



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

(21) **PI0706214-1 A2**

(22) Data de Depósito: 04/01/2007
(43) Data da Publicação: 15/03/2011
(RPI 2097)



(51) *Int.Cl.:*
H04N 7/32

(54) Título: **MÉTODO DE CODIFICAÇÃO E DECODIFICAÇÃO DE VÍDEO, APARELHOS PARA OS MESMOS, PROGRAMAS PARA OS MESMOS E MEIOS DE ARMAZENAMENTO PARA ARMAZENAR OS PROGRAMAS**

(30) Prioridade Unionista: 05/01/2006 JP 2006-000393

(73) Titular(es): National University Corporation Nagoya University,
Nippon Telegraph And Telephone Corporation

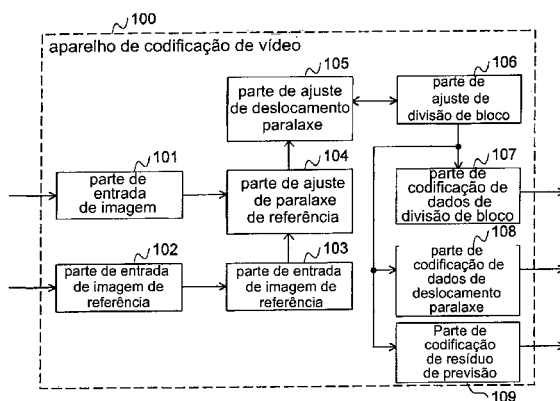
(72) Inventor(es): Hideaki Kimata, Kazuto Kamikura, Kenji
Yamamoto, Masaki Kitahara, Masayuri Tanimoto, Shinya Shimizu,
Toshiaki Fujii, Yoshiyuki Yashima

(74) Procurador(es): Dannemann, Siemsen, Bigler &
Ipanema Moreira

(86) Pedido Internacional: PCT JP2007050005 de 04/01/2007

(87) Publicação Internacional: WO 2007/077989de 12/07/2007

(57) Resumo: MÉTODO DE CODIFICAÇÃO E DECODIFICAÇÃO DE VÍDEO, APARELHOS PARA OS MESMOS, PROGRAMAS PARA OS MESMOS E MEIOS DE ARMAZENAMENTO PARA ARMAZENAR OS PROGRAMAS. A presente invenção refere-se ao uso de compensação paralaxe que realiza previsão pelo uso de paralaxe entre as imagens de vídeo, as imagens de vídeo são codificadas como uma imagem de vídeo única. É ajustada paralaxe de referência para uma imagem-alvo a ser codificada, em que a paralaxe de referência é estimada usando uma imagem de referência; é ajustada divisão de área em uma estrutura de imagem; é ajustado o deslocamento paralaxe para cada área dividida, em que o deslocamento paralaxe é a diferença entre a paralaxe de diferença e a paralaxe para a compensação paralaxe; são codificados os dados de divisão de área; e são codificados os dados para indicar o deslocamento paralaxe. Durante a decodificação, é ajustada a paralaxe de referência para uma imagem-alvo a ser decodificada, em que é estimada usando uma imagem de referência; são decodificados dados para indicar divisão de área, que estão incluídos em dados codificados; e são decodificados os dados de deslocamento paralaxe, que estão incluídos nos dados codificados, para cada área indicada pelos dados de divisão de área.





PI0706214-1

Relatório Descritivo da Patente de Invenção para "MÉTODO DE CODIFICAÇÃO E DECODIFICAÇÃO DE VÍDEO, APARELHOS PARA OS MESMOS, PROGRAMAS PARA OS MESMOS E MEIOS DE ARMAZENAMENTO PARA ARMAZENAR OS PROGRAMAS".

5 Campo da Técnica

A presente invenção refere-se a técnicas para codificação e decodificação de imagens de vídeo de múltiplos pontos de vista.

É reivindicada a prioridade do Pedido de Patente Nº JP 2006-000393, depositado em 05 de janeiro de 2006, cuja descrição encontra-se incorporada ao presente à guisa de referência.

10 Antecedentes da Técnica

As imagens de vídeo de múltiplos pontos de vista são imagens de vídeo obtidas pela fotografia do objeto e fundo das mesmas usando uma pluralidade de câmeras em diferentes posições. Abaixo, uma imagem de vídeo obtida por uma única câmera é chamada uma "imagem de vídeo bidimensional", e um conjunto de imagens de vídeo bidimensionais obtido pela fotografia do mesmo objeto e antecedente do mesmo é chamado uma "imagem de vídeo de múltiplos pontos de vista". Há uma forte correlação entre as imagens de vídeo bidimensionais (de câmeras diferentes) incluídas na imagem de vídeo de múltiplos pontos de vista. Se as câmeras estiverem sincronizadas umas com as outras, as estruturas (das câmeras) correspondentes ao mesmo tempo tiverem capturado o objeto e o fundo do mesmo totalmente no mesmo estado, de maneira que haja uma forte correlação entre as câmeras.

25 Primeiro, serão apresentadas as técnicas convencionais relativas à codificação de imagens de vídeo bidimensionais. Em muitos métodos conhecidos de codificação de imagens de vídeo bidimensionais, como, por exemplo, H.264, MPEG-4, MPEG-2 (que são padrões de codificação internacionais), e similares, é realizada a codificação altamente eficiente por meio de compensação de movimento, transformação ortogonal, quantização, codificação de extensão variável ou similar.

30 Por exemplo, em H.264, cada estrutura I pode ser codificada por

meio de correlação intra-estrutura; cada estrutura P pode ser codificada por meio de correlação intra-estrutura juntamente com uma pluralidade de estruturas anteriores; cada estrutura B pode ser codificada por meio de correlação intra-estrutural juntamente com uma pluralidade de estruturas anteriores ou futuras estruturas.inter-estrutura

Apesar do documento de não-patente I descrever as técnicas H.264 em detalhe, o perfil do mesmo será descrito baixo. Em cada estrutura I, a estrutura é dividida em blocos (chamados "macro-blocos", o tamanho de cada bloco é de 16x16 (pixels)), e é realizada a previsão de estrutura (intra-previsão) em cada macro-bloco. Na intraprevisão, cada macro-bloco é também dividido em blocos menores (denominados "sub-blocos", abaixo), e pode ser aplicado um método de intracodificação individual em cada sub-bloco.

