

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

(21) **PI0706803-4 A2**



(22) Data de Depósito: 30/01/2007
(43) Data da Publicação: 05/04/2011
(RPI 2100)

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

(54) Título: **MÉTODO E APARELHO PARA ESTIMATIVA DE ANIMAÇÃO USANDO BI-PREDIÇÃO DE REFERÊNCIAS COMBINADAS**

(30) Prioridade Unionista: 02/02/2006 US 60/764.452

(73) Titular(es): Thomson Licensing

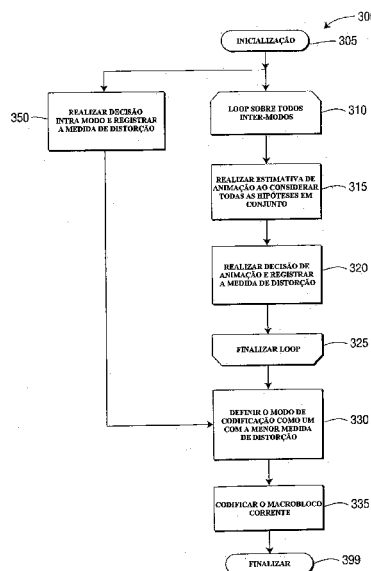
(72) Inventor(es): Alexandros Tourapis, Jill MacDonald Boyce, Peng Yin

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

(86) Pedido Internacional: PCT US2007002381 de 30/01/2007

(87) Publicação Internacional: WO 2007/092192 de 16/08/2007

(57) **Resumo:** MÉTODO E APARELHO PARA ESTIMATIVA DE ANIMAÇÃO USANDO BI-PREDIÇÃO DE REFERÊNCIAS COMBINADAS Um método e aparelho são providos para a estimativa de animação usando bi-predição de referências combinadas. O aparelho inclui um codificador (200) para a codificação de uma imagem de multi-predição a partir da combinação de duas ou mais imagens de referência ao se predizer respectivamente um vetor de animação para uma imagem particular dentre duas ou mais imagens de referência em um processo de estimativa de animação enquanto se inicializa os vetores de animação de modo a manter algumas dentre duas ou mais imagens de referência em um valor predefinido para uso pelo processo de estimativa de animação.





"MÉTODO E APARELHO PARA ESTIMATIVA DE ANIMAÇÃO
USANDO BI-PREDIÇÃO DE REFERÊNCIAS COMBINADAS"

REFERÊNCIA CRUZADA COM PEDIDOS RELACIONADOS

O presente pedido reivindica o benefício do Pedido
5 de Patente Provisório dos Estados Unidos N. de Série 60/764
452, depositado em 02 de fevereiro de 2006, incorporado à
guisa de referência ao presente documento em sua totalidade.

CAMPO DA INVENÇÃO

A presente invenção refere-se, de modo geral, a
10 uma codificação de vídeo e, mais particularmente, a um
método e aparelho para a estimativa de animação usando uma
bi-predição de referência combinada.

FUNDAMENTOS DA INVENÇÃO

É bem conhecido que a compensação de animação de
15 múltiplas hipóteses, conforme descrita em uma primeira
abordagem da técnica anterior, pode oferecer consideráveis
benefícios dentro dos codificadores e decodificadores de
vídeo compensados por animação. Em termos mais específicos,
fatias bi-preditivas (B) (ou imagens bi-preditivas em
20 padrões ou recomendações mais antigas), que consideram 2
hipóteses, são, em geral, as fatias codificadas mais
eficazes dentro do Codec para a Organização Internacional
para a Comissão de Padronização / Internacional de
Eletrotécnica (ISO/IEC) do Grupo de Especialistas em Imagens
25 de Animação-4 (MPEG-4) da União de Telecomunicação
Internacional / padrão de Codificação de Vídeo Avançado
(AVC) Parte 10, da recomendação H.264 do Setor de
Telecomunicações (ITU-T) (doravante o padrão MPEG-4 AVC).

Este comportamento se deve a estas fatias que conseguem mais eficientemente explorar a correlação temporal que existe dentro de uma seqüência por meio da combinação linear de duas ou mais hipóteses em conjunto, de modo a reduzir o seu erro associado. Em termos mais específicos, nas fatias B, um macrobloco ou bloco é codificado de tal modo que o mesmo possa ser predito por uma predição (lista0 ou lista1) ou pela combinação linear de duas predições (lista0 e lista1), embora ponderações associadas a cada lista possam prover benefícios adicionais na presença de fades ou cross-fades. Para realizar esta predição, o decodificador requer que apenas 1 ou 2 vetores de animação (MV), dependendo do tipo de predição, e suas referências associadas (uma para cada lista associada), sejam transmitidos dentro do fluxo de bits, ou que estes sejam inferidos como no caso de modos diretos.

Aparentemente, em muitos codificadores, tais como, por exemplo, na versão JM7.4 do Software de Referência do Time Conjunto de Vídeo (Joint Video Team) (JVT) do software de referência JM atual, a estimativa de animação para as fatias B (e para a codificação de múltiplas hipóteses em geral) considera cada referência de candidato dentro das listas de predição disponíveis separadamente durante uma estimativa de animação e não faz nenhuma presunção específica para uma bi-predição (ou para uma predição de múltiplas hipóteses, respectivamente). Para cada lista de predição, um codificador calcula os melhores vetores de animação de predição única e, em seguida, estes vetores de

animação candidatos bi-preditivos que serão mais adiante usados dentro de uma decisão de modo final, que decidirá qual modo (predição única ou múltipla ou mesmo intra) será usado.

5 Voltando à Figura 1, um codificador de vídeo que usa uma bi-predição é indicado de modo geral pelo numeral de referência 100. Uma entrada não invertida de um combinador 105, uma primeira entrada de uma Decisão de Modo (MD) & Compensação de Animação (MC) 175, uma primeira entrada de um
10 estimador de animação (ME) 165, e uma primeira entrada de um estimador de animação 170 encontram-se disponíveis como entradas no codificador de vídeo. Uma saída do combinador 105 é conectada em comunicação de sinal a uma entrada de um transformador 110. Uma saída do transformador 110 é
15 conectada em comunicação de sinal a uma entrada de um quantizador 115. Uma saída do quantizador 115 é conectada em comunicação de sinal a uma entrada de um codificador de tamanho variável (VLC) 120. Uma saída do codificador VLC encontra-se disponível como uma saída do codificador de
20 vídeo 100.

 A saída do quantizador 115 é também conectada em comunicação de sinal a uma entrada de um quantizador inverso 125. Uma saída do quantizador inverso 125 é conectada em comunicação de sinal a uma entrada de um transformador
25 inverso 130. Uma saída de preferência transformador inverso 130 é conectada em comunicação de sinal a uma primeira entrada não invertida de um combinador 180. Uma saída do combinador 180 é conectada em comunicação de sinal a uma

entrada de um armazenador de referência de imagem 140. Uma
 saída do armazenador de imagem de referência 140 é conectada
 em comunicação de sinal a uma entrada de um armazenador
 temporário de referência List0 145 e a uma entrada de um
 5 armazenador temporário de referência List1 150. Uma primeira
 saída do armazenador temporário de referência List0 145 é
 conectada em comunicação de sinal a uma primeira entrada de
 um multiplicador 155. Uma primeira saída do armazenador
 temporário de referência List1 150 é conectada em
 10 comunicação de sinal a uma primeira entrada de um
 multiplicador 160. Uma segunda saída do armazenador
 temporário de referência List0 145 e uma segunda saída do
 armazenador temporário de referência List1 150 são
 conectadas em comunicação de sinal a uma segunda entrada do
 15 MD&MC 175. Uma saída do multiplicador 155 é conectada em
 comunicação de sinal a uma segunda entrada do estimador de
 animação 165. Uma saída do multiplicador 160 é conectada em
 comunicação de sinal a uma segunda entrada do estimador de
 animação 170. Uma primeira saída do MD&MC 175 é conectada em
 20 comunicação de sinal a uma entrada invertida do combinador
 105. Uma segunda saída do MD&MC 175 é conectada em
 comunicação de sinal a uma segunda entrada não invertida do
 combinador 180.

O método acima se baseia na suposição que estes
 25 vetores de animação são bons o suficiente para serem usados
 dentro da bi-predição. Infelizmente, esta suposição nem
 sempre é verdade, potencialmente resultando em uma perda
 significativa da eficiência. Isto é particularmente

verdadeiro na presença de cross-fades (se dissolve) quando objetos sobrepostos a partir destas imagens podem ter características de luminância consideravelmente diferentes e possivelmente animação, e a consideração de cada lista separadamente pode potencialmente resultar em um desempenho relativamente fraco. Sendo assim, seria altamente desejável poder se considerar em conjunto as referências candidatas disponíveis dentro da fase de estimativa de animação, que poderiam resultar em uma maior eficiência de codificação.

10 Por outro lado, isto não implica que cada referência não deva ser considerada separadamente, uma vez que uma predição única poderia ainda prover bons resultados, especialmente ao considerar que apenas um único conjunto de vetores de animação (mv_0 e mv_1 para as referências x e y , respectivamente) precisa ser transmitido neste caso, o que é

15 muito importante em baixas taxas de bit.

É bem conhecido que a estimativa de animação para um único candidato é ele mesmo muito caro em termos computacionais. Quando uma abordagem de pesquisa completa é usada com uma janela de pesquisa de $(+N, +M)$, isto implicaria

20 que pontos de verificação $(2N+1) \times (2M+1)$ seriam testados. Obviamente, a força bruta e, em um sentido, a abordagem ótima para uma bi-predição requereria que pontos de verificação $(2N+1)^2 \times (2M+1)^2$, o que é proibido em qualquer

25 arquitetura. No caso mais geral de múltiplas hipóteses (k -predição), isto significaria que $(2N+1)^k \times (2M+1)^k$ precisa ser testado. Em uma alternativa, consideravelmente mais simples, uma arquitetura foi apresentada na abordagem da técnica

anterior acima mencionada, na qual, ao invés do método de força bruta, uma abordagem iterativa foi usada, na qual cada hipótese é examinada e refinada sequencialmente ao considerar a hipótese previamente estimada.

5 Este método pode, para um caso bi-preditivo, ser resumido como segue: Suponha que a imagem corrente seja z , e as duas referências sob consideração são as imagens x e y . Em tais imagens, as ponderações a e b são respectivamente selecionadas para uma predição ponderada (isto é, para uma
10 bi-predição normal $a=b=1/2$). mv_0 e mv_1 são os vetores de animação necessários para uma compensação de animação correspondente às referências x e y (ou às suas contrapartidas ponderadas), respectivamente. Para fins de simplificação, os desvios ponderados são ignorados neste
15 processo, embora uma consideração similar possa se aplicar. No procedimento a seguir, SAD (soma da diferença absoluta) é usada como uma medida de distorção.

Etapa 1. Definir $mv_0 = mv'_0 = mv_1 = mv'_1 = 0$

Etapa 2. Formar imagem de referência como ax

20 Etapa 3. Realizar estimativa de animação em ax para refinar os vetores de animação mv_0 usando a distorção $SAD = |z - ax(mv_0) - by(mv'_1)|$

Etapa 4. Definir $mv'_0 = mv_0$

Etapa 5. Formar imagem de referência como by

25 Etapa 6. Realizar estimativa de animação em by para encontrar vetores de animação mv_1 usando a distorção $SAD = |z - by(mv_1) - ax(mv'_0)|$

Etapa 7. Se $(mv_1 = mv'_1)$ sair

Etapa 8. Definir $mv'1=mv1$

Etapa 9. Refinar vetores de animação $mv0$ em ax usando distorção $SAD=|z-ax(mv0)-by(mv'1)|$

Etapa 10. Se $(mv0=mv'0)$ sair

5 Etapa 11. Definir $mv'0=mv0$

Etapa 12. Ir para Etapa 6

Isto pode ser generalizado para o caso de múltiplas hipóteses. O problema com este método é que o mesmo pode ainda requerer um grande número de iterações.

10 Além disso, embora seja muito provável que o desempenho seja aperfeiçoado, é também possível que a predição final possa não ser a melhor possível, especialmente se o esquema estiver aprisionado em um local mínimo. A implementação desta arquitetura é também muito complicada, especialmente

15 considerando que é necessário reconstruir uma nova hipótese usando uma compensação de animação para cada iteração do algoritmo. Uma abordagem alternativa, porém muito similar é proposta na abordagem da primeira técnica anterior acima mencionada, na qual os vetores de animação zero iniciais da

20 etapa 1 são substituídos pelos vetores de animação gerados ao considerar cada lista independentemente.

Os codificadores e decodificadores de compactação de vídeo ganham muito de sua eficiência de compactação ao formar uma predição da imagem corrente (ou fatia) $P_{current}$ que

25 deve ser codificada, e ao codificar ainda a diferença entre esta predição e a imagem em questão. Quanto mais proximamente correlacionada estiver a predição à imagem corrente, menos bits serão necessários para compactar a

imagem. É, portanto, desejável que a melhor predição de imagem possível seja formada. Esta predição pode ser gerada através de métodos espaciais de predição (intra-codificação) ou por métodos temporais (inter-codificação).

5 Os métodos temporais de predição basicamente empregam técnicas compensadas de animação a fim de gerar a referência de predição. Isto é geralmente feito por meio da divisão da imagem fonte em blocos não sobrepostos de um tamanho $N \times M$ e ao encontrar a melhor combinação dentro de uma
10 imagem de referência Preference, usando técnicas de estimativa de animação. Esta melhor combinação é associada a um conjunto de parâmetros de animação também codificados dentro do fluxo de bits. Padrões mais novos, tais como o padrão MPEG-4 AVC, também permitem a consideração de
15 múltiplas imagens de referência para a estimativa e seleção da melhor predição, ao sinalizar o índice da referência usada com os parâmetros de animação. Estes codificadores e decodificadores de múltiplas referências usam um armazenador temporário de referência, no qual cada referência candidata
20 potencial é armazenada e acessada durante a codificação ou decodificação.

Um método alternativo que pode consideravelmente melhorar o desempenho é considerar não somente uma imagem de referência única de uma vez, mas sim a possibilidade de usar
25 combinações de múltiplas hipóteses, como é feito, em particular, para uma imagem bi-preditiva (B) / codificação de fatias. Neste caso, a predição pode ser gerada novamente ao se considerar uma única referência selecionada a partir

de um conjunto de múltiplas referências, mas também por meio da combinação linear (isto é, realizando uma média ponderada) de duas referências disponíveis. Isto também requer que, quando necessário, dois conjuntos de parâmetros diferentes sejam estimados e transmitidos, correspondentes a cada referência. Este conceito pode ser generalizado nos codificadores que consideram mais de duas hipóteses, conforme descrito na abordagem da primeira técnica anterior acima mencionada. Outros parâmetros que também podem aperfeiçoar o desempenho incluem a consideração de uma predição ponderada conforme sugerido em uma segunda abordagem da técnica anterior, na qual um fator de ponderação diferente pode ser aplicado a cada hipótese, e modelos de animação mais complicados, tais como as técnicas de compensação de animação global, são usados.

