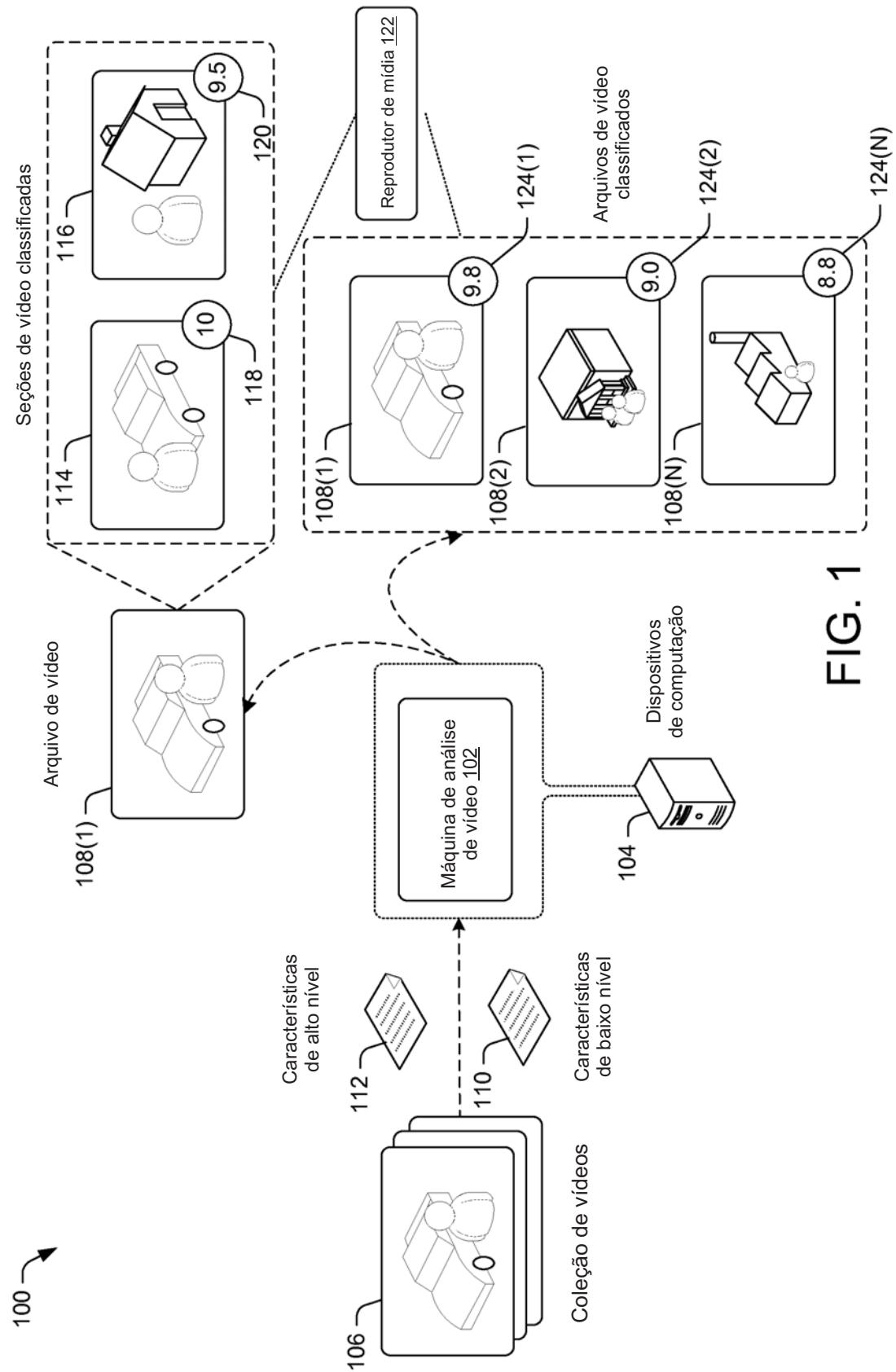