Em cada estrutura P, pode ser realizada intraprevisão ou "inter-prediction" (previsão inter-estrutura) em cada macro-bloco. A intraprevisão aplicada em uma estrutura P é similar àquela aplicada a estrutura I. Na "inter-prediction", é realizada a compensação de movimento. Ainda na compensação de movimento, cada macro-bloco é dividido em blocos menores, e cada sub-bloco pode ser dotado de um vetor de movimento individual e de uma imagem de referência individual.

Ainda na estrutura B, pode ser realiza a intraprevisão ou "inter-prediction". Na "inter-prediction" da estrutura B, além de uma estrutura anterior, uma estrutura futura pode ser referida como uma imagem de referência em compensação de movimento. Por exemplo, ao codificar uma seqüência de estrutura de "I→B→B→P", as estruturas podem ser codificadas na ordem "I→P→B→B". Ainda, em cada estrutura B, a compensação do movimento pode ser realizada por referência a uma estrutura I ou P. Adicionalmente, do mesmo modo que a estrutura P, cada sub-bloco (obtido pela divisão de um macro-bloco) pode ser dotado de um vetor de movimento individual.sub-blocos

Ao realizar uma inter- ou intraprevisão, é obtido um resíduo de previsão. Em cada macro-bloco, um bloco de resíduo de previsão é sujeitado a DCT (transformar cosseno discreto), de maneira a realizar a quantização.

Os valores quantizados obtidos de coeficientes DCT são então sujeitos a codificação de extensão variável.

É conhecido um método para codificar imagens de vídeo de múltiplos pontos de vista, as imagens de vídeo de múltiplos pontos de vista são codificadas com alta eficiência por meio de "compensação paralaxe" na qual é aplicada a compensação de movimento em imagens obtidas por câmeras diferentes ao mesmo tempo. Aqui "paralaxe" é a diferença entre as posições, para as quais é projetado o mesmo ponto em um objeto, em um plano de imagem ou câmeras que estejam dispostas em posições diferentes.

A Figura 9 é uma vista esquemática ilustrando o conceito de paralaxe gerado entre tais câmeras. Na vista esquemática, um plano de imagem de câmeras, cujos eixos geométricos ópticos são paralelos um ao outro, é bloqueado para baixo verticalmente. Geralmente, tais pontos, para os quais é projetado o mesmo ponto em um objeto, em um plano de imagem de câmeras diferentes, são chamados "pontos correspondentes". Um paralaxe pode ser representado como uma diferença posicional no plano de imagem relevante, o mesmo pode ser representado como dados de vetor bidimensional.

Na compensação paralaxe, o ponto correspondente em uma imagem de uma câmera de referência, que corresponde a um pixel-alvo em uma imagem de uma câmera-alvo para a codificação relevante, é estimado usando uma imagem de referência, e o valor do pixel do pixel-alvo é previsto pelo uso de um valor de pixel atribuído ao ponto correspondente. Abaixo, tal "paralaxe estimado" é também chamado "paralaxe" para conveniência da explanação.

O documento de não-patente 2 descreve um método de codificação usando compensação paralaxe, e em tal método, dados paralaxe e cada resíduo de previsão são codificados com relação ao pixels de uma imagem-alvo a ser codificada. Especificamente, no método relevante, a compensação paralaxe é realizada para cada bloco como uma unidade, onde tal paralaxe para cada bloco de unidade é representada usando um vetor bidimensional. A Figura 10 é uma vista esquemática usando um vetor paralaxe.

Isto é, nesse método os dados paralaxe são codificados como vetor bidimensional e o residual de previsão relevante. Como esse método não usa parâmetros de câmera na codificação, é eficaz quando os parâmetros de câmera são desconhecidos.

5 Além disso, quando há uma pluralidade de imagens de referência obtidas por câmeras diferentes, a compensação paralaxe pode ser realizada usando uma técnica de imagem de ponto de vista arbitrário. O documento de não-patente 3 descreve a compensação paralaxe usando uma técnica de imagem de ponto de vista arbitrário. Especificamente, cada valor
10 de pixel de uma imagem obtida por uma câmera-alvo para a codificação relevante é previsto por meio de interpolação usando os valores de pixel dos pontos correspondentes (pertencentes a câmeras diferentes) que correspondam ao pixel relevante. A Figura 11 é uma vista esquemática usando tal interpolação. "Na interpolação, o valor de pixel m em uma imagem-alvo a ser
15 codificada é previsto pela formação de interpolação entre m' e m'' das imagens de referência 1 e 2, onde os pixels m' e m'' correspondem ao pixel m .

Quando há duas ou mais imagens de referência obtidas por câmeras diferentes (conforme descrito no Documento de Patente 3), a paralaxe de cada pixel de uma imagem-alvo (a ser codificada) para cada imagem
20 de referência pode ser estimada sem o uso de imagem-alvo. A Figura 12 é uma vista esquemática para ilustrar o conceito de tal estimativa de paralaxe.

Conforme ilustrado na Figura 12, na verdadeira paralaxe, os valores de pixel dos pontos correspondentes entre as imagens de referência relevante devem ser quase idênticas umas as outras. Portanto, em muitos
25 métodos de estimativa de paralaxe, com relação a cada das várias profundidades, os valores de pixel dos pontos correspondentes entre as imagens de referência são comparados uns aos outros, e a paralaxe pode ser estimada com base na profundidade que conduz os valores de pixel mais próximos. Tal processo pode ser aplicado a cada pixel de uma imagem-alvo a ser codi-
30 ficada.

Conforme acima descrito, quando há duas ou mais imagens de referência obtidas por câmeras diferentes, e é possível a estimativa de para-

laxe na face de decodificação, então a compensação paralaxe pode ser realizada no lado de decodificação usando dados paralaxe para cada pixel, sem prover dados paralaxe, que são explicitamente codificados no lado de codificação, para o lado de decodificação.

- 5 Documento de não-patente 1: ITU-T Rec.H.264H.264/ISSO/IEC 11496-10, "Codificação de Vídeo Avançada", Minuta Final de Comitê, Documento JVT-E022, setembro de 2002.