Embora a consideração de uma compensação de animação de múltiplas hipóteses possa aperfeiçoar o desempenho de um codec de vídeo, uma estimativa apropriada dos parâmetros de animação para este caso é um problema muito complicado. Em particular, a solução ótima pode ser encontrada ao examinar todas as combinações possíveis de predição com as referências disponíveis, ou seja, examinar cada vetor de animação possível em uma referência a todos os outros vetores de animação e suas combinações nas referências restantes, o que obviamente computacionalmente proibido.

SUMÁRIO DA INVENÇÃO

Estas e outras falhas e desvantagem da técnica anterior são tratadas por meio dos presente princípios, direcionados a um método e aparelho para uma estimativa de animação usando uma predição de referência combinada.

5 De acordo com um aspecto da presente invenção, é provido um aparelho. O aparelho inclui um codificador para codificar uma imagem de multi-predição a partir de uma combinação de duas ou mais imagens de referência ao respectivamente predizer um vetor de animação para uma
10 imagem em particular dentre duas ou mais imagens de referência em um processo de estimativa de animação, enquanto se inicializa vetores de animação no sentido de manter algumas dentre duas ou mais imagens de referência em um valor predefinido para uso por meio do processo de
15 estimativa de animação.

De acordo com um outro aspecto da presente invenção, é provido um método. O método inclui a codificação de uma imagem de multi-predição a partir de uma combinação de duas ou mais imagens de referência ao respectivamente
20 predizer um vetor de animação para uma imagem em particular dentre duas ou mais imagens em referência de um processo de estimativa de animação, enquanto se inicializa os vetores de animação de modo a manter uma imagem dentre duas ou mais imagens de referência em um valor predefinido para uso por
25 meio do processo de estimativa de animação.

Estes e outros aspectos, recursos e vantagens da presente invenção tornar-se-ão aparentes a partir da

descrição detalhada a seguir de modalidades exemplares, que devem ser lidas com relação aos desenhos em anexo.

BREVE DESCRIÇÃO DOS DESENHOS

A presente invenção pode ser melhor entendida de acordo com as figuras exemplares a seguir, nas quais:

A Figura 1 mostra um diagrama para um codificador de vídeo que usa bi-predição, de acordo com a técnica anterior;

A Figura 2 mostra um diagrama para um codificador de vídeo exemplar que usa bi-predição com consideração de subtração, de acordo com uma modalidade da presente invenção; e

A Figura 3 mostra um método exemplar para a codificação de um macrobloco usando bi-predição, de acordo com uma modalidade da presente invenção;

A Figura 4 mostra um método exemplar para estimativa de animação usando bi-predição de referência combinada, na qual são feitas iterações paralelas a fim de otimizar a hipótese restante, de acordo com uma modalidade da presente invenção;

A Figura 5 mostra um método exemplar para a estimativa de animação usando uma bi-predição de referência combinada, na qual vetores de animação são iterativamente refinados com base em uma melhor referência de acordo com uma modalidade da presente invenção;

A Figura 6 mostra um método exemplar para uma estimativa de animação que considera todas as hipóteses como

tendo vetores de animação idênticos, de acordo com uma modalidade da presente invenção;

A Figura 7 mostra um método exemplar para a realização de uma estimativa de animação usando uma bi-
5 predição de referência combinada, na qual cada referência (x ou y) é considerada separadamente como sendo estacionária em comparação a uma imagem fonte x, de acordo com a uma modalidade da presente invenção;

A Figura 8 mostra um método exemplar para a
10 realização de uma estimativa de animação usando uma bi-predição de referência combinada, na qual apenas uma referência é considerada como sendo estacionária em comparação a uma imagem fonte, de acordo com uma modalidade da presente invenção; e

15 A Figura 9 mostra um método exemplar para a realização de uma estimativa de animação que usa uma bi-predição de referência, na qual ambas as referências são combinadas presumindo uma animação zero entre as referências, de acordo com uma modalidade da presente
20 invenção.

DESCRIÇÃO DETALHADA DA INVENÇÃO

A presente invenção trata de um método e aparelho para uma estimativa de animação usando uma bi-predição de referência combinada.

25 A presente descrição ilustra a presente invenção. será, portanto, apreciado que os versados na técnica serão capazes de conceber várias disposições que, embora não explicitamente descritas ou mostradas no presente documento,

incorporam a presente invenção e são incluídas no seu espírito e âmbito de aplicação.

Todos os exemplos e linguagem condicional aqui apresentados pretendem, para fins pedagógicos, ajudar o
5 leitor no entendimento dos princípios da presente invenção e dos conceitos contribuídos pela presente invenção no sentido de fomentar a técnica, e devem ser construídos como tais, sem se limitarem aos exemplos e condições especificamente apresentados.

10 Além disso, todas as explicações no presente documento que descrevem os princípios, os aspectos, e as modalidades da presente invenção, assim como os exemplos específicos da mesma, pretendem abranger seus equivalentes estruturais e funcionais. Adicionalmente, é um objetivo que
15 tais equivalentes incluam os equivalentes tanto correntemente conhecidos como os equivalentes a serem desenvolvidos no futuro, isto é, quaisquer elementos desenvolvidos que realizem a mesma função, independentemente de sua estrutura.

20 Sendo assim, por exemplo, será apreciado pelos versados na técnica que os diagramas em bloco aqui apresentados representam visões conceituais de um circuito ilustrativo que incorpora os princípios da presente invenção. De maneira similar, será apreciado que quaisquer
25 painéis, fluxogramas, diagramas de transição de estado, pseudo-códigos, ou coisa do gênero representam vários processos que podem ser substancialmente representados em um meio legível em computador e assim executados por um

computador ou processador, quer ou não tal computador ou processador seja explicitamente mostrado.

As funções dos vários elementos mostrados nas figuras podem ser providas através do uso de um hardware dedicado, assim como através de um hardware capaz de executar um software em associação com o apropriado software. Quando providas por um processador, as funções podem ser providas por meio de um processador dedicado, por meio de um único processador compartilhado, ou por meio de uma pluralidade de processadores individuais, alguns dos quais podendo ser compartilhados. Além disso, o uso explícito do termo "processador" ou "controladora" não deve ser construído no sentido exclusivo de um hardware capaz de executar um software, e pode implicitamente incluir, sem limitação, um hardware de processador de sinais digitais ("DSP"), uma memória de leitura ("ROM") para armazenar o software, uma memória de acesso aleatório ("RAM"), e um armazenamento não volátil.

Outros hardwares, convencionais e/ou personalizados, podem também ser incluídos. De maneira similar, quaisquer comutadores mostrados nas figuras são apenas conceituais. Suas funções podem ser executadas através da operação de uma lógica de programa, através de uma lógica dedicada, através da interação do controle de programa e da lógica dedicada, ou ainda manualmente, a técnica em particular sendo executável pelo implementador, conforme mais especificamente entendido a partir do contexto.

Nas presentes reivindicações, qualquer elemento expresso como um meio para a execução de uma função específica pretende abranger qualquer maneira de se executar aquela função, incluindo, por exemplo, a) uma combinação de elementos de circuito que realiza aquela função, ou b) um software de qualquer forma, incluindo, portanto, um firmware, um micro-código, ou coisa do gênero, combinado com um circuito apropriado para a execução daquele software no sentido de realizar a função. A presente invenção conforme definida por tais reivindicações é inerente ao fato de que as funcionalidades providas pelos diversos meios apresentados são combinadas e reunidas da maneira que as reivindicações sugerem. É, por conseguinte, considerado que qualquer meio que possa prover estas funcionalidades seja equivalente aos mostrados no presente documento.

Referência, no relatório descritivo, ao termo "uma modalidade" da presente invenção significa que um aspecto, estrutura, característica em particular, e assim por diante, descrito com relação à modalidade é incluído em pelo menos uma modalidade da presente invenção. sendo assim, as citações à expressão "em uma modalidade" que aparecem em vários lugares em todo o relatório descritivo não se referem todas necessariamente à mesma modalidade.

Deve-se apreciar que um pseudocódigo provido no presente documento se presta para fins ilustrativos, e, deste modo, dados os ensinamentos da presente invenção, aqui providos, uma pessoa com habilidade simples nesta e em técnicas relacionadas contemplará que as seguintes

seqüências de pseudocódigo, assim como suas variações, se mantêm no âmbito de aplicação da presente invenção.

Voltando para a Figura 2, um codificador de vídeo exemplar usando bi-predição com consideração de subtração é
5 indicado de modo geral pelo numeral de referência 200.

Uma entrada não invertida de um combinador 205, uma primeira entrada de uma Decisão de Modo (MD) & Compensação de Animação (MC) 275, uma primeira entrada de um estimador de animação (ME) 265, uma primeira entrada de um
10 estimador de animação 270, e uma entrada não invertida de um combinador 286, e uma entrada não invertida de um combinador 288 encontram-se disponíveis como entradas em um codificador de vídeo. Uma saída de um combinador 205 é conectada em comunicação de sinal a uma entrada de um transformador 210.
15 Uma saída do combinador 205 é conectada em comunicação de sinal a uma entrada de um transformador 210. Uma saída do transformador 210 é conectada em comunicação de sinal a uma entrada de um quantizador 215. Uma saída do quantizador 215 é conectada em comunicação de sinal a uma entrada de um
20 codificador de tamanho variável (VLC) 220. Uma saída do codificador VLC 220 encontra-se disponível como uma saída do codificador de vídeo 200.

A saída do quantizador 215 é também conectada em comunicação de sinal a uma entrada de um quantizador inverso
25 225. Uma saída do quantizador inverso 225 é conectada em comunicação de sinal a uma entrada de um transformador inverso 230. Uma saída do transformador inverso 230 é conectada em comunicação de sinal a uma primeira entrada não

invertida de um combinador 280. Uma saída do combinador 280 é conectada em comunicação de sinal a uma entrada de um filtro de loop 235. Uma saída do filtro de loop 235 é conectada em comunicação de sinal a uma entrada de um armazenador de referência de imagem 240. Uma saída do armazenador de imagem de referência 240 é conectada em comunicação de sinal a uma entrada de um armazenador temporário de referência List0 245 e a uma entrada de um armazenador temporário de referência List1 150. Uma primeira saída do armazenador temporário de referência List0 245 é conectada em comunicação de sinal a uma primeira entrada de um multiplicador 255. Uma primeira saída do armazenador temporário de referência List1 250 é conectada em comunicação de sinal a uma primeira entrada de um multiplicador 260. uma segunda saída do armazenador temporário de referência List0 245 e uma segunda saída do armazenador temporário de referência List1 250 são conectadas em comunicação de sinal a uma segunda entrada do MD&MC 275. Uma saída do multiplicador 255 é conectada em comunicação de sinal a uma segunda entrada do estimador de animação 265. Uma saída do multiplicador 260 é conectada em comunicação de sinal a uma segunda entrada do estimador de animação 270. Uma primeira saída do MD&MC 275 é conectada em comunicação de sinal a uma entrada invertida do combinador 205. Uma segunda saída do MD&MC 275 é conectada em comunicação de sinal a uma segunda entrada não invertida do combinador 280.

A saída do multiplicador 255 é também conectada em comunicação de sinal a uma primeira entrada do estimador de animação 284 e a uma entrada invertida do combinador 286. A saída do multiplicador 260 é também conectada em comunicação de sinal a uma entrada invertida do combinador 288 e a uma primeira entrada de um estimador de animação 282. Uma saída do combinador 286 é conectada em comunicação de sinal a uma segunda entrada invertida do estimador de animação 282. Uma saída do estimador de animação 284 é conectada em comunicação de sinal a uma terceira entrada do MD&MC 275. Uma saída do estimador de animação 282 é conectada em comunicação de sinal a uma quarta entrada do MD&MC 275.

De acordo com a presente invenção, um método e aparelho são descritos, os quais podem significativamente aumentar o desempenho de codificação de imagens bi-preditivas ou fatias dentro de uma arquitetura de codificação de vídeo. Embora a presente invenção seja basicamente descrita com relação ao padrão MPEG-4 AVC, deve-se apreciar que a presente invenção não se limita apenas ao padrão acima, e pode ser utilizada com relação a outros padrões de codificação de vídeo e a recomendações, e ainda manter o âmbito de aplicação da presente invenção.

Em uma modalidade da presente invenção, considera-se uma imagem de referência adicional dentro do processo de estimativa de animação, que é basicamente uma combinação linear de duas outras referências. Esta nova referência permite se refinar ainda mais a referência de predição dentro das imagens bi-preditivas (B) (a bi-predição de

referência combinada), uma vez que uma consideração mais precisa de uma bi-predição é realizada, levando a uma eficiência de codificação maior. Em uma modalidade da presente invenção, uma extensão deste método em uma predição de múltiplas hipóteses é provida. Com vantagem, a modalidade diminui a complexidade da estimativa de animação de múltiplas hipóteses sem afetar adversamente a qualidade de qualquer grau significativa.

Na abordagem da primeira técnica anterior acima mencionada, a estimativa de animação para uma compensação de animação de múltiplas hipóteses é feita usando um método iterativo no qual, em cada iteração, os vetores de animação de uma referência em particular ou hipótese são refinados ao mesmo tempo considerando os vetores de animação estimados anteriores de todas as outras listas de referência ou hipóteses. É igualmente sugerido que, como uma estimativa inicial, os vetores de animação (MV) de predição única possam ser usados.

Em uma modalidade da presente invenção, propõe-se o uso de um método diferente para a estimativa dos vetores de animação de uma referência em particular, que, embora ainda baseados em um método iterativo, podem melhor estimar os vetores de animação para cada referência, e ainda podem ser muito mais simples sob algumas condições ou arquiteturas de implementar.

Uma modalidade se baseia na suposição de que imagens adjacentes tendem a incluir um grande percentual de regiões estacionárias, mesmo durante fades e cross-fades.

Diferentemente da abordagem da primeira técnica anterior acima mencionada, na qual é feita uma sequência de iterações no sentido de estimar os vetores de animação de uma referência em particular com base nos vetores de animação calculados de todas as outras referências já estimadas. Em uma primeira modalidade, primeiramente considera-se a predição dos vetores de animação de cada referência ao supor que todas as outras referências restantes são estacionárias (vide Figura 2). Sendo assim, permite-se separar melhor as regiões estacionárias e as não estacionárias e, portanto, pode-se aumentar a precisão da estimativa de animação ou a convergência de um esquema iterativo, como feito na abordagem da primeira técnica anterior. Mais especificamente, esta solução apresenta diversas possibilidades alternativas.