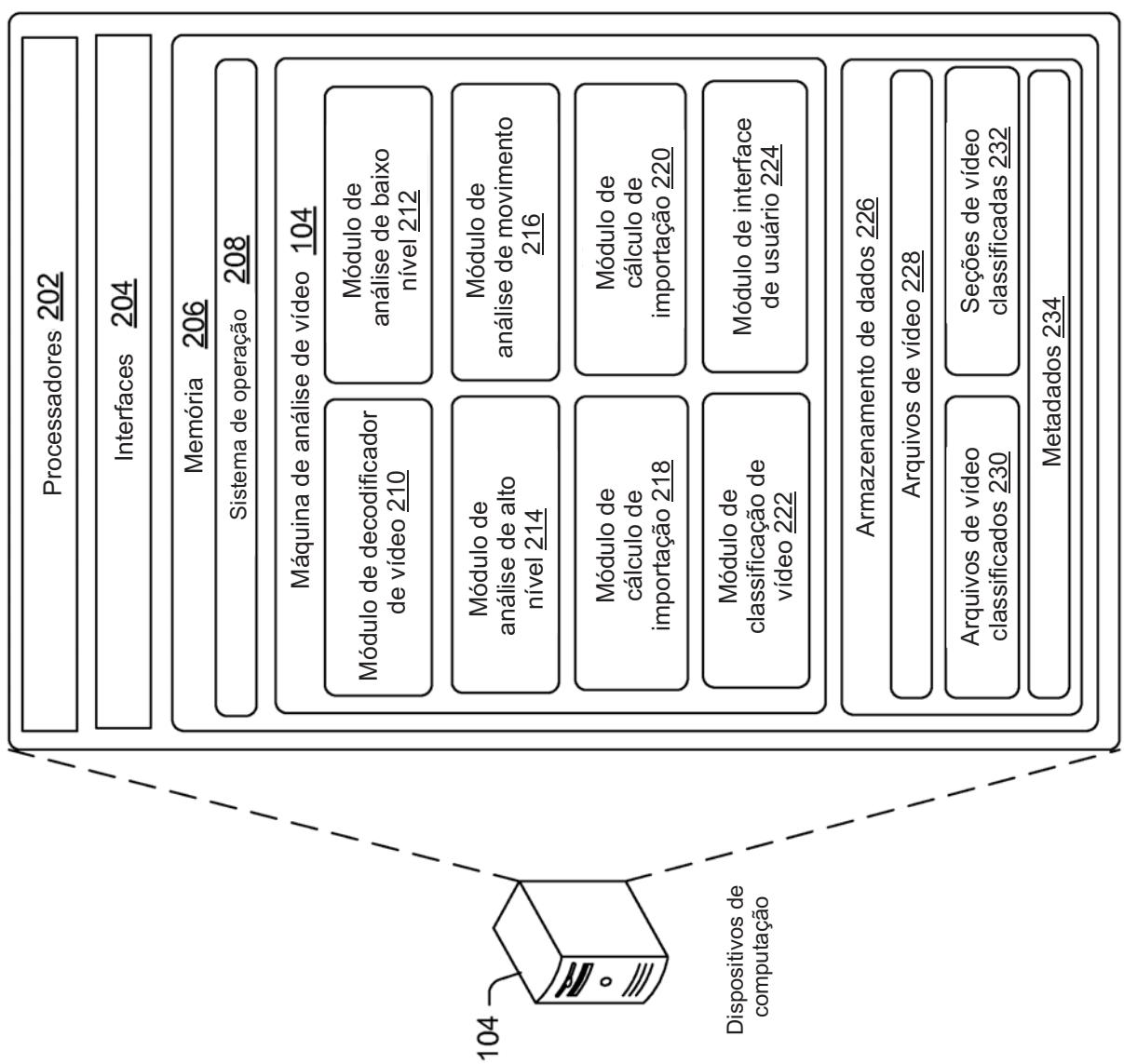