- Documento de não-patente 2: Hediaki Kimata e Masaki Kitahara, "Resultados Preliminares em codificação de vídeo de múltiplas vistas (3DAV)", Reunião de Redmond documento M10976 MPEG, julho de 2004.

Documento de não-patente 3: Masayuki Tanimoto, Toshiaki Fujii, "Resposta ao Chamado para Evidencia em Codificação de Vídeo de Múltiplas Vistas", documento Mxxxxx Reunião de Hong Kong, janeiro de 2005.

- Nas técnicas convencionais, quando há duas ou mais imagens de referência obtidas por câmeras diferentes, e estimativa paralaxe é possível no lado de decodificação, então a compensação paralaxe pode ser realizada na face de decodificação pelo uso de dados paralaxe para cada pixel, sem proporcionar dados paralaxe, que estão explicitamente codificados no lado de codificação, para o lado de decodificação. Tal paralaxe com relação a uma imagem-alvo a ser codificada ou decodificada, que pode ser estimada no lado de codificação ou decodificação sem o uso de imagem-alvo (para a decodificação, sem decodificar a imagem relevante), é chamada "paralaxe de referência".

- Contudo, a paralaxe de referência, que é estimada no lado de decodificação, é uma ótima em razão da eficiência da previsão, e a quantidade do código atribuído à previsão relevante residual pode ser aumentada. Quando a paralaxe para maximizar a eficiência da previsão é computada no lado de codificação, e a diferença (chamada abaixo "deslocamento paralaxe") entre a paralaxe computada e a paralaxe de referência é codificada para cada pixel, pode ser aperfeiçoada a eficiência da previsão, aperfeiçoando por meio disso a eficiência da previsão com relação à resíduo de previsão.

Contudo, em tal técnica que pode ser facilmente antecipada, à

medida que o deslocamento paralaxe é codificado para cada pixel, aumenta a quantidade de código para os dados paralaxe, de maneira que a eficiência da codificação não pode ser alta.

Descrição da Invenção

5 Para solucionar o problema acima, é um objetivo da presente invenção diminuir a quantidade do código de dados paralaxe sem fazer sacrificar consideravelmente a precisão da compensação paralaxe.

A característica mais evidente da presente invenção em comparação com as técnicas convencionais é que a fim de solucionar os problemas acima descritos, são realizados um processo de determinação de divisão de bloco e um processo de codificação dos dados de divisão de bloco 10 determinados e dados de deslocamento paralaxe para cada bloco, e tais dados são lidados como dados codificados.

Em um primeiro modo da codificação de vídeo e métodos de decodificação da presente invenção, com base na divisão de área, que é 15 ajustada de acordo com a característica de uma imagem-alvo a ser codificada, podem ser codificados os dados de deslocamento paralaxe para cada área dividida.

Geralmente, o deslocamento paralaxe é dotado de correlação espacial em uma estrutura de imagem. Portanto, quando é ajustada a divisão de área ajustada, e é codificado o deslocamento paralaxe para cada 20 área dividida, então a quantidade de dados paralaxe pode ser reduzida sem degradar a eficiência de previsão de compensação paralaxe.

No lado de codificação de vídeo, em uma etapa de ajuste de referência paralaxe, pode ser codificada a paralaxe para uma imagem-alvo a 25 ser codificada (isto é, paralaxe de referência) é ajustada usando os dados já codificados (isto é, imagem de referência), e em uma etapa de ajuste de divisão de área, é ajustada a divisão de área em uma estrutura da imagem-alvo. Em uma etapa de ajuste de deslocamento paralaxe, o deslocamento paralaxe é ajustado para cada área dividida na divisão de área. Adicionalmente, em uma etapa de codificação de dados de divisão de dados, são co- 30 dificados os dados de divisão de área para indicar a divisão de área, e em

uma etapa de codificação de dados de deslocamento paralaxe, são codificados os dados de deslocamento paralaxe.

Como a paralaxe de referência ajustada na etapa de ajuste da paralaxe de referência, (1) a paralaxe estimada usando imagens de referência no princípio conforme ilustrado na Figura 12 (referir a um terceiro modo como explicado posteriormente); (ii) paralaxe com base em uma imagem paralaxe ou um modelo tridimensional, que é codificado por outro dispositivo de maneira a, por exemplo, gerar uma imagem de ponto de vista arbitrário, e enviar para o lado de decodificação; ou (iii) paralaxe provida por outro método.

No lado de decodificação, em uma etapa de ajuste de paralaxe de referência, similar ao lado de codificação, a paralaxe de referência para uma imagem-alvo a ser decodificada é ajustada usando dados já decodificados (isto é, imagem de referência). Em uma etapa de decodificação de dados de divisão de área, são decodificados os dados de divisão de área, e em uma etapa de decodificação de dados de deslocamento paralaxe, são decodificados os deslocamento paralaxe para cada área de acordo com os dados de divisão de área.

Um segundo modo de métodos de codificação e decodificação de vídeo da presente invenção é basicamente similar ao primeiro modo acima descrito. Contudo, no segundo modo, é formada uma divisão de área de maneira a gerar blocos retangulares. Especificamente, a divisão de bloco retangular (referir à Figura 10 explicada posteriormente) é definida para cada macro-bloco. Os dados para tal divisão de bloco podem ser eficientemente codificados pelo uso de técnica de codificação de entropia conhecida (por exemplo, codificação de entropia em H.264).

Um terceiro modo de métodos de codificação e decodificação de vídeo da presente invenção é basicamente similar aos primeiro e segundo modos acima descritos. Contudo, no terceiro modo, é usada a paralaxe de referência, que é estimada por uma pluralidade de imagens de referência sem usar uma imagem-alvo a ser codificada (isto é, de acordo com o princípio acima descrito conforme ilustrado pela Figura 12). Quando a paralaxe de

referência é estimada usando imagens de referência conforme explicado acima, pode ser estimada pelo uso apenas de dados que já tenham sido enviados para o lado de decodificação. Portanto, é possível não gerar dados adicionais a serem codificados, de maneira a estimar a paralaxe de referência.

De acordo com a presente invenção, nos métodos de codificação e decodificação de vídeo para estimar paralaxe pelo uso de imagens de referência, é usada a correlação espacial dos dados de deslocamento paralaxe, de maneira que a quantidade de código de deslocamento paralaxe possa ser reduzida sem degradar a eficiência de previsão na compensação paralaxe, por meio disso aperfeiçoando a eficiência de codificação total.