Em uma modalidade, uma primeira abordagem de acordo com a presente invenção para a realização de uma estimativa de animação usando uma bi-predição de referência combinada envolve a realização de iterações paralelas a fim de otimizar as hipóteses restantes. A paralelização é feita ao iniciar o desempenho da estimativa de animação com diferentes índices de referência, ao mesmo tempo presumindo os vetores de animação para os demais índices de referência como 0. este conceito pode ser explicado ao estudar o seguinte pseudocódigo, que pode ser usado para cada bloco e pressupondo que se têm N números de hipóteses. No pseudocódigo a seguir, as seguintes designações são usadas: x_i , que indica a imagem de referência com um índice

interface; w_i que indica a sua ponderação associada a uma predição ponderada; $mv_{j,k,i}$ indica o vetor de animação computado para a referência com o índice j , se ME for realizado começando pelo índice de referência i ; e $k=\{0,1\}$ é
 5 usado para verificar se qualquer mudança nos vetores de animação ocorreu durante o processo de iteração e se o algoritmo precisa terminar ou não. Em particular, se $N=2$, tem-se o seguinte:

```

      para (i = 0; i < 2; i++)
10      {
          mv0,0,i=mv0,1,i=mv1,0,i=mv1,t,i=(0,0);
          / realiza início de ME usando o índice de
referência i
          referência = w1 * xi;
15      se (i==0)j=1;
          ainda j=0;
          fonte = z - wj * xj;
          //! Realiza ME usando referência para refinar
vetores de animação mvi,0
20      //! Usando distorção SAD = | fonte - referência
(mvm,0,i) |.
          Motion_Estimation(source,    reference,    mvi,0,i,
minSADi)
          Fazer
25      {
          para (m=0; m<2; m++)
          {
              referência = Wm * Xm;

```

```

        fonte=z;
        se (m==0)j=1;
        ainda j=0;
        fonte = z - wj * xj(mvj,0,i);
5      //! Realiza ME usando referência para refinar
        mvm,0,i
        //! usando distorção SAD = | fonte - referência
        (mvm,0,i)|.
        Motion_Estimation(source,      reference,      mvi,0,i,
10    minSADi)
        }
        para (j = 0; j < 2; j++)
        {
        se (mvj,0,i!=mvj,1,i)
15    {
        iterate=1
        mvj,1,i=mvj,0,i
        break;
        }
20    ainda iterate=0
        }
        }
        enquanto (iterate==1)
        }

```

25 O pseudocódigo acima resulta em potencialmente 2 combinações de vetores de animação para cada hipótese diferente. Pode-se então selecionar a melhor combinação usando o que leva à menor distorção, ou seja:

```

se (minSAD0<minSAD1)
{
best_mv={mv0,0,0, mv1,0,0}
minSAD= minSAD0;
5      }
    ainda
    {
best_mv={mv0,0,1, mv1,0,1}
minSAD= minSAD1;
10      O algoritmo mais geral para N hipóteses é como
segue:
para (i = 0; i < N; i++)
{
referência = wj * xi;
15      fonte = z
para (j = 0; j < N;j++)
{
mvj,0,i=mvj,1,i=(0,0);
se (j!=i) source=source wj * wj(mvj,0,i);
20      }
    //! Realiza ME usando referência para refinar
vetores de animação mvi,0,i
    //! usando distorção SAD = | fonte - referência
(mvm,0,i)|.
25      Motion_Estimation(source,      reference,      mvi,0,i,
minSADi)
Fazer
{

```

```

    para (m=0; m<N; m++)
    {
        referência = Wm * Xm;
        fonte=z;
5      para (j = D; j < N; j++)
        {
            se (j!=m) source=source - wj * xj(mvj,0,i);
        }
        //! Realiza ME usando referência para refinar
10     mvm,0
            //! usando distorção SAD = | source -
            reference(mvm,0,i) |.
            Motion_Estimation(source, reference, mvm,0,i,
minSADi)
15     }
        para (j = 0; j < N; j++)
        {
            se (mvj,0,i!=mvj,1,i)
            {
20         iterate=1
            mvj,1,i=mvj,0,i
            break;
            }
            ainda iterate=0
25         }
        }
        enquanto (iterate==1)
    }

```

Neste caso, pode-se ter até N combinações possíveis de vetores de animação para cada hipótese diferente. Pode-se então selecionar a melhor combinação usando o seguinte método:

```

5      minSAD=minSAD0;
      best_mv={mv0,0,0, mv1,0,0, ..., mvm,0,0, ..., mvN-1,0,0}
      para (i = 1; j < N-1; i++)
      {
        se (minSADi < minSAD)
10     {
        minSAD=minSADi;
        best_mv={ mv0,0,i, mv1,0,i, ..., mvm,0,i, ..., mvN-1,0,i}
        }
      }

```

15 Esta decisão pode também ser feita dentro do processo de decisão de modo (usando a Soma do Erro Quadrado (SSE) como a métrica da distorção, combinado com os bits em questão requeridos para codificar os vetores de animação com um dado parâmetro Lagrangiano), enquanto se pode também

20 permitir combinações de vetores de animação a partir de cada conjunto, ou mesmo de outros preditores, ou vetores de animação previamente estimados, tais como os vetores de animação diretos espaciais / temporais, os vetores de animação de direção única, os vetores de animação zero, os

25 vetores de animação colocados temporariamente, e assim por diante, os quais podem levar a hipóteses extras dentro da decisão de modo. Observa-se que, dependendo da implementação, o número máximo de iterações, ou mesmo o

número de refinamentos paralelos feitos, pode ser predefinido, reduzindo, assim, a complexidade de um dado sistema.

Em uma modalidade, uma segunda abordagem de acordo com a presente invenção para a realização de uma estimativa de animação usando bi-predição de referência combinada será descrita a seguir. Deve-se apreciar que a realização de todas as iterações possíveis conforme descritas na primeira abordagem de acordo com a presente invenção é ainda muito complexa. a segunda abordagem, de acordo com a presente invenção, envolve a realização apenas da primeira etapa para cada hipótese; e em seguida a realização de um refinamento iterativo com base apenas no melhor (com base em um critério de distorção mínima). Isto pode reduzir consideravelmente a complexidade, e ainda obter um desempenho similar. O pseudocódigo para a segunda abordagem é como segue:

```

    para (i = 0; i < N; i++)
    {
        referência = wi * xi;
        fonte = z;
    20   para (j = 0 < N; j++)
        {
            mvj,0,j=mvj,1,i=(0,0);
            se (j!=i) source=source wj * wj(mvj,0,i);
    25   }

    //! Realiza ME usando referência para refinar
    vetores de animação mvi,
```

```

        //! usando distorção SAD = | fonte - referência
(mvm,0,i)|.
        Motion_Estimation(source,    reference,    mvi,0,i,
minSADi)
5          }
        minSAD= minSADi;
        best_mv={mv0,0,0,0,0}
        best_ref=0;
        para (i = 1; I < N-1; i++)
10       {
            se (minSADi < minSAD)
            {
                minSAD= minSADi;
                best_mv={0,0, ..., mvi,0,i,0,0}
15             best_ref=i;
            }
            fazer
            {
                para (m=0; m< N; m++)
20             {
                    referência = Wm * Xm;
                    fonte=z;
                    para (j = 0; j < N; j++)
                    {
25             se (j!=m) source=source- wj * wj(mvj,0, best_ref);
                    }

                //! Realiza ME usando referência para refinar
vetoires de animação mvi,0

```

```

        //!< usando distorção SAD = | fonte - referência
(mvm,0,i)|.

        Motion_Estimation(source,                reference,
mvm,0,best_ref, minSADbest_ref)
5          }

        para (j = 0; j < N; j++)
        {
            se (mvj,0,best_ref!=mvj,1,best_ref)
            {
10              iterate=1
              mvj,1,best_ref=mvj,0,best_refi
              break;
            }

            ainda iterate=0
15          }:
          }

          enquanto (iterate==1)
            minSAD=minSADbest_ref;

            best_mv={mv0,0,best_ref, mv1,0,best_ref, mvM,0,best_ref, ... ,
20 mvN-1,0,best_ref}

```

A seleção de referência inicial pode também se basear em quão similar (ou mais especificamente dissimilar) é uma referência para a imagem fonte. Por exemplo, pode-se iniciar o refinamento da referência com a mais alta

25 distorção em comparação à fonte, enquanto todos os outros são inicialmente considerados estacionários, e as etapas de iteração são feitas conforme previamente descritas com

relação à primeira abordagem de acordo com a presente invenção.

Em uma modalidade, uma terceira abordagem de acordo com a presente invenção para a realização de uma
5 estimativa de animação usando bi-predição de referência combinada será descrita a seguir. Observa-se que foram considerados vetores de animação zero para a etapa inicial do algoritmo exemplar nas duas abordagens acima de acordo com a presente invenção. no entanto, é possível considerar
10 também os vetores de animação previamente gerados (por exemplo, único ou qualquer número M de hipóteses) como soluções alternativas e selecionar o melhor para o processo de refinamento. Isto pode ser feito examinando qual dentre estes preditores, tais como os vetores de animação diretos
15 temporais / espaciais, os vetores de animação de direção única, os vetores de animação zero, os vetores de animação colocados temporariamente, e assim por diante, os quais podem resultar em uma imagem fonte com a menor distorção com relação à referência inicial sem a necessidade de realizar
20 etapas adicionais de estimativa / compensação. Este preditor é em seguida usado no lugar do preditor zero em todas as etapas anteriores para o refinamento de todos os demais vetores de animação.

Em uma modalidade, uma quarta abordagem de acordo
25 com a presente invenção para a realização de uma estimativa de animação usando bi-predição de referência combinada será descrita a seguir. A quarta modalidade representa uma abordagem mais simples que pode, no entanto, resultar em

certa perda de eficiência em comparação com os métodos acima, e envolve a consideração de todas as hipóteses como tendo uma animação idêntica (isto é, todas as referências são associadas aos mesmos vetores de animação idênticos) e no refinamento de sua predição em conjunto em comparação à imagem fonte. Isto pode ser feito como segue:

```

referência = 0;
fonte = z;
para (j = 0 < N; j++)
10  {
    mvj,0,j=mvj,1,i=(0,0);
    reference=reference + wj * xj(mvj,0,i);
  }
  //! Realiza ME usando referência para refinar
15  vetores de animação mvi,0
    //! usando distorção SAD = | fonte - referência
    (mv) |.
    Motion_Estimation(source,    reference,    mvi,0,i,
minSADi)

20  best_mv={mv,mv, ..., mv}
    Destaca-se que os algoritmos exemplares acima
    estimam os vetores de animação de combinação potencialmente
    melhores quando N referências são usadas. No entanto, um
    codificador pode escolher o número de referências para um
25  dado macrobloco, desde que possa resultar em um melhor
    desempenho, ou devido a uma melhor predição ou devido a um
    overhead reduzido. Obviamente, embora este processo possa se

```

repetir para todas as combinações de referência possíveis, no fim, no entanto, a combinação de referência que leva a um melhor desempenho poderá ser selecionada usando um determinado critério. Em particular, quando o número de referências / hipóteses possíveis é dois, então, baseado no
 5 dito acima, pode-se considerar os seguintes três casos.

Com relação ao primeiro caso relativo à quarta abordagem, de acordo com a presente invenção, considera-se cada referência (x ou y) separadamente como sendo
 10 estacionária em comparação a uma imagem fonte z, (caso 1 sem nenhuma iteração), computando a animação da outra referência usando a seguinte métrica de distorção, embora ainda considerando as ponderações apropriadas (a e b, respectivamente) a partir do processo de estimativa de
 15 ponderação, como segue:

$$_SAD = |(z - by) - ax(mv_2)|,$$

$$SAD = |(z - ax) - by(mv_3)|.$$

Durante a decisão de modo, pode-se considerar os seguintes vetores de animação / modos: predição única; modo
 20 bi-preditivo usando mv_0 e mv_1 computados usando uma predição única ME; modo direto; modo bi-preditivo usando mv_2 e (0,0); modo bi-preditivo usando (0,0) e mv_3 ; modo bi-preditivo usando mv_2 e mv_3 ; e consideração opcional de modos bi-preditivos usando mv_2 e mv_1 , ou mv_0 , e mv_3 , ou ainda
 25 combinações com vetores de animação diretos.

Com relação ao segundo caso relativo à quarta abordagem de acordo com a presente invenção, considera-se apenas uma referência como sendo estacionária em comparação

à imagem fonte, sendo computada a animação da outra referência (caso 1 limitado, sem nenhuma iteração), usando a seguinte métrica de distorção, considerando também as ponderações apropriadas a partir do processo de estimativa
 5 de ponderação, como segue:

$$SAD = |(z-ax) - by(mv)|.$$

Durante a decisão de modo, pode-se considerar os seguintes vetores de animação / modos: predição única; modo bi-preditivo usando mv_0 e mv_1 computados usando uma predição
 10 única ME; modo direto; modo bi-preditivo usando mv para ambas as referências; e consideração opcional (recomendada) de modos bi-preditivos usando mv e mv_1 , ou mv_0 , e mv , ou ainda combinações com vetores de animação diretos.

Voltando para a Figura 3, um método exemplar para
 15 a codificação de um macrobloco usando bi-predição é indicado de modo geral com o numeral de referência 300. O método 300 inclui um bloco de inicialização 305 que passa o controle para um bloco de função 350 e para um bloco de limite de loop 310.

20 O bloco de função realiza uma decisão intra-modos, registra a medida de distorção, e passa o controle para um bloco de função 330.

O bloco de limite de laço 310 começa um laço sobre todos os inter-modos, e passa o controle para um bloco de
 25 função 315. O bloco de função 315 realiza uma estimativa de animação ao considerar em conjunto todas as hipóteses, e passa o controle para um bloco de função 320. O bloco de função 315 realiza a estimativa de animação ao considerar

junto todas as hipóteses, e passa o controle para um bloco de limite de laço 325 que finaliza o laço e passa o controle para o bloco de função 330.

O bloco de função 330 define o modo de codificação para ser uma medida com a menor medida de distorção, e passa o controle para um bloco de função 335. O bloco de função 335 codifica o macrobloco em questão, e passa o controle para um bloco de finalização 399.

Voltando para a Figura 4, um método exemplar para a estimativa de animação usando uma bi-predição de referência combinada, na qual iterações paralelas são feitas no sentido de otimizar a hipótese restante, e é indicado de modo geral com o numeral de referência 400.

O método 400 inclui um bloco de inicialização 405 que passa o controle para um bloco de função 410, um bloco de função 415, e um bloco de função 420.