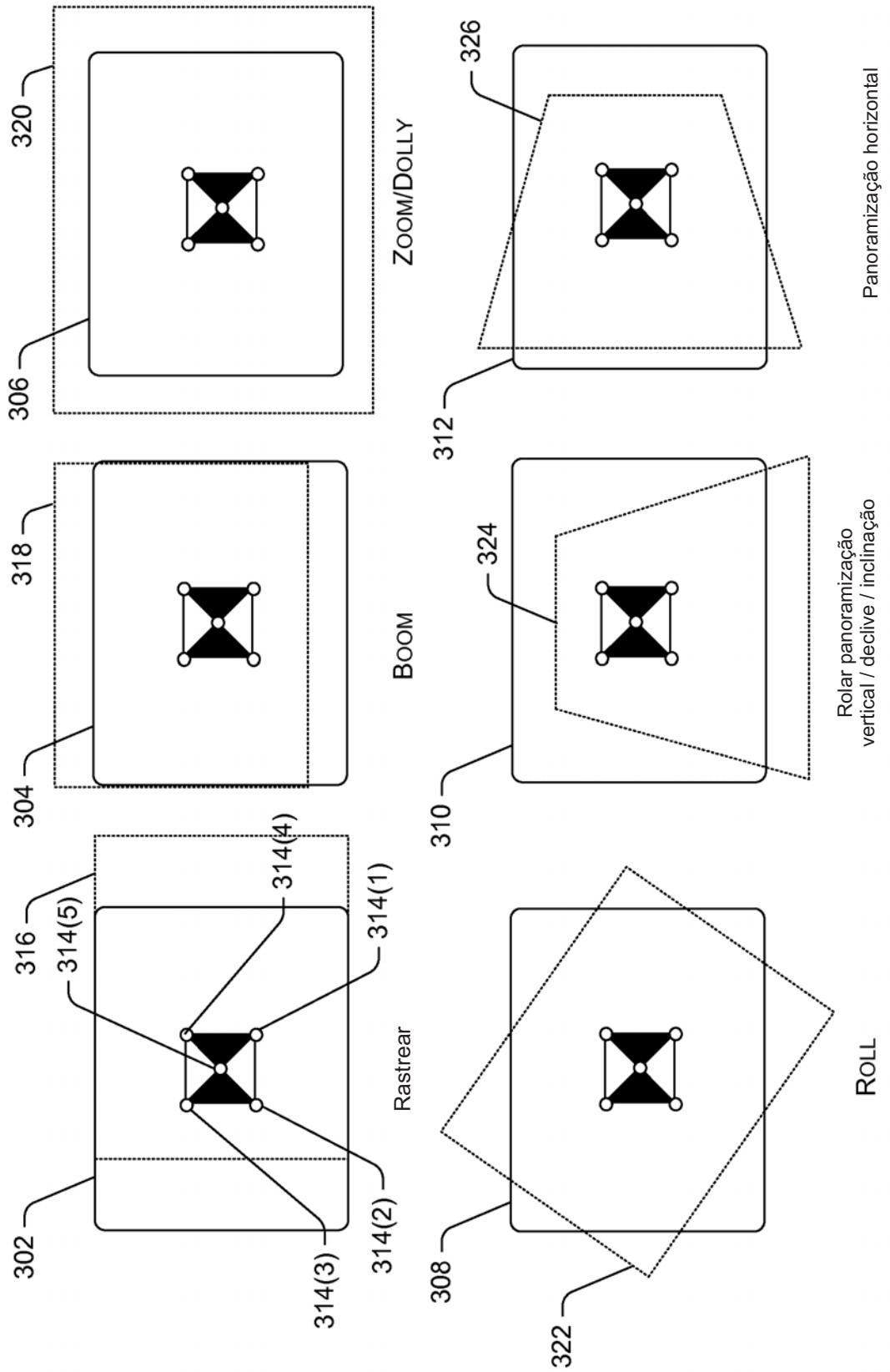