Breve Descrição dos Desenhos

A Figura 1 é um diagrama ilustrando um aparelho de codificação de vídeo como uma modalidade da presente invenção.

A Figura 2 é um diagrama ilustrando a relação entre câmeras na modalidade.

A Figura 3 é um diagrama ilustrando a disposição das câmeras na modalidade.

A Figura 4 é um fluxograma de codificação na modalidade.

A Figura 5 é um fluxograma detalhado da etapa S106 na Figura 4.

A Figura 6 é um diagrama ilustrando exemplos de divisão de bloco em cada macro-bloco.

A Figura 7 é um diagrama ilustrando um aparelho de codificação de vídeo na modalidade.

A Figura 8 é um diagrama de decodificação na modalidade.

A figura 9 é uma vista esquemática ilustrando o conceito de paralaxe gerada entre as câmeras.

A Figura 10 é uma vista esquemática ilustrando um vetor paralaxe.

A Figura 11 é uma vista esquemática ilustrando interpolação de valor de pixel.

A Figura 12 é uma vista esquemática ilustrando o conceito de estimativa paralaxe.

Melhores Modos para Realizar a Invenção

5 A Figura 1 é um diagrama ilustrando a estrutura de um aparelho de codificação de vídeo como uma modalidade da presente invenção.

O aparelho de codificação de vídeo 100 inclui uma parte de entrada de imagem 101 na qual é entrada cada imagem original da câmera C (isto é, imagem-alvo a ser codificada), uma parte de entrada de imagem de referência 102 na qual são entradas as imagens decodificadas (como ima-
10 gens de referência) das câmeras A e B; uma memória de imagem de referência 103 para armazenar cada imagem de referência; uma parte de ajuste de paralaxe de referência 104 para obter paralaxe de referência pelo uso de imagens de referência; uma parte de ajuste de deslocamento paralaxe 105 para obter deslocamento paralaxe; uma parte de ajuste de divisão de bloco
15 106 para ajustar um estado de divisão de bloco, uma parte de codificação de dados de divisão de bloco 106 para codificar dados de divisão de bloco de codificação; uma parte de codificação de dados de deslocamento paralaxe 108 para codificar dados de deslocamento paralaxe; e uma parte de previsão de codificação residual 109 para codificar o previsão relevante residual.

20 A Figura 2 é um diagrama ilustrando relação de referência entre as câmeras na presente modalidade. Conforme ilustrado na Figura 2, na presente modalidade para codificar imagens de vídeo de múltiplos pontos de vista obtidas por três câmeras, as imagens de vídeo da câmera C são codificados usando imagens de decodificação das câmeras A e B, como imagens
25 de referência.

As setas na Figura 2 representam as relações de referência em compensação paralaxe. A fim de codificar cada imagem da câmera C, as imagens decodificadas (das câmeras A e B) sendo dotadas do mesmo tempo de exibição são codificadas como imagens de referência. No processo
30 relevante, uma imagem prevista é gerada pela computação de uma média entre os valores de pixel dos pontos correspondentes pertencentes às câmeras A e B (cada ponto correspondente é um pixel indicado pelo vetor que é

proporcionado pela soma do vetor de deslocamento paralaxe relevante e o vetor paralaxe de referência).

A Figura 3 é um diagrama ilustrando a disposição das câmeras na presente modalidade. Na presente modalidade, conforme ilustrado na
5 Figura 3, os pontos de vista das três câmeras se alinham ao longo de uma linha reta e intervalos regulares, e os eixos geométricos ópticos das câmeras são perpendiculares à linha reta. Isto é, os eixos geométricos ópticos das três câmeras são paralelos uns aos outros.

Além disso, o sistema de coordenada x-y no plano de imagem
10 relevante pode ser obtido pela translação paralela (não é realizada nenhuma rotação ou similar) com relação à linha reta na qual as câmeras estão dispostas, e os pixels são formados pela divisão de cada eixo geométrico x e y do plano de imagem em intervalos regulares para cada câmera. Isto é, cada câmera é dotada da mesma resolução, e uma paralaxe de pixels P entre as
15 câmeras C e A corresponde a uma paralaxe de pixels P entre as câmeras C e B.

A Figura 4 é um fluxograma de codificação na presente modalidade. A Figura 5 é um fluxograma detalhado da etapa S 106 na Figura 4.

Na presente modalidade, é usado um macro-bloco consistindo
20 em pixels 16x16 (nas direções vertical e horizontal) como uma unidade, e é atribuída a divisão (estado) de bloco em cada macro-bloco. São codificados os dados de deslocamento paralaxe para cada bloco definido na divisão de bloco (simplesmente denominado "bloco", por conveniência).

A paralaxe de referência é representada usando um vetor bidi-
25 mensional. Portanto, no lado de decodificação (e no lado de codificação), um vetor bidimensional (isto é, paralaxe de referência) para cada imagem de referência é obtido para cada pixel, pelo uso da imagem de referência.

Por outro lado, é codificado um único deslocamento paralaxe (para a câmera A) representado por um vetor bidimensional para cada bloco.
30 Aqui, similar à computação da paralaxe de referência, é presumida uma condição de restrição física na qual cada (paralaxe) vetor (usado para compensação paralaxe), que é proporcionada como a soma do vetor de paralaxe

de referência e um vetor de paralaxe de referência para cada pixel de cada imagem de referência, designa a mesma posição do objeto. Nesse caso, se for conhecido o vetor de deslocamento paralaxe para a câmera A, é unicamente determinado o vetor de deslocamento paralaxe para outra câmera.

5 Em vez disso, o deslocamento paralaxe para cada imagem de referência pode ser computado independentemente e codificado sem presumir tal condição de restrição. Contudo, tal método pode ser facilmente implementado como uma variação da presente modalidade, e são omitidas explicações específicas do mesmo.

10 Há muitos estados de divisão de bloco possíveis aplicáveis a cada macro-bloco. A Figura 6 ilustra exemplos dos mesmos. Conforme ilustrado na Figura 6, "Modo blk" é um índice para indicar o tipo de divisão de bloco, e o número de blocos para cada "Modo blk" é indicado por "max-Blk[Modo blk]".