O bloco de função 410 inicia a realização da estimativa de animação pelo índice de referência 0, pressupondo vetores de animação para outros índices de referência como 0 ou algum valor predefinido, e passa o controle para um bloco de função 425. O bloco de função 425 refina os vetores de animação iterativamente para todas as referências, salva os vetores de animação como $\{mv_{0,0,0}, mv_{1,0,0}, mv_{m,0,0}, \dots, mv_{n-1,0,0}\}$ e a distorção como $minsad_0$, e passa o controle para um bloco de função 445.

O bloco de função 415 inicia a realização da estimativa de animação para o índice de referência m , presumindo que os vetores de animação para os outros índices

de referência são 0 ou algum valor predefinido, e passa o controle para um bloco de função 430. O bloco de função 430 refina os vetores de animação iterativamente para todas as referências, salva os vetores de animação como $\{mv_{0,0,m},$
 5 $mv_{1,0,m}, mv_{m,0,0}, \dots, mv_{n-1,0,m}\}$ e a distorção como $minsad_m$, e passa o controle para o bloco de função 445.

O bloco de função 420 inicia a realização da estimativa de animação para o índice de referência N-1, presumindo que os vetores de animação para outros índices de
 10 referência são 0 ou algum valor predefinido, e passa o controle para um bloco de função 440. O bloco de função 440 refina os vetores de animação iterativamente para todas as referências, salva os vetores de animação como $\{mv_{0,0,N-1},$
 $mv_{1,0,N-1}, mv_{m,0,N-1}, \dots, mv_{N-1,0,N-1}\}$ e a distorção como $minSAD_{N-}$
 15 $_1$, e passa o controle para o bloco de função 445.

O bloco de função 445 seleciona a melhor combinação de vetores de animação com um valor $minSAD$ mínimo, e passa o controle para um bloco de finalização 499.

Voltando para a Figura 5, um método exemplar para
 20 a estimativa de animação usando uma bi-predição de referência combinada, na qual vetores de animação são iterativamente refinados com base em uma melhor referência, é indicado de modo geral pelo numeral de referência 500.

O método 500 inclui um bloco de inicialização 505
 25 que passa o controle para um bloco de função 510, um bloco de função 515, e um bloco de função 520.

O bloco de função 510 inicia a realização da estimativa de animação para o índice de referência 0,

pressupondo vetores de animação para outros índices de referência como 0 ou algum valor predefinido, e passa o controle para um bloco de função 525.

O bloco de função 515 inicia a realização da
 5 estimativa de animação para um índice de referência m , pressupondo vetores de animação para outros índices de referência como 0 ou algum valor predefinido, e passa o controle para o bloco de função 525.

O bloco de função 520 inicia a realização da
 10 estimativa de animação para o índice de referência $N-1$, pressupondo vetores de animação para outros índices de referência como 0 ou algum valor predefinido, e passa o controle para o bloco de função 525.

O bloco de função 525 seleciona a melhor
 15 referência ($best_ref$) com um valor $minSAD$ mínimo, e passa o controle para um bloco de função 530. O bloco de função 530 refina os vetores de animação iterativamente para todas as referências, salva os vetores de animação como $\{mv_{0,0,best_ref}, mv_{1,0,best_ref}, mv_{m,0,best_ref}, \dots, mv_{N-1,0,best_ref}\}$, e passa o
 20 controle para um bloco de finalização 599.

Voltando para a Figura 6, um método exemplar para a estimativa de animação, considerando todas as hipóteses com tendo vetores de animação idênticos, é indicado de modo geral pelo numeral de referência 600.

25 O método 600 inclui um bloco de inicialização 605 que passa o controle para um bloco de função 610. O bloco de função 610 adiciona todas as imagens de referência juntas para formar uma única imagem de referência, e passa o

controle para o bloco de função 615. O bloco de função 615 realiza a estimativa de animação na imagem de referência recém-formada, e passa o controle para um bloco de função 620. O bloco de função 620 define o vetor de animação resultante para a imagem de referência recém-formada como os vetores de animação para todas as imagens de referência, e passa o controle para um bloco de finalização 699.

Voltando para a Figura 7, um método exemplar para a realização de uma estimativa de animação usando uma bi-predição de referência combinada, na qual cada referência (x ou y) é considerada separadamente como sendo estacionária em comparação a uma imagem fonte x, é indicado de modo geral pelo numeral de referência 700.

O método 700 inclui um bloco de inicialização 705 que passa o controle para um bloco de função 710. O bloco de função 710 realiza uma estimativa de animação (que gera o vetor de animação mv_0) para a referência x, e passa o controle para um bloco de função 715. O bloco de função 715 realiza a estimativa de animação (que gera o vetor de animação mv_1) para a referência y, e passa o controle para um bloco de função 720. O bloco de função 720 realiza a estimativa de animação (que gera o vetor de animação mv_2) usando $SAD = |(z-by) - ax(mv_2)|$, e passa o controle para um bloco de função 725. O bloco de função 725 realiza a estimativa de animação (mv_2) usando $SAD = |(z-ax) - by(mv_3)|$, e passa o controle para o bloco de função 730. O bloco de função 730 realiza a decisão de modo, e passa o controle para um bloco de função 735. Com relação à decisão de modo

realizada pelo bloco de função 730, tal decisão de modo pode ser feita considerando, por exemplo: uma predição única; um modo bi-preditivo usando mv_0 e mv_1 ; um modo direto; um modo bi-preditivo usando mv_2 e $(0,0)$; um modo bi-preditivo usando $(0,0)$ e mv_3 ; um modo bi-preditivo usando mv_2 e mv_3 ; e a consideração opcional de modos bi-preditivos usando mv_2 e mv_1 , ou mv_0 , e mv_3 , ou ainda combinações com vetores de animação diretos. As considerações precedentes são exemplares e, por conseguinte, dados os ensinamentos da presente invenção aqui providos, estas e outras considerações podem também ser utilizadas com relação à decisão de modo feita pelo bloco de função 730, ao mesmo tempo mantendo o âmbito de aplicação da presente invenção.

O bloco de função 735 seleciona o modo com uma medida de distorção mínima, e passa o controle para um bloco de finalização 799.

Voltando para a Figura 8, um método exemplar para a realização de uma estimativa de animação usando uma bi-predição de referência combinada, na qual apenas uma referência é considerada como sendo estacionária em comparação a uma imagem fonte, é indicado de modo geral pelo numeral de referência 800. O método 800 inclui um bloco de inicialização 805 que passa o controle para um bloco de função 810. O bloco de função 810 realiza uma estimativa de animação (que gera o vetor de animação mv_0) para a referência x, e passa o controle para um bloco de função 815. O bloco de função 815 realiza a estimativa de animação (que gera o vetor de animação mv_1) para a referência y, e

passa o controle para um bloco de função 820. O bloco de função 820 realiza a estimativa de animação (que gera o vetor de animação mv_2) usando $SAD = |(z-by) - ax(mv_2)|$, e passa o controle para um bloco de função 825. O bloco de função 825 realiza a decisão de modo, e passa o controle para o bloco de função 830. Com relação à decisão de modo realizada pelo bloco de função 830, tal decisão de modo pode ser feita considerando, por exemplo: uma predição única; um modo bi-preditivo usando mv_0 e mv_1 ; um modo direto; um modo bi-preditivo usando mv_2 e $(0,0)$; e uma consideração opcional de modos bi-preditivos usando mv_2 e mv_1 , ou ainda combinações com vetores de animação diretos. As considerações precedentes são exemplares e, por conseguinte, dados os ensinamentos da presente invenção aqui providos, estas e outras considerações podem também ser utilizadas com relação à decisão de modo feita pelo bloco de função 825, ao mesmo tempo mantendo o âmbito de aplicação da presente invenção.

O bloco de função 830 seleciona o modo com uma medida de distorção mínima, e passa o controle para um bloco de finalização 899.

Voltando para a Figura 9, um método exemplar para a realização de uma estimativa de animação usando uma bi-predição de referência combinada, na qual ambas as referências são combinadas, pressupondo uma animação zero entre as referências, é indicado de modo geral pelo numeral de referência 900. O método 900 inclui um bloco de inicialização 905 que passa o controle para um bloco de função 910. O bloco de função 910 realiza uma estimativa de

animação (que gera o vetor de animação mv_0) para a referência x , e passa o controle para um bloco de função 925. O bloco de função 925 realiza a estimativa de animação (que gera o vetor de animação mv_1) para a referência y , e

5 passa o controle para um bloco de função 920. O bloco de função 920 realiza a estimativa de animação (que gera o vetor de animação mv) usando $SAD = |(z-by) - ax(mv)|$, e passa o controle para um bloco de função 925. O bloco de função 925 realiza a decisão de modo, e passa o controle para o bloco

10 de função 930. Com relação à decisão de modo realizada pelo bloco de função 930, tal decisão de modo pode ser feita considerando, por exemplo: uma predição única; um modo bi-preditivo usando mv_0 e mv_1 ; um modo direto; um modo bi-preditivo usando mv para ambas as referências; e a

15 consideração opcional de modos bi-preditivos usando mv e mv_1 , ou mv_0 e mv , ou ainda combinações com vetores de animação diretos. As considerações precedentes são exemplares e, por conseguinte, dados os ensinamentos da presente invenção aqui providos, estas e outras

20 considerações podem também ser utilizadas com relação à decisão de modo feita pelo bloco de função 925, ao mesmo tempo mantendo o âmbito de aplicação da presente invenção.

O bloco de função 930 seleciona o modo com uma medida de distorção mínima, e passa o controle para um bloco

25 de finalização 999.

Será feita a seguir uma descrição de algumas vantagens / características auxiliares da presente invenção, sendo que algumas destas já forma acima mencionadas. Por

exemplo, uma vantagem / característica é um aparelho que inclui um codificador para codificar uma imagem de multipredição a partir de uma combinação de duas ou mais imagens de referência ao se predizer respectivamente um vetor de animação para uma imagem particular dentre duas ou mais imagens de referência em um processo de estimativa de animação enquanto se inicializa os vetores de animação de modo a manter algumas dentre duas ou mais imagens de referência em um valor predefinido para uso pelo processo de estimativa de animação.

Uma outra vetor de animação / característica é o aparelho tendo o codificador conforme descrito acima, cujo codificador prediz o vetor de animação para uma referência particular dentre duas ou mais referências ao considerar as demais imagens dentre duas ou mais imagens de referência como estacionárias.

Ainda, uma outra vantagem / característica é o aparelho tendo o codificador conforme descrito acima, cujo codificador prediz os vetores de animação para cada qual dentre duas ou mais referências nas respectivas porções paralelas de um primeiro dentre múltiplos processos iterativos, e refina os vetores de animação nos subseqüentes processos dentre os múltiplos processos iterativos. Ainda, uma outra vantagem / característica é o aparelho tendo o codificador que prediz os vetores de animação para cada qual dentre duas ou mais referências nas respectivas porções paralelas de um primeiro processo dentre múltiplos processos iterativos conforme acima descrito, cujo codificador

seleciona uma melhor combinação de vetor de animação resultante dos múltiplos processos iterativos com base em uma medida de distorção.

Além disso, uma outra vantagem / característica é o aparelho tendo o codificador conforme descrito acima, no qual pelo menos um vetor dentre os vetores de animação para pelo menos uma imagem dentre duas ou mais imagens de referência é inicializado no valor predefinido com base nos vetores de animação previamente computados para a pelo menos uma dentre as restantes duas ou mais imagens de referência. Ainda, uma outra vantagem / característica é o aparelho tendo o codificador, no qual o pelo menos um vetor dentre os vetores de animação para o a pelo menos uma das imagens restantes dentre duas ou mais imagens de referência é inicializado com base nos vetores de animação previamente computados para pelo menos uma imagem restante dentre duas ou mais imagens de referência conforme descrito acima, em que os vetores de animação previamente computados compreendem pelo menos um vetor de animação direto e um vetor de animação previamente computado a partir da combinação de um número menor de imagens de referência do que o das duas ou mais imagens de referência, em que o pelo menos um vetor de animação direto inclui pelo menos um dentre pelo menos um vetor de animação direto temporal e pelo menos um vetor de animação direto espacial. Ainda, uma outra vantagem / característica é o aparelho tendo o codificador, no qual o pelo menos um dos vetores de animação para a pelo menos uma das imagens restantes dentre duas ou

mais imagens de referência é inicializado com base nos vetores de animação previamente computados para a pelo menos uma das imagens restantes dentre duas ou mais imagens de referência conforme descrito acima, cujo dito codificador
5 utiliza um processo iterativo a fim de predizer os vetores de animação para as duas ou mais imagens de referência, e sintetiza uma decisão relativa aos vetores de animação a serem inicialmente usados no processo iterativo, a partir de pelo menos um vetor dentre os vetores de animação
10 inicializados no valor predefinido e os vetores de animação previamente computados, com base em uma métrica de correlação.

Adicionalmente, uma outra vantagem / característica é o aparelho tendo o codificador conforme
15 descrito acima, no qual o codificador prediz os vetores de animação para as duas ou mais imagens de referência usando o processo iterativo conforme acima descrito, em que o codificador limita um número máximo de iterações realizadas pelo processo iterativo. Além disso, uma outra vantagem /
20 característica é o aparelho tendo o codificador que prediz os vetores de animação para as duas ou mais imagens de referência usando o processo iterativo conforme acima descrito, cujo codificador limita um número máximo dentre duas ou mais imagens de referência a serem testadas em
25 paralelo pelo processo iterativo.

Ainda, uma outra vantagem / característica é o aparelho tendo o codificador conforme acima descrito, cujo codificador estima os vetores de animação para as duas ou

mais imagens de referência com base nas ponderações usadas para as duas ou mais imagens de referência de um processo de estimativa de ponderação anterior.

Adicionalmente, uma outra vantagem /
5 característica é o aparelho tendo o codificador conforme acima descrito, cujo codificador estima os vetores de animação para as duas ou mais imagens de referência usando múltiplos processos iterativos em paralelo, os múltiplos processos iterativos tendo diferentes pontos de
10 inicialização, correspondendo, respectivamente, a diferentes imagens dentre duas ou mais imagens de referência. Além disso, uma outra vantagem / característica é o aparelho tendo o codificador que realiza os múltiplos processos iterativos conforme acima descrito, em que o codificador
15 realiza um primeiro dentre múltiplos processos iterativos a fim de prover valores iniciais para os vetores de animação para cada qual dentre duas ou mais imagens de referência, seleciona a melhor referência inicial dentre duas ou mais referências com base na menor medida de distorção, e realiza
20 os subseqüentes processos dentre os múltiplos processos iterativos de modo a refinar os valores iniciais obtidos para a melhor referência inicial dentre duas ou mais referências.