FIG. 1

FIG. 2



**FIG. 3**

Panoramização horizontal

Rolar panoramização
vertical / declive / inclinação

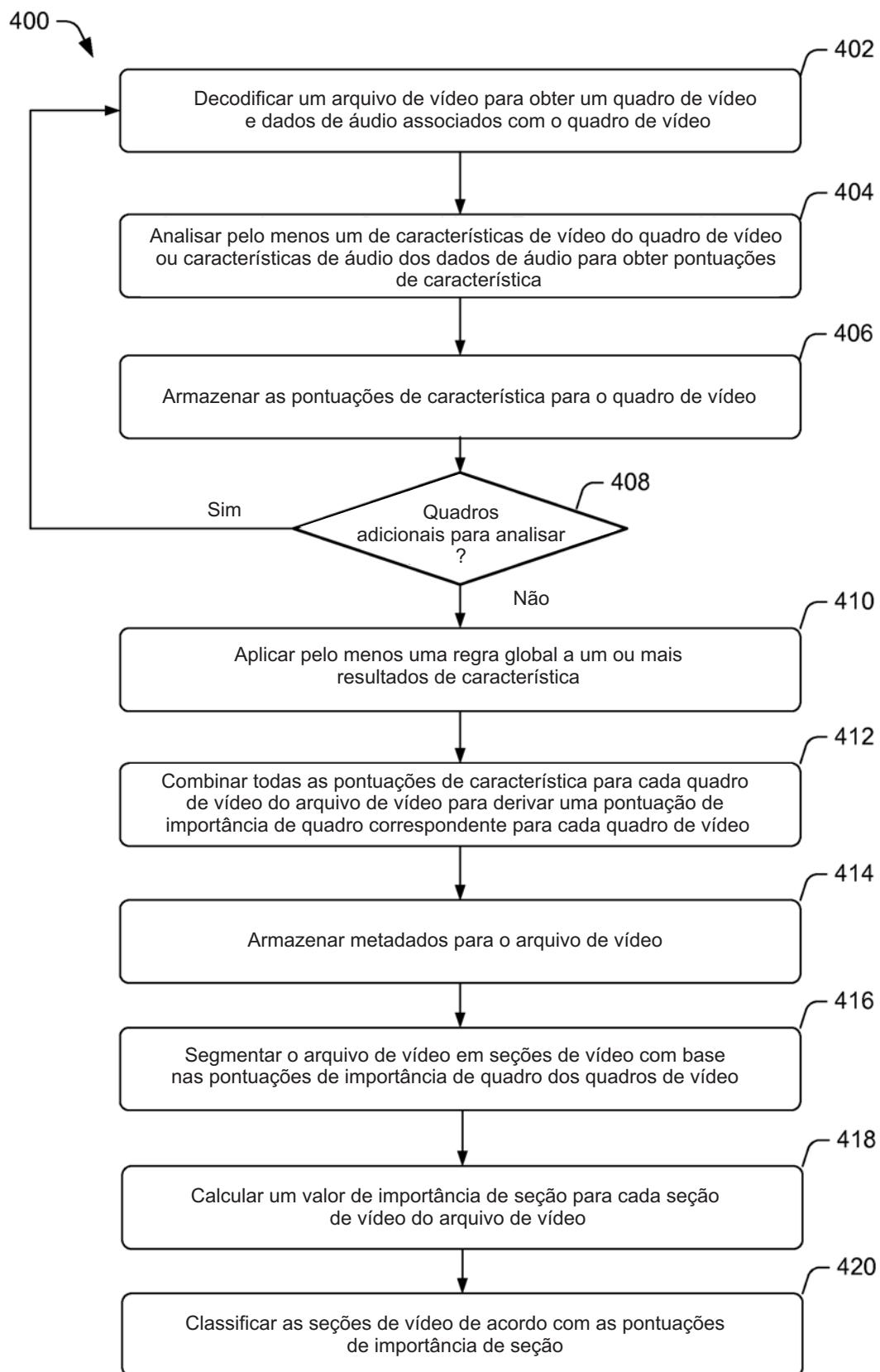