15 Sob as condições acima, a operação de codificação será explicada com relação ao fluxograma da Figura 4.

 Primeiro, uma imagem da câmera C é entrada na parte de entrada de imagem 101 (ver etapa S101), onde as imagens decodificadas, que pertencem às câmeras A e B e são dotadas do mesmo tempo de exibição,
20 foram armazenadas na memória de imagem de referência 103 por via da parte de entrada de imagem de referência 102.

 A seguir, as imagens decodificadas relevantes das câmeras A e B são entradas a partir da memória de imagem de referência 103 (ver etapa S102), e a paralaxe de referência para cada pixel de uma imagem-alvo a ser
25 codificada é obtida para cada imagem de referência de entrada pelo uso das imagens de referência (ver etapa S103). Isto é, as duas imagens de referência, que pertencem às câmeras A e B, são entradas na parte de ajuste paralaxe de referência 104, e é obtida a paralaxe de referência para cada pixel da imagem-alvo da câmera C.

30 Portanto, são obtidos os dois "vetores bidimensionais" para cada pixel da imagem-alvo da câmera C. Para as coordenadas (x,y) no plano de imagem da câmera C (x e y são um inteiro que indicam um valor de coorde-

nada do pixel relevante), os vetores paralaxe de referência para as câmeras A e B são respectivamente indicados $d_A[x,y]$ e $d_B[x,y]$.

Além disso, "MBBlk" é um índice para indicar cada macro-bloco, e o número de macro-blocos é indicado por "maxMBBlk".

- 5 Após o índice MBBlk para cada macro-bloco ser inicializado para zero (ver etapa S104), as etapas que seguintes (S105 e S111) são repetidamente aplicadas em cada macro-bloco ao mesmo tempo em que é adicionado "1" a cada índice MBBlk (ver etapa S110).

- 10 Primeiro, após a divisão de bloco (índice) o Modo blk é inicializado para zero (ver etapa S105), um custo de distorção de taxa para o macro-bloco relevante e a divisão de bloco Modo blk é computado (ver etapa S106) ao mesmo tempo em que é adicionado "1" à divisão de bloco Modo blk (ver etapa S107), até que a divisão de bloco Modo blk alcance o valor de [índice máximo "maxModo blk" (ver etapa S108). Isto é, na parte de ajuste
15 de deslocamento paralaxe 105, é computado o custo de distorção de taxa "Custo MB" para cada divisão de bloco Modo blk.

- O custo de distorção de taxa Custo MB para cada divisão de bloco Modo blk é obtido pela computação de um custo de distorção de taxa "Custo blk" para cada bloco (o número total de blocos é indicado por "max-
20 Blk[Modo blk]"), e pela computação da soma total dos custos de distorção de taxa de Custo blk.

- O custo de distorção de taxa de Custo blk para cada bloco pode ser obtido como um valor mínimo do custo de distorção de taxa "custo" obtido ao usar um vetor de deslocamento paralaxe. Portanto, devem ser compu-
25 tados, para cada bloco, o vetor de deslocamento paralaxe para minimizar "custo" e o próprio "custo".

- Aqui, a busca para o deslocamento paralaxe usado para codificar é realizada na adjacência da paralaxe de referência relevante. Isto é, " $e_0, e_1, \dots, e_{N-1}$ " são definidos como candidatos para o vetor de deslocamento
30 paralaxe, e um processo de computação de custo de distorção de taxa para o vetor paralaxe " $d_A[x,y] + e_n$ " é repetido para " $n=0,1,\dots,N-1$ ", de maneira a codificar o vetor paralaxe ótimo para o presente bloco.

Para computar o "custo", primeiro, é computada a soma total "SAD[e_n]" dos valores absolutos da resíduo de previsão para o bloco relevante, que é obtida ao usar um vetor de deslocamento paralaxe e_n. Além disso, é também computada uma quantidade estimada R[e_n] do código do vetor de deslocamento paralaxe, que é obtido ao codificar um vetor de deslocamento paralaxe e_n. O "custo" é computado pela seguinte fórmula:

$$\text{Custo} = \text{SAD}[e_n] + \lambda R[e_n] \quad (1)$$

A seguir, na parte de ajuste de divisão de bloco 106, a divisão de bloco "melhor Modo blk" para minimizar o custo de distorção de taxa Custo MB é computado para cada macro-bloco MBBlk (o vetor de deslocamento paralaxe correspondente já foi obtido).

A Figura 5 ilustra o processo detalhado da etapa S106.

Primeiro, o custo de distorção de taxa Custo MB do macro-bloco relevante e o valor de índice do bloco "blk" são inicializados para zero (ver etapa S1061), e é obtido o vetor de deslocamento paralaxe para minimizar o custo de distorção de taxa de Custo blk do bloco "blk" (ver etapa S1062).

Após a obtenção do custo de distorção de taxa de Custo blk do bloco relevante é adicionado ao custo de distorção de taxa Custo MB do macro-bloco (ver etapa S1063), é adicionado "1" ao bloco "blk", e são repetidas as etapas S1062 a S1064 até que o bloco "blk" alcance o número total maxBlk[Modo blk] de blocos (ver etapa S1065).

Quando o bloco "blk" alcança o número total maxBlk[Modo blk] de blocos, é determinado se o custo de distorção de taxa presentemente computado Custo MB do macro-bloco é menor do que o custo de distorção de taxa mínima computado minCusto MB do macro-bloco (ver etapa S1066). Se for menor, o custo de distorção de taxa Custo MB é armazenado como o custo de distorção de taxa mínima minCusto MB, e a divisão de bloco correspondente Modo blk é armazenada como a divisão de bloco ótima melhor Modo blk (ver etapa S1067).

De acordo com a operação acima, são obtidos os dados da divisão de bloco "melhor Modo blk", os dados de deslocamento paralaxe (isto é o vetor de deslocamento paralaxe), e a resíduo de previsão gerado pela

compensação paralaxe, que são usados para a compensação paralaxe do macro-bloco MBBIk. A seguir, são codificados os dados da divisão de bloco "melhor Modo blk" na parte de codificação dos dados de divisão de bloco 107, e são codificados os dados de deslocamento paralaxe correspondentes
 5 ao melhor Modo blk na parte de codificação de dados de deslocamento paralaxe 108 (ver etapa S109).