Ainda, uma outra vantagem / característica é o
25 aparelho tendo o codificador conforme acima descrito, cujo codificador prediz informações de animação para a imagem de multi-predição com base nos vetores de animação para duas ou mais imagens de referência idênticas e refina, em conjunto,

os vetores de animação comparados à imagem de multi-predição.

Ainda, uma outra vantagem / característica é o aparelho tendo o codificador conforme acima descrito, cujo
5 codificador estima informações de vetor de animação para a imagem de multi-predição ao combinar as duas ou mais imagens de referência em uma imagem de referência combinada, e ao usar animação zero entre as duas ou mais imagens de referência.

10 Adicionalmente, uma outra vantagem / característica é o aparelho tendo o codificador conforme acima descrito, cujo codificador prediz um respectivo vetor de animação para cada qual dentre duas ou mais imagens de referência ao considerar separadamente as demais imagens
15 dentre duas ou mais imagens de referência como sendo estacionárias em comparação à imagem de multi-predição. Ademais, uma outra vantagem / característica é o aparelho tendo o codificador que prediz o respectivo vetor de animação para cada qual dentre duas ou mais imagens de
20 referência ao considerar, separadamente, as demais imagens dentre duas ou mais imagens de referência como sendo estacionárias em comparação à imagem de multi-predição conforme descrita acima, em que o codificador prediz o respectivo vetor de animação para cada qual dentre duas ou
25 mais imagens de referência por meio da seleção de uma respectiva ponderação para as mesmas.

Ainda, uma outra vantagem / característica é o aparelho tendo o codificador conforme acima descrito, cujo

codificador realiza uma decisão de modo para a imagem de multi-predição a fim de selecionar um modo com uma medida de distorção mínima dentre um modo preditivo único, um modo preditivo direto, um modo bi-preditivo usando os vetores de animação para as duas ou mais imagens de referência respectivamente calculadas usando uma predição única, pelo menos um modo bi-preditivo usando um vetor de animação refinado para uma imagem dentre duas ou mais imagens de referência e um vetor de animação estacionário, e um modo bi-preditivo usando um respectivo vetor de animação refinado para cada qual dentre duas ou mais imagens de referência, em que os vetores de animação para cada qual dentre duas ou mais imagens de referência são respectivamente calculados com base nos vetores de animação para os restantes dentre duas ou mais imagens de referência separadamente consideradas como estacionárias. Ainda, uma outra vantagem / característica é o aparelho tendo o codificador que realiza a decisão de modo conforme acima descrito, cujo codificador seleciona ainda o modo com a medida de distorção mínima dentre pelo menos um modo bi-preditivo usando o vetor de animação para uma imagem dentre duas ou mais imagens de referência, e o vetor de animação refinado para uma imagem dentre duas ou mais imagens de referência, e pelo menos uma combinação incluindo pelo menos um vetor de animação direto.

Ainda, uma outra vantagem / característica é o aparelho tendo o codificador conforme acima descrito, cujo codificador prediz um respectivo vetor de animação para cada qual dentre duas ou mais imagens de referência ao considerar

apenas uma imagem dentre duas ou mais imagens de referência como sendo estacionárias em comparação à imagem de multi-predição. Adicionalmente, uma outra vantagem / característica é o aparelho tendo o codificador que prediz o respectivo vetor de animação para cada qual dentre duas ou mais imagens de referência ao considerar apenas uma imagem dentre duas ou mais imagens de referência como sendo estacionárias em comparação à imagem de multi-predição conforme acima descrito, em que o codificador prediz o respectivo vetor de animação para cada qual dentre duas ou mais imagens de referência ao selecionar uma respectiva ponderação para as mesmas. Ainda, uma outra vantagem / característica é o aparelho tendo o codificador que prediz o respectivo vetor de animação para cada qual dentre duas ou mais imagens de referência ao considerar apenas uma imagem dentre duas ou mais imagens de referência como sendo estacionárias em comparação à imagem de multi-predição conforme descrita acima, em que apenas uma imagem dentre duas ou mais imagens de referência é selecionada com base em pelo menos uma dentre uma lista de referência particular à qual a apenas uma imagem dentre duas ou mais imagens de referência corresponde, em termos de proximidade temporal, e de maior similaridade, à imagem de multi-predição.

Ainda, uma outra vantagem / característica é o aparelho tendo o codificador conforme acima descrito, cujo codificador realiza uma decisão de modo para a imagem de multi-predição a fim de selecionar um modo com uma medida de distorção mínima dentre um modo preditivo único, um modo

preditivo direto, um modo bi-preditivo usando os vetores de animação para as duas ou mais imagens de referência respectivamente calculadas usando uma predição única, um modo bi-preditivo usando um vetor de animação estacionário e um vetor de animação refinado, no qual um respectivo vetor dentre os vetores de animação para apenas uma imagem dentre duas ou mais imagens de referência é calculado com base em um vetor de animação estacionário que é usado para as demais imagens dentre duas ou mais imagens de referência. Ainda, uma outra vantagem / característica é o aparelho tendo o codificador que realiza a decisão de modo conforme acima descrita, em que o codificador seleciona ainda o modo com a medida de distorção mínima dentre pelo menos um modo bi-preditivo usando um vetor dentre os vetores de animação de uma imagem dentre duas ou mais imagens de referência e o vetor de animação refinado, e pelo menos uma combinação de um vetor de animação de direção e o vetor de animação calculado de acordo com a medida de distorção.

Ainda, uma outra vantagem / característica é o aparelho tendo o codificador conforme acima descrito, cujo codificador computa, em conjunto, os vetores de animação para as duas ou mais imagens de referência com base nas duas ou mais imagens de referência tendo animação idêntica com relação à imagem de multi-predição a ser codificada. Além disso, uma outra vantagem / característica é o aparelho tendo o codificador que, em conjunto, computa os vetores de animação para as duas ou mais imagens de referência conforme acima descrito, em que o codificador computa, em conjunto,

os vetores de animação para as duas ou mais imagens de referência por meio da seleção de uma respectiva ponderação para as mesmas.

Ainda, uma outra vantagem / característica é o
5 aparelho tendo o codificador conforme acima descrito, cujo
codificador realiza uma decisão de modo para a imagem de
multi-predição a fim de selecionar um modo com uma medida de
distorção mínima dentre um modo preditivo único, um modo
preditivo direto, um modo bi-preditivo usando os vetores de
10 animação para as duas ou mais imagens de referência
respectivamente calculadas usando uma predição única, um
modo bi-preditivo usando um respectivo vetor de animação
refinado para cada qual dentre duas ou mais imagens de
referência, e um modo de predição usando um vetor de
15 animação para uma combinação das duas ou mais imagens de
referência calculadas ao se considerar os vetores de
animação para as duas ou mais imagens de referência se
tornarem idênticas. Adicionalmente, uma outra vantagem /
característica é o aparelho tendo o codificador que realiza
20 a decisão de modo conforme acima descrito, cujo codificador
seleciona ainda o modo com a medida de distorção mínima
dentre pelo menos um modo bi-preditivo usando um vetor
dentre os vetores de animação de uma imagem dentre duas ou
mais imagens de referência, e o vetor de animação calculado
25 ao considerar os vetores de animação para as duas ou mais
imagens de referência se tornarem idênticas, e pelo menos
uma combinação de um vetor de animação direto e o vetor de

animação calculado ao considerar os vetores de animação para as duas ou mais imagens de referência se tornarem idênticas.

Estas e outras características e vantagens da presente invenção podem ser prontamente obtidas por uma
5 pessoa com habilidade simples na técnica com base nos presentes ensinamentos. Deve-se entender que os ensinamentos da presente invenção podem ser implementados em várias formas de hardware, software, firmware, processadores de uso especial, ou suas combinações.

10 Mais preferivelmente, os ensinamentos da presente invenção são implementados como uma combinação de um hardware e um software. Além disso, o software pode ser implementado como um programa de aplicação tangivelmente incorporado em uma unidade de armazenamento de programa. O
15 programa de aplicação pode ser transferido para, e executado por, uma máquina que compreenda qualquer arquitetura adequada. De preferência, a máquina é implementada em uma plataforma de computador tendo um hardware, como, por exemplo, uma ou mais unidades de processamento central
20 ("CPU"), uma memória de acesso aleatório ("RAM"), e interfaces de entrada / saída ("I/O"). A plataforma de computador pode também incluir um sistema operacional e um código de micro-instrução. Os vários processos e funções descritos no presente documento podem fazer parte de um
25 código de micro-instrução ou de um programa de aplicação, ou qualquer combinação dos mesmos, que possam ser executados por uma unidade CPU. Além disso, várias outras unidades periféricas podem ser conectadas à plataforma de computador,

como, por exemplo, uma unidade de armazenamento de dados adicional ou uma unidade de impressão.

Deve-se entender que, uma vez que alguns dos componentes e métodos de sistema constituintes ilustrados nos desenhos em anexo são de preferência implementados em um software, as conexões em questão entre os componentes de sistema e os blocos de função de processo podem diferir dependendo da maneira na qual a presente invenção é programada. Dados os ensinamentos da presente invenção, uma pessoa com habilidade simples na técnica em questão será capaz de contemplar estas implementações ou implementações e configurações similares às da presente invenção.

Embora as modalidades ilustrativas tenham sido descritas no presente documento com referência aos desenhos em anexo, deve-se entender que a presente invenção não se limita a estas precisas modalidades, e que várias mudanças e modificações podem ser feitas por uma pessoa com habilidade simples na técnica pertinente sem se afastar do âmbito ou espírito da presente invenção. Todas estas mudanças e modificações pretendem estar incluídas dentro do âmbito da presente invenção conforme apresentada nas reivindicações em apenso.

REIVINDICAÇÕES

1. Aparelho, **CARACTERIZADO** pelo fato de compreender:

- um codificador (200) para codificar uma imagem de multi-predição a partir de uma combinação de duas ou mais imagens de referência ao se predizer respectivamente um vetor de animação para uma imagem particular dentre duas ou mais imagens de referência em um processo de estimativa de animação enquanto se inicializa os vetores de animação de modo a manter algumas imagens dentre duas ou mais imagens de referência em um valor predefinido para uso pelo processo de estimativa de animação.

2. Aparelho, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o dito codificador (200) prediz o vetor de animação para uma referência particular dentre duas ou mais referências ao considerar as demais imagens dentre duas ou mais imagens de referência como estacionárias.

3. Aparelho, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o dito codificador (200) prediz os vetores de animação para cada referência dentre duas ou mais referências nas respectivas porções paralelas de um primeiro processo dentre múltiplos processos iterativos, e refina os vetores de animação nos processos subseqüentes dentre múltiplos processos iterativos.

4. Aparelho, de acordo com a reivindicação 3, **CARACTERIZADO** pelo fato de que o dito codificador (200) seleciona uma melhor combinação de vetor de animação

resultante dos múltiplos processos iterativos com base em uma medida de distorção.

5. Aparelho, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que pelo menos um vetor dentre os vetores de animação para pelo menos uma imagem dentre duas ou mais imagens de referência é inicializado no valor predefinido com base nos vetores de animação previamente computados para pelo menos uma imagem dentre as demais imagens de duas ou mais imagens de referência.

10 6. Aparelho, de acordo com a reivindicação 5, **CARACTERIZADO** pelo fato de que os vetores de animação previamente computados compreendem pelo menos um vetor de animação direto e um vetor de animação previamente computado a partir da combinação de um número menor de imagens de
15 referência do que o das duas ou mais imagens de referência, em que o pelo menos um vetor de animação direto inclui pelo menos um vetor dentre pelo menos um vetor de animação direto temporal e pelo menos um vetor de animação direto espacial.

7. Aparelho, de acordo com a reivindicação 5, **CARACTERIZADO** pelo fato de que o dito codificador (200)
20 utiliza um processo iterativo a fim de prever os vetores de animação para as duas ou mais imagens de referência, e sintetiza uma decisão relativa aos vetores de animação a serem inicialmente usados no processo iterativo, a partir de
25 pelo menos um vetor dentre os vetores de animação inicializados no valor predefinido e os vetores de animação previamente computados, com base em uma métrica de correlação.

8. Aparelho, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o dito codificador (200) prediz os vetores de animação para as duas ou mais imagens de referência usando um processo iterativo.

5 9. Aparelho, de acordo com a reivindicação 8, **CARACTERIZADO** pelo fato de que o dito codificador (200) limita um número máximo de iterações realizadas pelo processo iterativo.

10 10. Aparelho, de acordo com a reivindicação 8, **CARACTERIZADO** pelo fato de que o dito codificador (200) limita um número máximo dentre duas ou mais imagens de referência a serem testadas em paralelo pelo processo iterativo.

15 11. Aparelho, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o dito codificador (200) estima os vetores de animação para as duas ou mais imagens de referência com base nas ponderações usadas para as duas ou mais imagens de referência de um processo de estimativa de ponderação anterior.

20 12. Aparelho, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o dito codificador (200) estima os vetores de animação para as duas ou mais imagens de referência usando múltiplos processos iterativos em paralelo, os múltiplos processos iterativos tendo diferentes
25 pontos de inicialização, correspondendo, respectivamente, a diferentes imagens dentre duas ou mais imagens de referência.

13. Aparelho, de acordo com a reivindicação 12, **CARACTERIZADO** pelo fato de que o dito codificador (200) realiza um primeiro processo dentre múltiplos processos iterativos de modo a prover valores iniciais para os vetores de animação para cada qual dentre duas ou mais imagens de referência, seleciona uma melhor referência inicial dentre duas ou mais referências com base em uma medida de distorção mais baixa, e realiza os processos subsequentes dentre múltiplos processos iterativos de modo a refinar os valores iniciais obtidos para a melhor referência inicial dentre duas ou mais referências.

14. Aparelho, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o dito codificador (200) prediz informações de animação para a imagem de multi-predição com base nos vetores de animação para as duas ou mais imagens de referência idênticas e refina, em conjunto, os vetores de animação comparados à imagem de multi-predição.

15. Aparelho, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o dito codificador (200) estima informações de vetor de animação para a imagem de multi-predição ao combinar as duas ou mais imagens de referência em uma imagem de referência combinada, e ao usar animação zero entre as duas ou mais imagens de referência.

16. Aparelho, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o dito codificador (200) prediz um respectivo vetor de animação para cada qual dentre duas ou mais imagens de referência ao considerar

separadamente as demais imagens dentre duas ou mais imagens de referência como sendo estacionárias em comparação à imagem de multi-predição.

17. Aparelho, de acordo com a reivindicação 16,
5 **CARACTERIZADO** pelo fato de que o dito codificador (200) prediz o respectivo vetor de animação para cada qual dentre duas ou mais imagens de referência ao selecionar uma respectiva ponderação para as mesmas.