FIG. 4

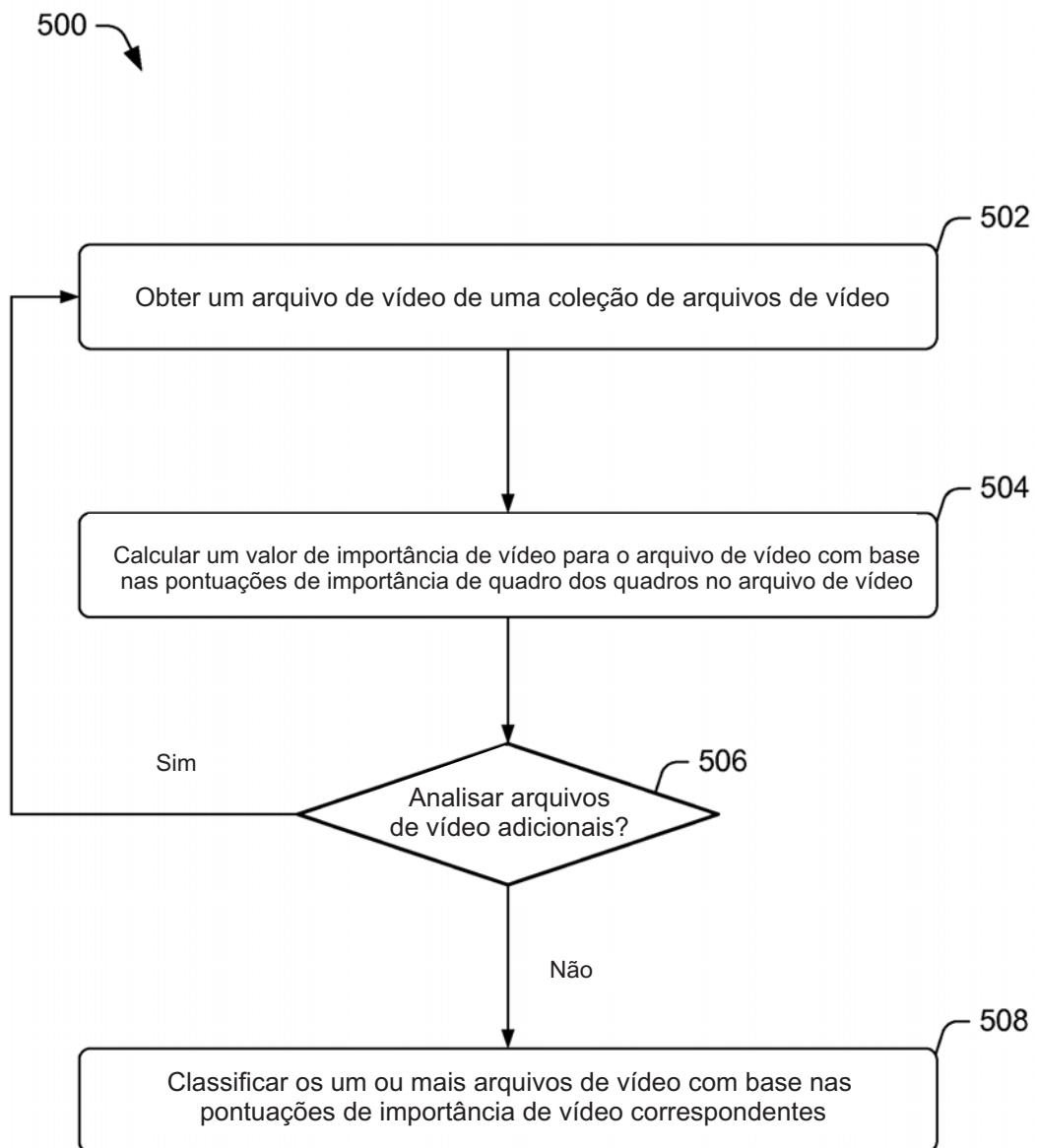


FIG. 5

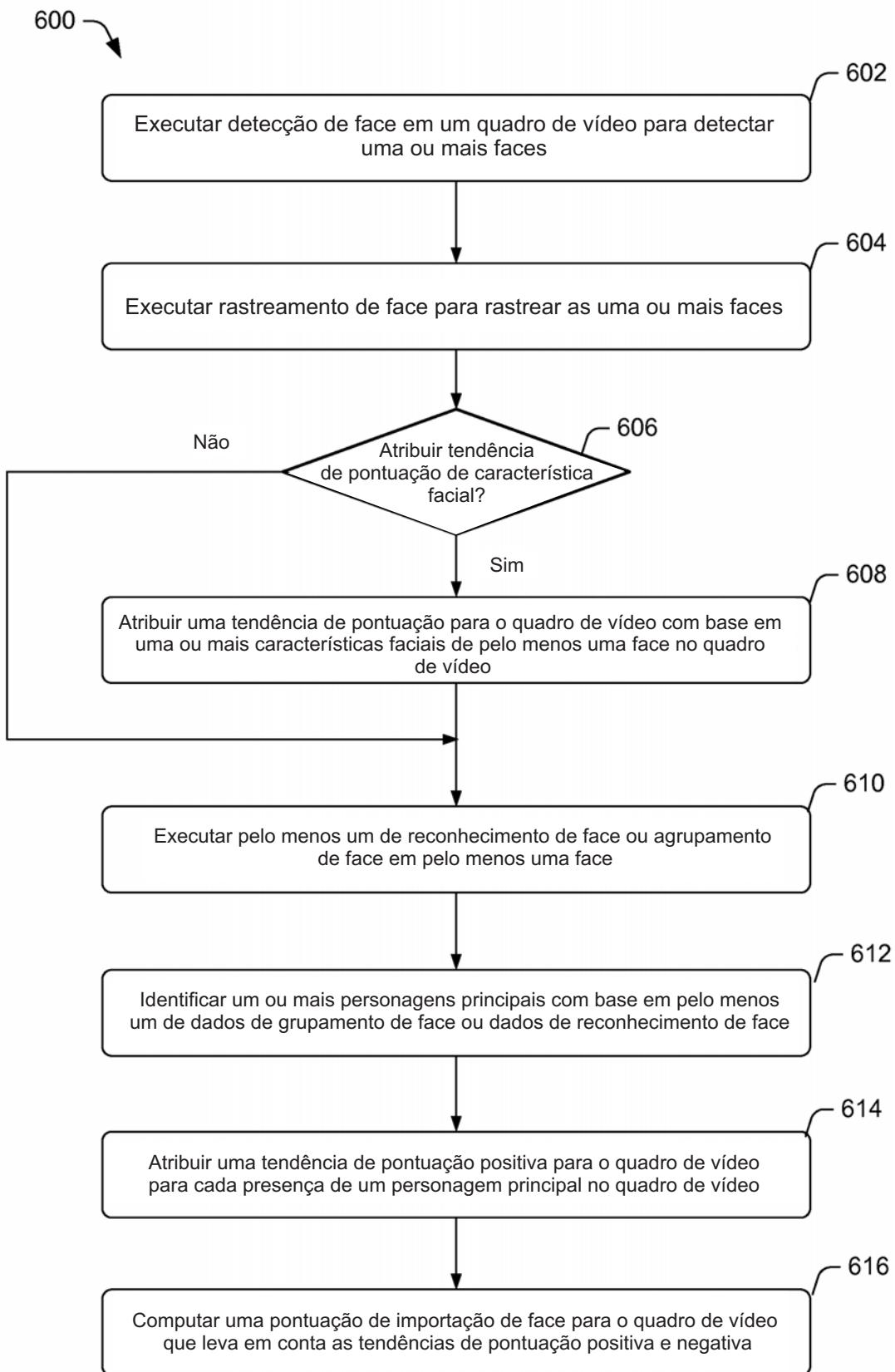