Além disso, o resíduo de previsão correspondente aos dados de deslocamento paralaxe é codificada na parte previsão de codificação residual 109.

10 A figura 7 ilustra um aparelho de decodificação de vídeo na presente modalidade. O aparelho de decodificação de vídeo 200 inclui uma parte de decodificação de dados de divisão de bloco 201; uma parte de decodificação de dados de deslocamento paralaxe 202; uma parte de previsão de decodificação de residual 203, uma parte de compensação paralaxe 204 e
 15 uma memória de imagem de referência 205.

A Figura 8 é um fluxograma da decodificação usando o aparelho de decodificação de vídeo 200 na presente modalidade. O fluxograma ilustra a operação para decodificar uma estrutura da câmera C, e será explicado detalhadamente abaixo. Aqui, as estruturas das câmeras A e B do mesmo
 20 período foram antecipadamente decodificadas, e as imagens decodificadas foram armazenadas na memória de imagem de referência 205.

Primeiro, as imagens decodificadas relevantes das câmeras A e B são entradas a partir da memória de imagem de referência 205 (ver etapa S201), e a paralaxe de referência para cada pixel da imagem-alvo a ser codificada é obtida usando cada imagem de referência (ver etapa S202). Isto é,
 25 duas imagens de referência pertencentes às câmeras A e B são entradas na parte de compensação paralaxe 204, e é obtida a compensação paralaxe para cada pixel da imagem relevante da câmera C. Portanto, são obtidos os "vetores bidimensionais" para cada pixel na imagem da câmera C.

30 A seguir o índice MBBIk do macro-bloco é inicializado para zero (ver etapa S203), as etapas seguintes (S204 e S212) são repetidas para cada macro-bloco da uma estrutura relevante ao mesmo tempo em que é adi-

cionado "l" ao MBBlk (ver etapa S211) ("maxMBBlk" indica o número de macro-blocos em uma estrutura). Portanto, é decodificada a uma estrutura relevante da câmera C.

Na decodificação de cada macro-bloco, primeiro, os dados de divisão de bloco melhor Modo blk para o macro-bloco MBBlk é decodificado pela parte de decodificação de dados de divisão de bloco 201 (ver etapa S204). A seguir, após o valor de índice de loco "blk" ser inicializado para zero (ver etapa S205), são repedidas as etapas seguintes (S206 a S209) para cada bloco "blk" ("maxBlk[Modo blk] indica o número máximo de bloco).

Primeiro, na parte de decodificação de dados de deslocamento paralaxe 202, são decodificados os dados de deslocamento paralaxe para o bloco "blk" (ver etapa S206), e é gerada uma imagem de previsão para o bloco "blk" usando os dados de deslocamento paralaxe e a paralaxe de referência (isto é, pela interpolação entre os valores de pixel das câmeras A e B) (ver etapa S207).

De acordo com o processo de repetição acima para o bloco "blk", é gerada uma imagem prevista para o macro-bloco MBBlk. Portanto, a resíduo de previsão para o macro-bloco MBBlk é decodificada na parte de decodificação de resíduo de previsão 203, e a parte de compensação paralaxe 204 computa a soma da imagem prevista e a resíduo de previsão, obtendo, por meio disso, uma imagem decodificada do macro-bloco MBBlk (ver etapa S210).

A operação acima é repetida para cada macro-bloco na uma armção relevante (ver etapas S211 e S212), por meio disso decodificando a uma estrutura da câmera C.

Na presente modalidade, uma imagem da câmera C é codificada por referência às outras câmeras (A e B). Contudo, tal caso de realização de codificação por referência a outras câmeras e outro caso de realização de compensação de movimento por referência a uma imagem decodificada da câmera C pode ser adaptativamente comutada em uma estrutura de imagem (isto é, para uma imagem).

Especificamente, a comutação entre a compensação paralaxe e

a compensação de movimento pode ser realizada para cada bloco dividido ilustrado na Figura 6. Nesse caso, o lado de codificação de vídeo deveria codificar dados para indicar qual compensação paralaxe e compensação de movimento foi usada, e o lado de decodificação de vídeo deveria decodificar os dados.

Os processos de codificação e decodificação de vídeo acima descritos podem ser implementados pelo uso de um computador e um programa de software. Tal programa pode ser provido pelo armazenamento do mesmo em um meio de armazenamento legível de computador, ou através de uma rede.

Aplicabilidade Industrial

Nos métodos de codificação e decodificação de vídeo para estimar paralaxe pelo uso de imagens de referência, são usados correlação espacial de dados de deslocamento paralaxe, de maneira que a quantidade de código do deslocamento paralaxe possa ser reduzido sem degradar a eficiência de previsão na compensação paralaxe, por meio disso aperfeiçoando a eficiência de codificação total.

REIVINDICAÇÕES

1. Método de codificação de vídeo para codificar imagens de vídeo como uma imagem de vídeo única pelo uso de compensação paralaxe que realiza previsão pelo uso de paralaxe entre as imagens de vídeo, o método compreendendo:

uma etapa de ajuste paralaxe de referência de ajuste paralaxe de referência para uma imagem-alvo a ser codificada, em que a paralaxe de referência é estimada usando uma imagem de referência;

uma etapa de ajuste de divisão de área de ajuste de divisão de área em uma estrutura de imagem;

uma etapa de ajuste de deslocamento paralaxe de ajuste de deslocamento paralaxe para cada área dividida ajustada na etapa de ajuste de divisão de área, em que o deslocamento paralaxe é a diferença entre a paralaxe de referência e a paralaxe usada para a compensação paralaxe;

uma etapa de codificação de dados de divisão de dados de divisão de área de codificação para indicar a divisão de área na qual é ajustada na etapa de ajuste de divisão de área; e

uma etapa de codificação de dados de deslocamento paralaxe de codificação de dados de deslocamento paralaxe para indicar o deslocamento paralaxe que é ajustado na etapa de ajuste de deslocamento paralaxe.

2. Método de codificação de vídeo, de acordo com a reivindicação 1, em que:

o ajuste de divisão de área na etapa de ajuste de divisão de área é realizado por um método selecionado dentre uma pluralidade de métodos de divisão de área para realizar a divisão de área em blocos retangulares.