18. Aparelho, de acordo com a reivindicação 1,
10 **CARACTERIZADO** pelo fato de que o dito codificador (200) realiza uma decisão de modo para a imagem de multi-predição a fim de selecionar um modo com uma medida de distorção mínima dentre um modo preditivo único, um modo preditivo direto, um modo bi-preditivo usando os vetores de animação
15 para as duas ou mais imagens de referência respectivamente calculadas usando uma predição única, pelo menos um modo bi-preditivo usando um vetor de animação refinado para uma imagem dentre duas ou mais imagens de referência e um vetor de animação estacionário, e um modo bi-preditivo usando um
20 respectivo vetor de animação refinado para cada qual dentre duas ou mais imagens de referência, em que os vetores de animação para cada qual dentre duas ou mais imagens de referência são respectivamente calculados com base nos vetores de animação para os restantes dentre duas ou mais
25 imagens de referência separadamente consideradas como estacionárias.

19. Aparelho, de acordo com a reivindicação 18,
CARACTERIZADO pelo fato de que o dito codificador (200)

seleciona o modo com a medida de distorção mínima dentre pelo menos um modo bi-preditivo usando o vetor de animação para uma imagem dentre duas ou mais imagens de referência, e o vetor de animação refinado para uma imagem dentre duas ou mais imagens de referência, e pelo menos uma combinação incluindo pelo menos um vetor de animação direto.

20. Aparelho, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o dito codificador (200) prediz um respectivo vetor de animação para cada qual dentre duas ou mais imagens de referência ao considerar apenas uma imagem dentre duas ou mais imagens de referência como sendo estacionárias em comparação à imagem de multi-predição.

21. Aparelho, de acordo com a reivindicação 20, **CARACTERIZADO** pelo fato de que o dito codificador (200) prediz o respectivo vetor de animação para cada qual dentre duas ou mais imagens de referência ao selecionar uma respectiva ponderação para as mesmas.

22. Aparelho, de acordo com a reivindicação 20, **CARACTERIZADO** pelo fato de que a apenas uma imagem dentre duas ou mais imagens de referência é selecionada com base em pelo menos uma dentre uma lista de referência particular à qual a apenas uma imagem dentre duas ou mais imagens de referência corresponde, em termos de proximidade temporal, e maior similaridade, à imagem de multi-predição.

23. Aparelho, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o dito codificador (200) realiza uma decisão de modo para a imagem de multi-predição a fim de selecionar um modo com uma medida de distorção

mínima dentre um modo preditivo único, um modo preditivo direto, um modo bi-preditivo usando os vetores de animação para as duas ou mais imagens de referência respectivamente calculadas usando uma predição única, um modo bi-preditivo
5 usando um vetor de animação estacionário e um vetor de animação refinado, no qual um respectivo vetor dentre os vetores de animação para apenas uma imagem dentre duas ou mais imagens de referência é calculado com base em um vetor de animação estacionário que é usado para as demais imagens
10 dentre duas ou mais imagens de referência.

24. Aparelho, de acordo com a reivindicação 23, **CARACTERIZADO** pelo fato de que o dito codificador (200) seleciona ainda o modo com a medida de distorção mínima dentre pelo menos um modo bi-preditivo usando um vetor
15 dentre os vetores de animação de uma imagem dentre duas ou mais imagens de referência e o vetor de animação refinado, e pelo menos uma combinação de um vetor de animação de direção e o vetor de animação calculado de acordo com a medida de distorção.

20 25. Aparelho, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o dito codificador (200) computa, em conjunto, os vetores de animação para as duas ou mais imagens de referência com base nas duas ou mais imagens de referência tendo animação idêntica com relação à imagem
25 de multi-predição a ser codificada.

26. Aparelho, de acordo com a reivindicação 25, **CARACTERIZADO** pelo fato de que o dito codificador (200) computa, em conjunto, os vetores de animação para as duas ou

mais imagens de referência por meio da seleção de uma respectiva ponderação para as mesmas.

27. Aparelho, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o dito codificador (200) realiza uma decisão de modo para a imagem de multi-predição a fim de selecionar um modo com uma medida de distorção mínima dentre um modo preditivo único, um modo preditivo direto, um modo bi-preditivo usando os vetores de animação para as duas ou mais imagens de referência respectivamente calculadas usando uma predição única, um modo bi-preditivo usando um respectivo vetor de animação refinado para cada qual dentre duas ou mais imagens de referência, e um modo de predição usando um vetor de animação para uma combinação das duas ou mais imagens de referência calculadas ao se considerar os vetores de animação para as duas ou mais imagens de referência se tornarem idênticas.

28. Aparelho, de acordo com a reivindicação 27, **CARACTERIZADO** pelo fato de que o dito codificador (200) seleciona ainda o modo com a medida de distorção mínima dentre pelo menos um modo bi-preditivo usando um vetor dentre os vetores de animação de uma imagem dentre duas ou mais imagens de referência, e o vetor de animação calculado ao considerar os vetores de animação para as duas ou mais imagens de referência se tornarem idênticas, e pelo menos uma combinação de um vetor de animação direto e o vetor de animação calculado ao considerar os vetores de animação para as duas ou mais imagens de referência se tornarem idênticas.

29. Método, **CARACTERIZADO** pelo fato de compreender as etapas de:

- codificar uma imagem de multi-predição a partir de uma combinação de duas ou mais imagens de referência ao se predizer respectivamente um vetor de animação para uma
5 imagem particular dentre duas ou mais imagens de referência em um processo de estimativa de animação enquanto se inicializa os vetores de animação de modo a manter algumas dentre duas ou mais imagens de referência em um valor
10 predefinido para uso pelo processo de estimativa de animação (400).

30. Método, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que a dita etapa de codificar prediz o vetor de animação para uma referência particular
15 dentre duas ou mais referências ao considerar as demais imagens dentre duas ou mais imagens de referência como estacionárias (410, 415, 420).

31. Método, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que a dita etapa de codificar
20 prediz os vetores de animação para cada qual dentre duas ou mais referências nas respectivas porções paralelas de um primeiro processo dentre múltiplos processos iterativos, e refina os vetores de animação processos subsequentes dentre os múltiplos processos iterativos (410, 425, 415, 430, 420,
25 440).

32. Método, de acordo com a reivindicação 31, **CARACTERIZADO** pelo fato de que a dita etapa de codificar seleciona uma melhor combinação de vetor de animação

resultante dos múltiplos processos iterativos com base em uma medida de distorção (445).

33. Método, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que pelo menos um vetor dentre os vetores de animação para pelo menos uma imagem dentre duas ou mais imagens de referência é inicializado no valor predefinido com base nos vetores de animação previamente computados para pelo menos uma imagem dentre as demais imagens de referência.

34. Método, de acordo com a reivindicação 33, **CARACTERIZADO** pelo fato de que os vetores de animação previamente computados compreendem pelo menos um vetor de animação direto e um vetor de animação previamente computado a partir da combinação de um número menor de imagens de referência do que o das duas ou mais imagens de referência, em que o pelo menos um vetor de animação direto inclui pelo menos um dentre pelo menos um vetor de animação direto temporal e pelo menos um vetor de animação direto espacial.

35. Método, de acordo com a reivindicação 33, **CARACTERIZADO** pelo fato de que a dita etapa de codificar utiliza um processo iterativo a fim de predizer os vetores de animação para as duas ou mais imagens de referência, e sintetiza uma decisão relativa aos vetores de animação a serem inicialmente usados no processo iterativo, a partir de pelo menos um vetor dentre os vetores de animação inicializados no valor predefinido e os vetores de animação previamente computados, com base em uma métrica de correlação.

36. Método, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que a dita etapa de codificar prediz os vetores de animação para as duas ou mais imagens de referência usando um processo iterativo (425, 430, 440).

5 37. Método, de acordo com a reivindicação 36, **CARACTERIZADO** pelo fato de que a dita etapa de codificar limita um número máximo de iterações realizadas pelo processo iterativo.

10 38. Método, de acordo com a reivindicação 36, **CARACTERIZADO** pelo fato de que a dita etapa de codificar limita um número máximo dentre duas ou mais imagens de referência a serem testadas em paralelo pelo processo iterativo.

15 39. Método, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que a dita etapa de codificar estima os vetores de animação para as duas ou mais imagens de referência com base nas ponderações usadas para as duas ou mais imagens de referência de um processo de estimativa de ponderação anterior.

20 40. Método, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que a dita etapa de codificar estima os vetores de animação para as duas ou mais imagens de referência usando múltiplos processos iterativos em paralelo, os múltiplos processos iterativos tendo diferentes
25 pontos de inicialização, correspondendo, respectivamente, a diferentes imagens dentre duas ou mais imagens de referência (510, 515, 520).

41. Método, de acordo com a reivindicação 40, **CARACTERIZADO** pelo fato de que a dita etapa de codificar realiza um primeiro dentre os múltiplos processos iterativos de modo a prover valores iniciais para os vetores de animação para cada qual dentre duas ou mais imagens de referência, seleciona a melhor referência inicial dentre duas ou mais referências com base em uma medida de distorção mais baixa, e realiza os processos subsequentes dentre múltiplos processos iterativos de modo a refinar os valores iniciais obtidos para a melhor referência inicial dentre duas ou mais referências (500).

42. Método, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que a dita etapa de codificar prediz informações de animação para a imagem de multi-predição com base nos vetores de animação para as duas ou mais imagens de referência idênticas e refina, em conjunto, os vetores de animação comparados à imagem de multi-predição (600).

43. Método, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que a dita etapa de codificar estima informações de vetor de animação para a imagem de multi-predição ao combinar as duas ou mais imagens de referência em uma imagem de referência combinada, e ao usar animação zero entre as duas ou mais imagens de referência (610).

44. Método, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que a dita etapa de codificar prediz um respectivo vetor de animação para cada qual dentre

duas ou mais imagens de referência ao considerar separadamente as demais imagens dentre duas ou mais imagens de referência como sendo estacionárias em comparação à imagem de multi-predição (720, 725).

5 45. Método, de acordo com a reivindicação 44, **CARACTERIZADO** pelo fato de que a dita etapa de codificar prediz o respectivo vetor de animação para cada qual dentre duas ou mais imagens de referência ao selecionar uma respectiva ponderação para as mesmas.

10 46. Método, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que a dita etapa de codificar realiza uma decisão de modo para a imagem de multi-predição a fim de selecionar um modo com uma medida de distorção mínima dentre um modo preditivo único, um modo preditivo
15 direto, um modo bi-preditivo usando os vetores de animação para as duas ou mais imagens de referência respectivamente calculadas usando uma predição única, pelo menos um modo bi-preditivo usando um vetor de animação refinado para uma
20 imagem dentre duas ou mais imagens de referência e um vetor de animação estacionário, e um modo bi-preditivo usando um respectivo vetor de animação refinado para uma imagem dentre
25 duas ou mais imagens de referência e um vetor de animação estacionário, e um modo bi-preditivo usando um respectivo vetor de animação refinado para cada qual dentre duas ou
mais imagens de referência, em que os vetores de animação para cada qual dentre duas ou mais imagens de referência são respectivamente calculados com base nos vetores de animação para as demais imagens dentre duas ou mais imagens de

referência separadamente consideradas como estacionárias (730).

47. Método, de acordo com a reivindicação 46, **CARACTERIZADO** pelo fato de que a dita etapa de codificar
5 seleciona o modo com a medida de distorção mínima dentre pelo menos um modo bi-preditivo usando o vetor de animação para uma imagem dentre duas ou mais imagens de referência, e o vetor de animação refinado para uma imagem dentre duas ou mais imagens de referência, e pelo menos uma combinação
10 incluindo pelo menos um vetor de animação direto.

48. Método, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que a dita etapa de codificar prediz um respectivo vetor de animação para cada qual dentre duas ou mais imagens de referência ao considerar apenas uma
15 imagem dentre duas ou mais imagens de referência como sendo estacionárias em comparação à imagem de multi-predição (820).

49. Método, de acordo com a reivindicação 48, **CARACTERIZADO** pelo fato de que a dita etapa de codificar
20 prediz o respectivo vetor de animação para cada qual dentre duas ou mais imagens de referência ao selecionar uma respectiva ponderação para as mesmas.

50. Método, de acordo com a reivindicação 48, **CARACTERIZADO** pelo fato de que a apenas uma imagem dentre
25 duas ou mais imagens de referência é selecionada com base em pelo menos uma dentre uma lista de referência particular à qual a apenas uma imagem dentre duas ou mais imagens de

referência corresponde, em termos de proximidade temporal, e maior similaridade, à imagem de multi-predição.

51. Método, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que a dita etapa de codificar
5 realiza uma decisão de modo para a imagem de multi-predição a fim de selecionar um modo com uma medida de distorção mínima dentre um modo preditivo único, um modo preditivo direto, um modo bi-preditivo usando os vetores de animação para as duas ou mais imagens de referência respectivamente
10 calculadas usando uma predição única, um modo bi-preditivo usando um vetor de animação estacionário e um vetor de animação refinado, no qual um respectivo vetor dentre os vetores de animação para apenas uma imagem dentre duas ou mais imagens de referência é calculado com base em um vetor
15 de animação estacionário que é usado para as demais imagens dentre duas ou mais imagens de referência (825).

52. Método, de acordo com a reivindicação 51, **CARACTERIZADO** pelo fato de que a dita etapa de codificar seleciona ainda o modo com a medida de distorção mínima
20 dentre pelo menos um modo bi-preditivo usando um vetor dentre os vetores de animação de uma imagem dentre duas ou mais imagens de referência e o vetor de animação refinado, e pelo menos uma combinação de um vetor de animação de direção e o vetor de animação calculado de acordo com a medida de
25 distorção.

53. Método, de acordo com a reivindicação 29, **CARACTERIZADO** pelo fato de que a dita etapa de codificar computa, em conjunto, os vetores de animação para as duas ou

mais imagens de referência com base nas duas ou mais imagens de referência tendo uma animação idêntica com relação à imagem de multi-predição a ser codificada (920).

54. Método, de acordo com a reivindicação 53,
5 **CARACTERIZADO** pelo fato de que a dita etapa de codificar computa, em conjunto, os vetores de animação para as duas ou mais imagens de referência por meio da seleção de uma respectiva ponderação para as mesmas.

55. Método, de acordo com a reivindicação 29,
10 **CARACTERIZADO** pelo fato de que a dita etapa de codificar realiza uma decisão de modo para a imagem de multi-predição a fim de selecionar um modo com uma medida de distorção mínima dentre um modo preditivo único, um modo preditivo direto, um modo bi-preditivo usando os vetores de animação
15 para as duas ou mais imagens de referência respectivamente calculadas usando uma predição única, um modo bi-preditivo usando um respectivo vetor de animação refinado para cada qual dentre duas ou mais imagens de referência, e um modo de predição usando um vetor de animação para uma combinação das
20 duas ou mais imagens de referência calculadas ao se considerar os vetores de animação para as duas ou mais imagens de referência se tornarem idênticas (925).