FIG. 6

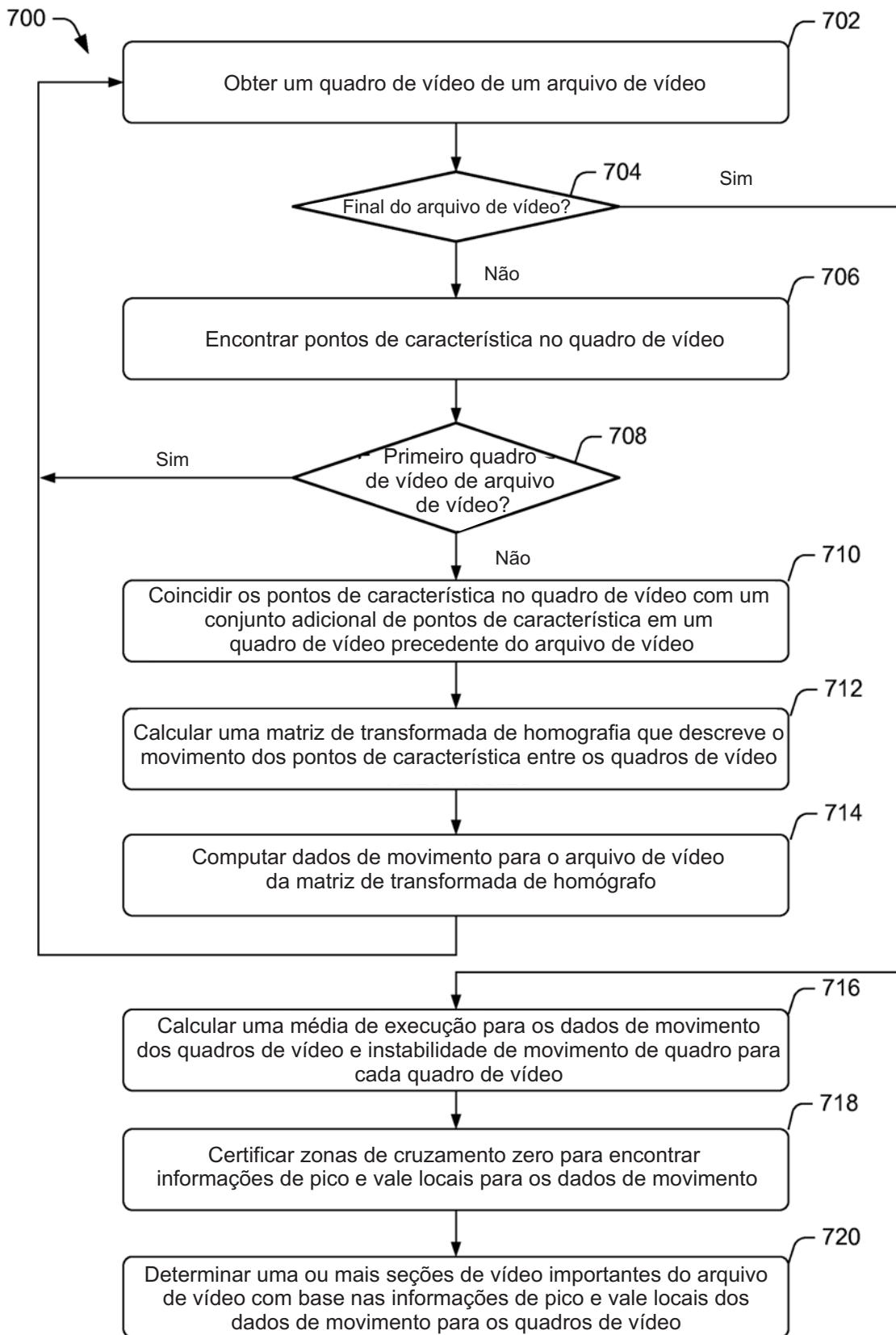


FIG. 7