3. Método de codificação de vídeo, de acordo com a reivindicação 1, em que:

o ajuste de paralaxe de referência na etapa de ajuste de paralaxe de referência é estimado usando uma pluralidade de imagens de referência sem usar a imagem-alvo a ser codificada.

4. Método de codificação de vídeo, de acordo com a reivindica-

ção 3, em que:

a paralaxe de referência é estimada para cada pixel como uma unidade, pelo uso de imagens de referência.

5 5. Método de codificação de vídeo para decodificar imagens de vídeo como uma única imagem de vídeo pelo uso de compensação paralaxe que realiza previsão pelo uso de paralaxe entre as imagens de vídeo, o método compreendendo:

10 uma etapa de ajuste de paralaxe de referência para uma imagem-alvo a ser decodificada, em que a paralaxe de referência é estimada usando uma imagem de referência;

uma etapa de decodificação de dados de divisão de dados de divisão de área de decodificação de dados de divisão e área para indicar divisão de área, em que os dados de divisão de área estão incluídos em dados codificados; e

15 uma etapa de decodificação de dados de deslocamento paralaxe de dados de decodificação de deslocamento paralaxe para cada área indicada pelos dados de divisão de área que são decodificados na etapa de decodificação de dados de divisão de área, em que os dados do deslocamento paralaxe estão incluídos nos dados codificados, e o deslocamento paralaxe
20 é a diferença entre a paralaxe de referência e a paralaxe usada para a compensação paralaxe.

6. Método de codificação de vídeo, de acordo com a reivindicação 5, em que:

25 os dados de divisão de área decodificados na etapa de decodificação de dados de divisão de área indicam um método selecionado dentre uma pluralidade de métodos de divisão de área para realizar divisão de área em blocos retangulares.

7. Método de decodificação de vídeo, de acordo com a reivindicação 5, em que:

30 os dados de deslocamento paralaxe decodificados na etapa de decodificação de dados de deslocamento paralaxe são dados da diferença entre a paralaxe de referência, que é estimada usando uma pluralidade de

imagens de referência sem usar a imagem-alvo a ser decodificada, e a paralaxe usada para a compensação paralaxe.

8. Método de decodificação de vídeo, de acordo com a reivindicação 7, em que:

5 a paralaxe de referência é estimada para cada pixel como uma unidade, pelo uso de imagens de referência.

9. Aparelho para codificação de vídeo para codificar imagens de vídeo como uma imagem de vídeo única pelo uso de compensação paralaxe que realiza previsão pelo uso de paralaxe entre as imagens de vídeo, o aparelho compreendendo:

um dispositivo de ajuste de paralaxe de referência para ajustar paralaxe de referência para uma imagem-alvo a ser codificada, em que a paralaxe de referência é estimada usando uma imagem de referência;

15 um dispositivo de ajuste de divisão de área para ajustar divisão de área em uma estrutura de imagem;

um dispositivo de ajuste de deslocamento paralaxe para ajustar deslocamento paralaxe para cada área dividida ajustada pelo dispositivo de ajuste de divisão de área, em que o deslocamento paralaxe é a diferença entre a paralaxe de referência e a paralaxe usada para a compensação paralaxe;

20 um dispositivo de codificação de dados de divisão de área para codificar dados de divisão de área para indicar a divisão de área que é ajustada pelo dispositivo de ajuste de divisão de área; e

25 um dispositivo de codificação de dados de deslocamento paralaxe para codificar dados de deslocamento paralaxe para indicar o deslocamento paralaxe que é ajustado pelo dispositivo de ajuste de deslocamento paralaxe.

30 10. Aparelho de decodificação de vídeo para decodificar imagens de vídeo como uma imagem de vídeo única pelo uso de compensação paralaxe que realiza previsão pelo uso de paralaxe entre as imagens de vídeo, o aparelho compreendendo:

um dispositivo de ajuste de paralaxe de referência para ajustar

paralaxe de referência para uma imagem-alvo a ser decodificada, em que a paralaxe de referência é estimada usando uma imagem de referência;

um dispositivo de decodificação de dados de divisão de área para decodificar dados de divisão de área para indicar a divisão de área, em
5 que os dados de divisão de área estão incluídos nos dados codificados; e

um dispositivo de decodificação de dados de deslocamento paralaxe para decodificar dados de deslocamento paralaxe para cada área indicada pelos dados de divisão de área que são decodificados pelo dispositivo de decodificação de dados de divisão de área, em que os dados de deslocamento paralaxe estão incluídos nos dados codificados, e o deslocamento paralaxe é a diferença entre a paralaxe de referência e a paralaxe usada para a compensação paralaxe.
10

11. Programa de codificação de vídeo para fazer com que um computador execute o método de codificação de vídeo como definido em
15 qualquer uma das reivindicações 1,2 e 3.

12. Programa de decodificação de vídeo para fazer com que um computador execute o método de decodificação de vídeo como definido em qualquer das reivindicações 5, 6 e 7.

13. Meio de armazenamento legível de computador que armazena um programa de codificação de vídeo para fazer com que um computador execute o método de codificação de vídeo como definido em qualquer das reivindicações 1, 2 e 3.
20

14. Meio de armazenamento legível de computador que armazena um programa de decodificação de vídeo para fazer com que um computador execute o método de decodificação de vídeo de acordo com qualquer das reivindicações 5, 6 e 7.
25