56. Método, de acordo com a reivindicação 55,
25 **CARACTERIZADO** pelo fato de que a dita etapa de codificar seleciona ainda o modo com a medida de distorção mínima dentre pelo menos um modo bi-preditivo usando um vetor dentre os vetores de animação de uma imagem dentre duas ou mais imagens de referência, e o vetor de animação calculado

ao considerar os vetores de animação para as duas ou mais
imagens de referência se tornarem idênticas, e pelo menos
uma combinação de um vetor de animação direto e o vetor de
animação calculado ao considerar os vetores de animação para
5 as duas ou mais imagens de referência se tornarem idênticas.

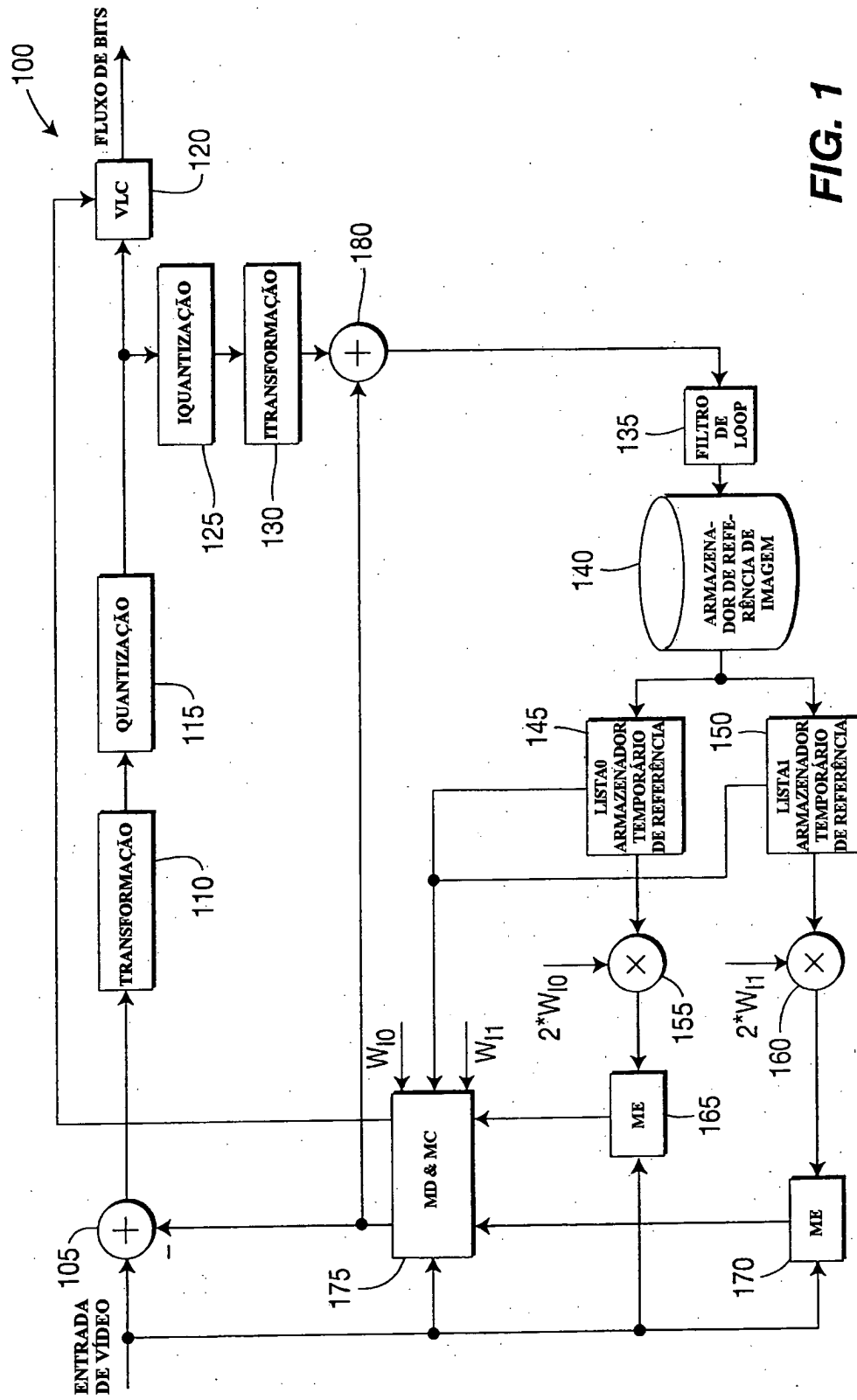


FIG. 1

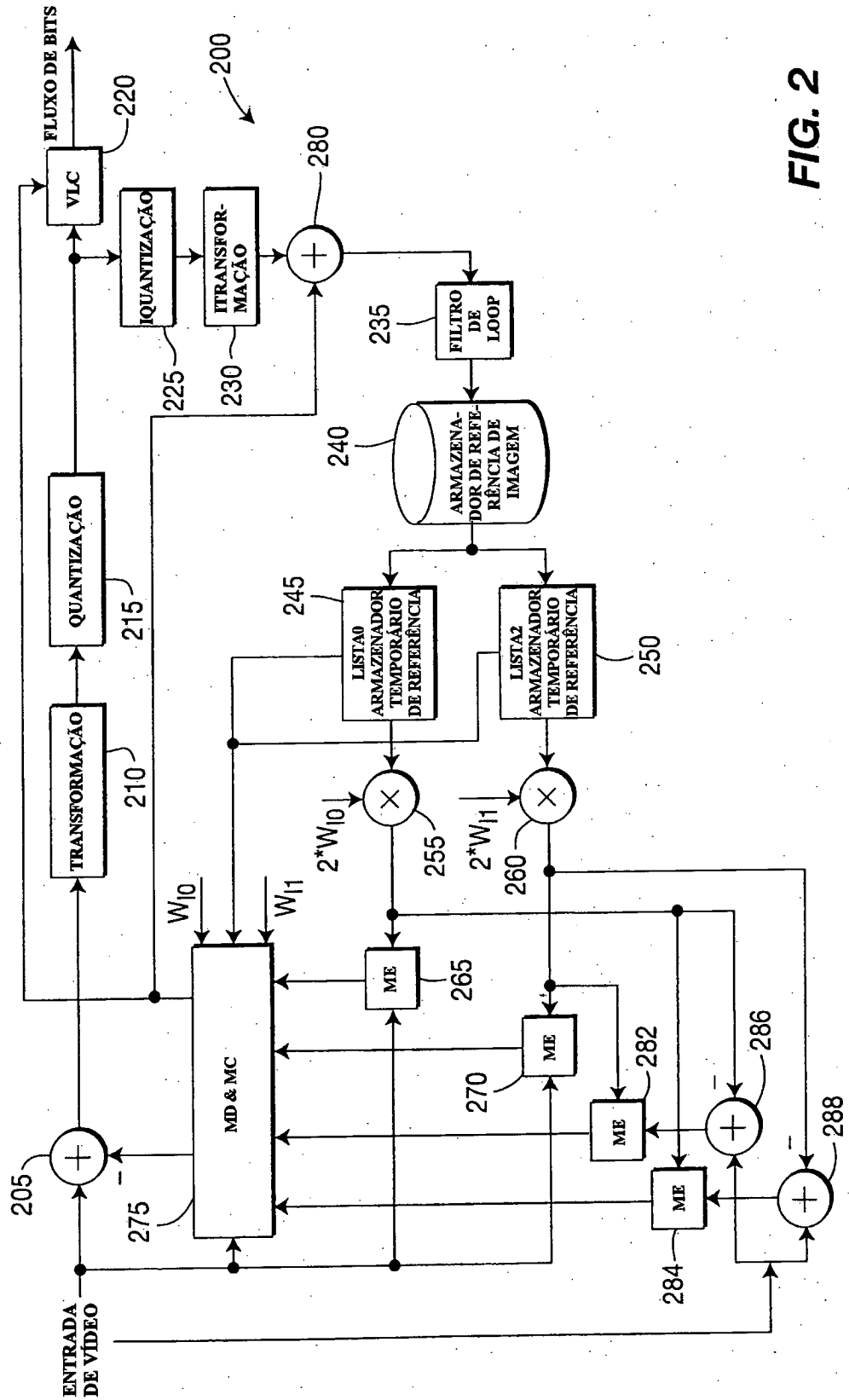
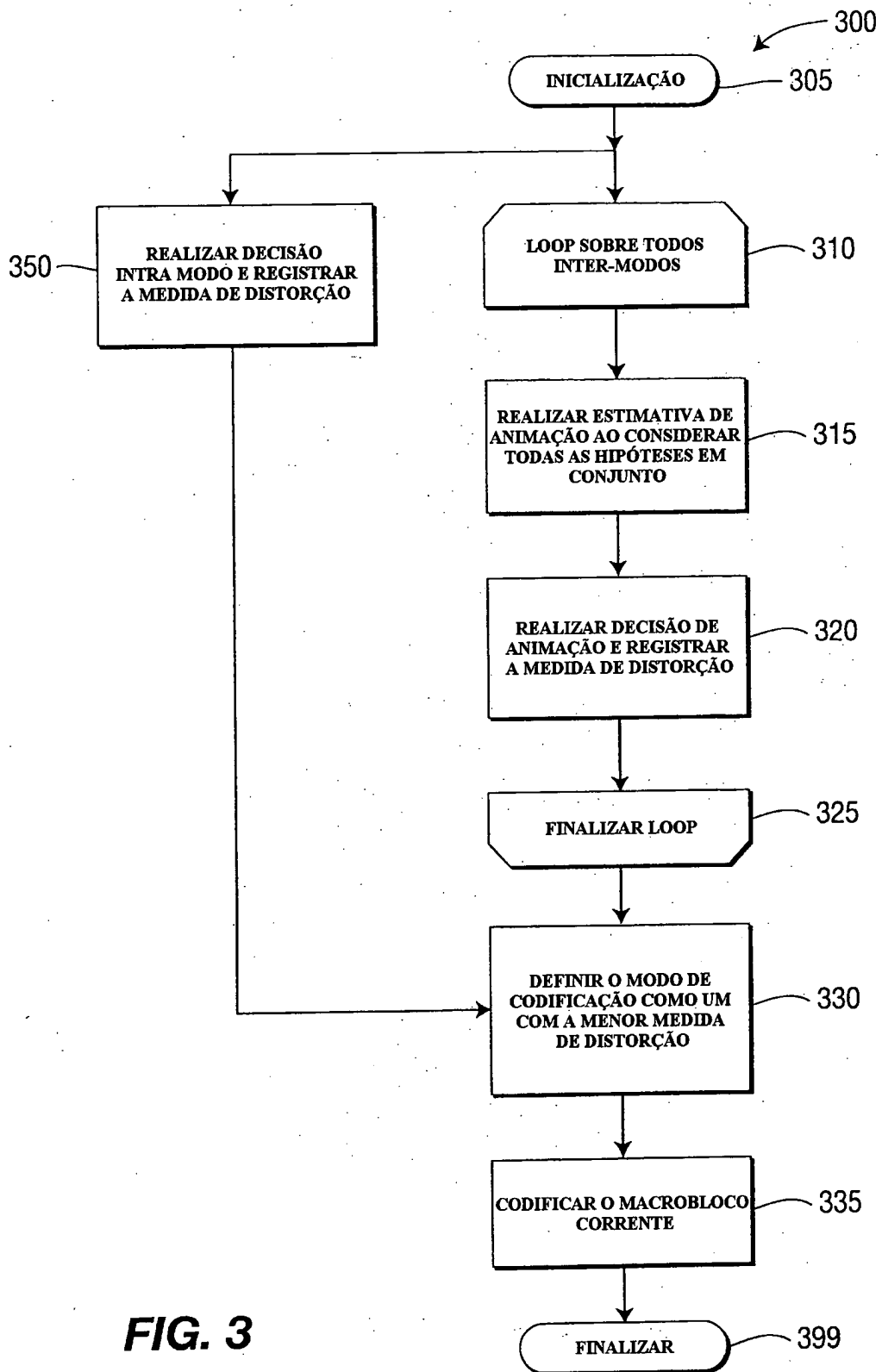


FIG. 2



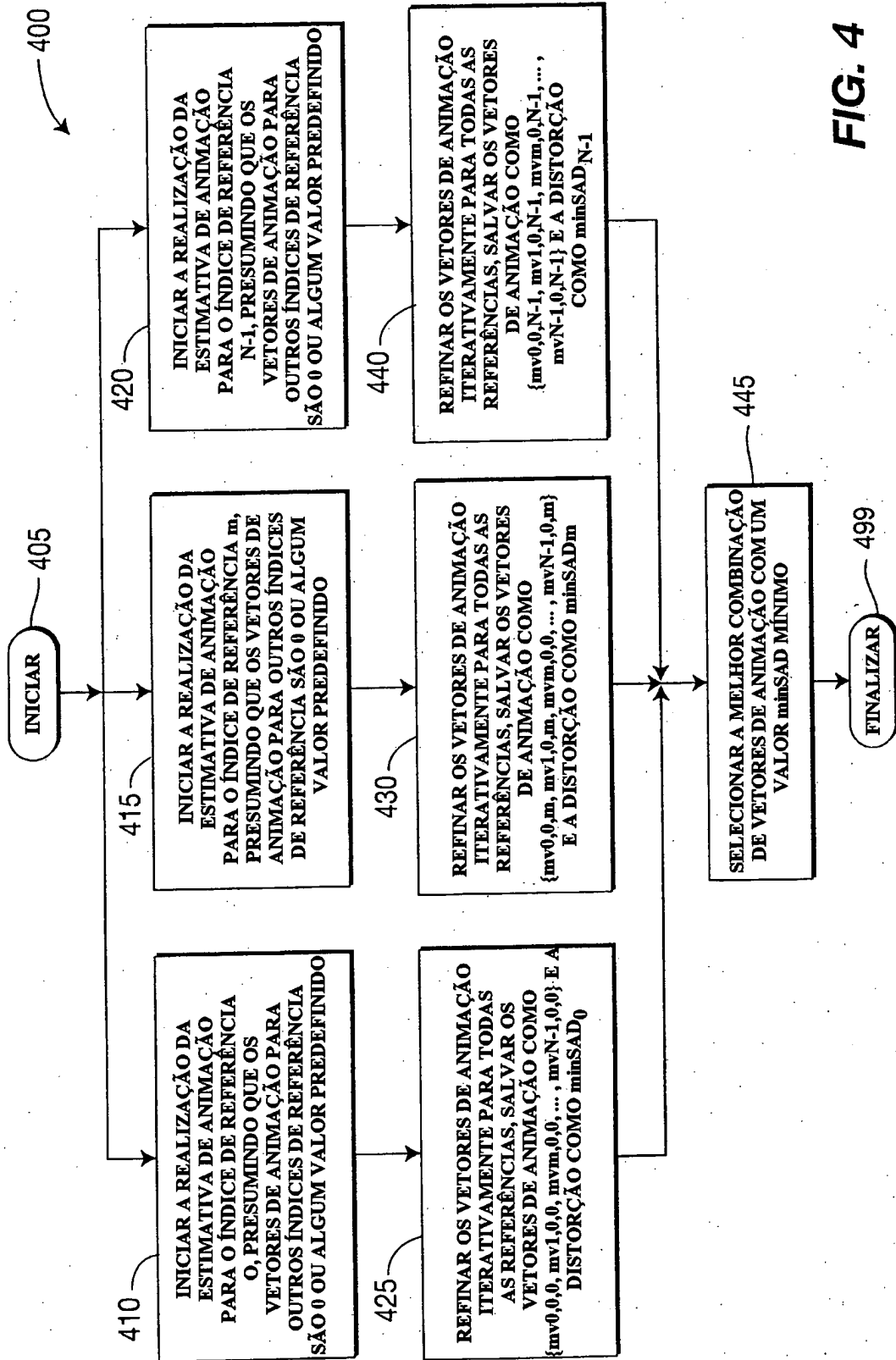
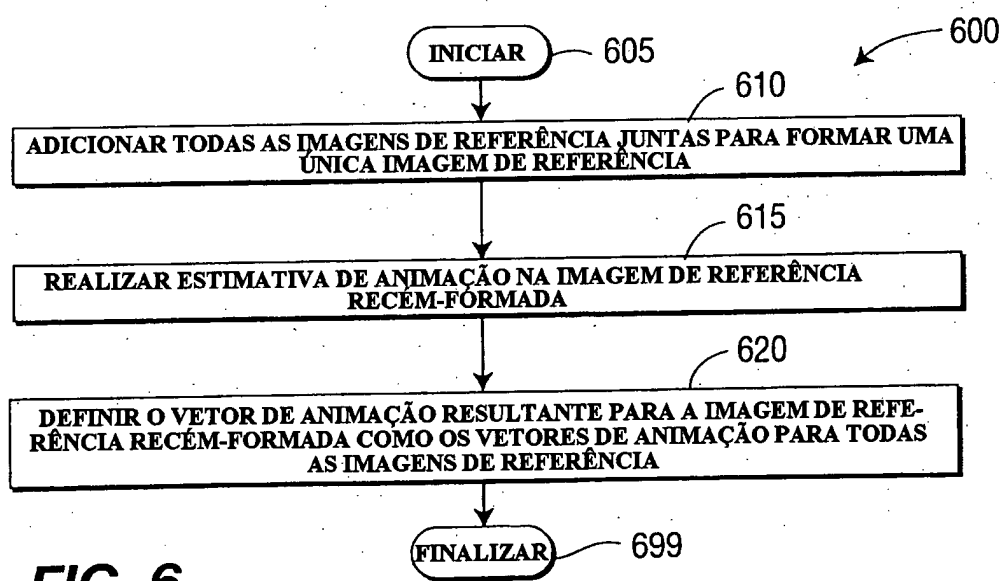
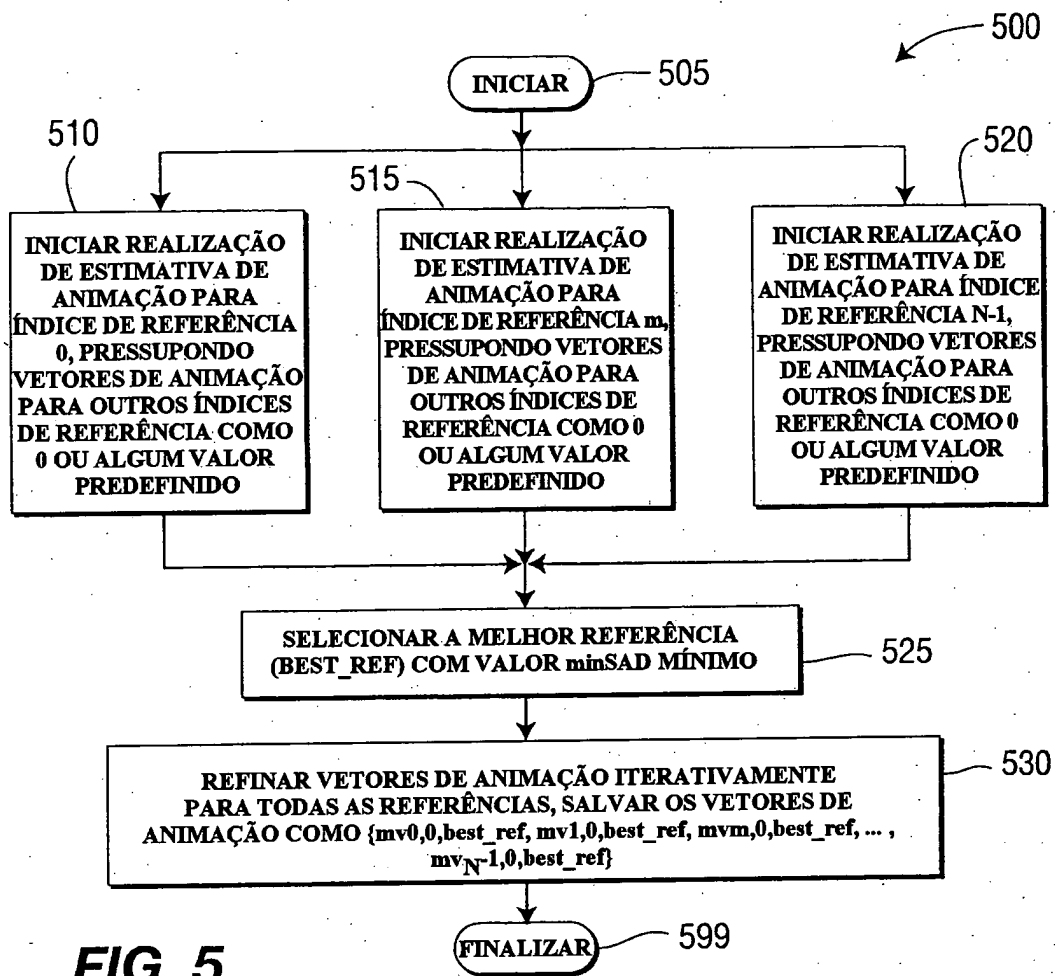


FIG. 4



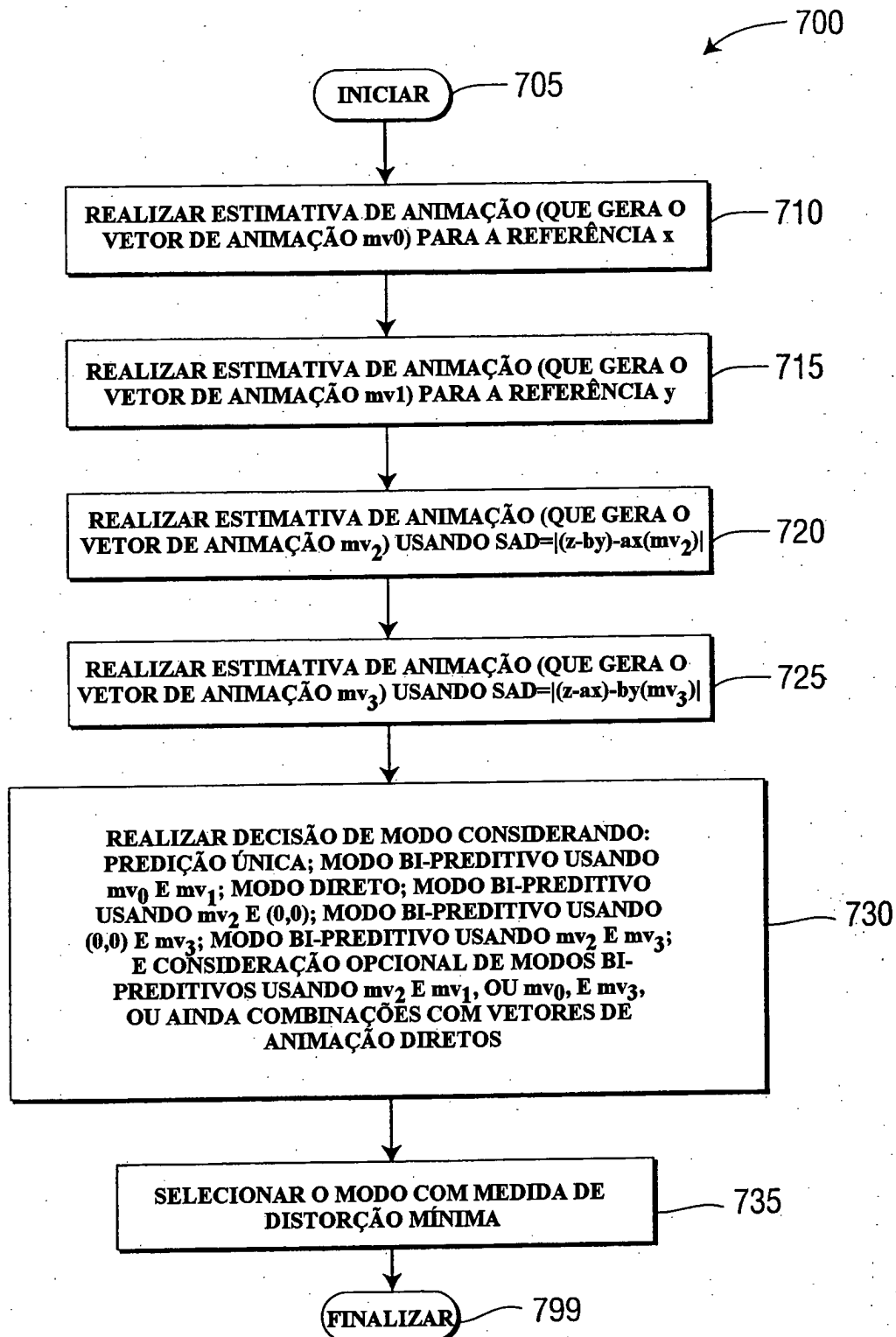
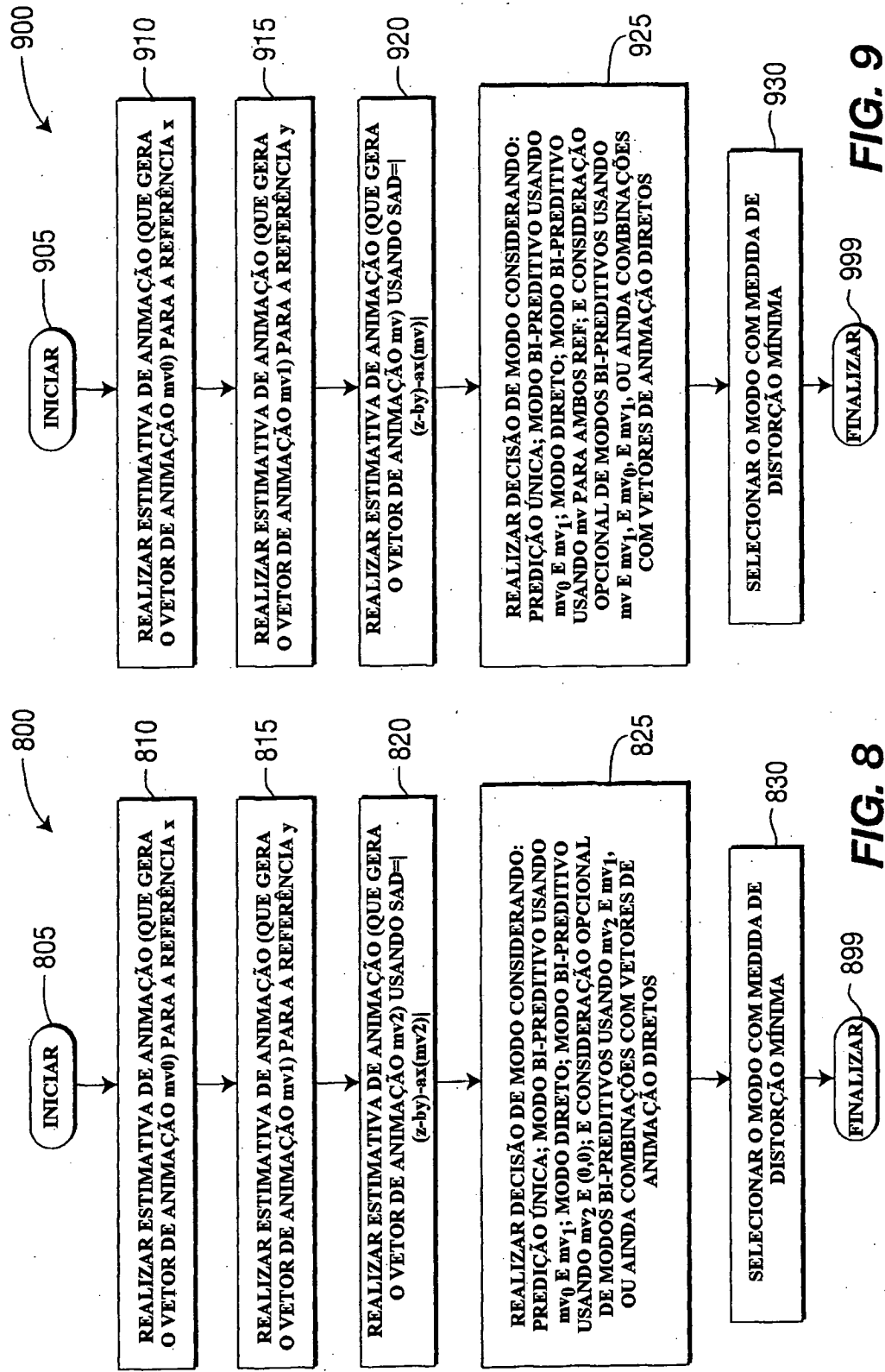


FIG. 7



RESUMO

"MÉTODO E APARELHO PARA ESTIMATIVA DE ANIMAÇÃO
USANDO BI-PREDIÇÃO DE REFERÊNCIAS COMBINADAS"

Um método e aparelho são providos para a
5 estimativa de animação usando bi-predição de referências
combinadas. O aparelho inclui um codificador (200) para a
codificação de uma imagem de multi-predição a partir da
combinação de duas ou mais imagens de referência ao se
predizer respectivamente um vetor de animação para uma
10 imagem particular dentre duas ou mais imagens de referência
em um processo de estimativa de animação enquanto se
inicializa os vetores de animação de modo a manter algumas
dentre duas ou mais imagens de referência em um valor
predefinido para uso pelo processo de estimativa de
15 animação.