FIG. 1

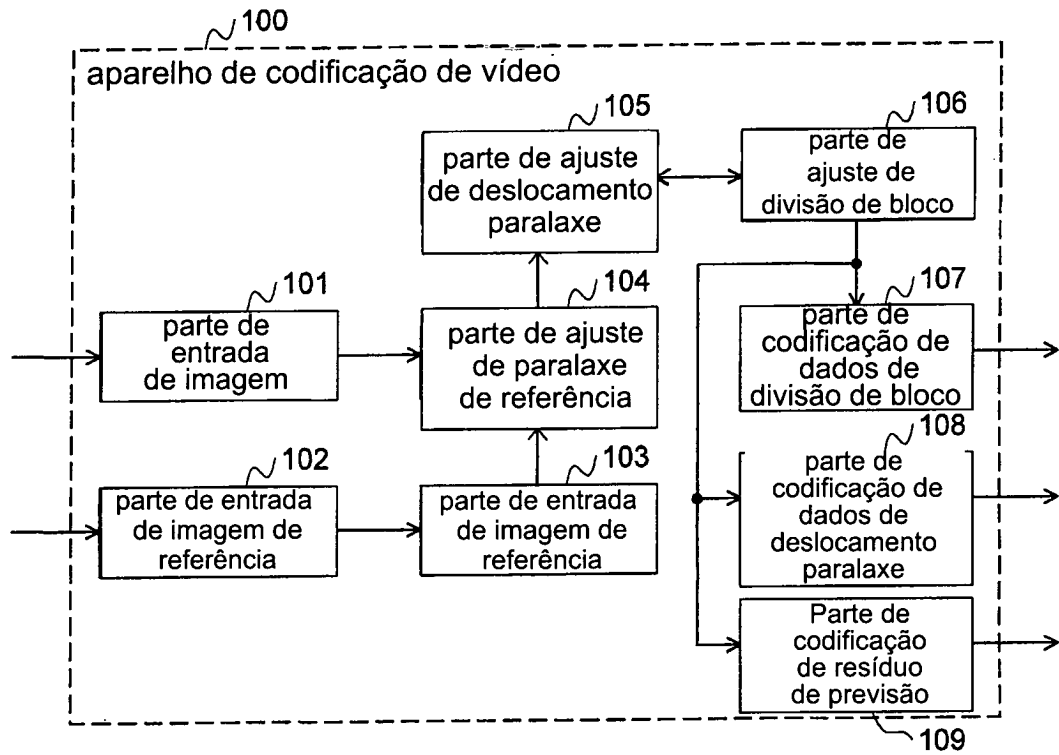


FIG. 2

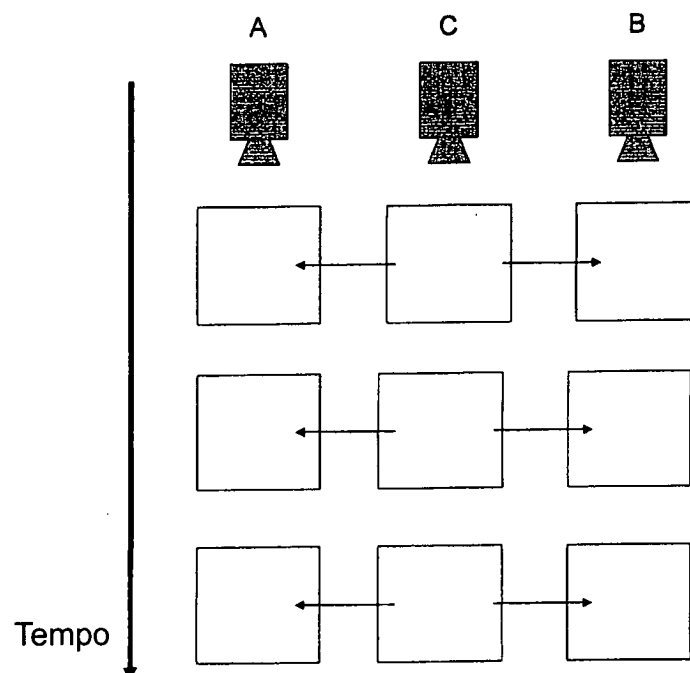


FIG. 3

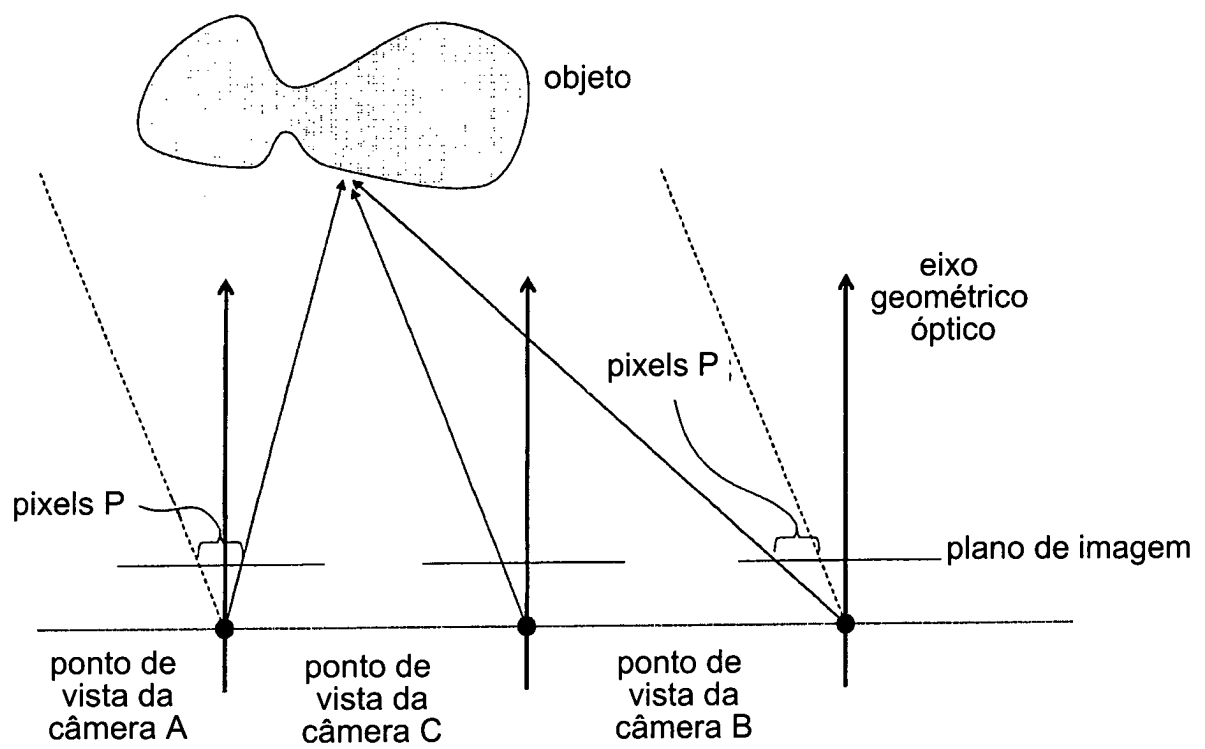


FIG. 4

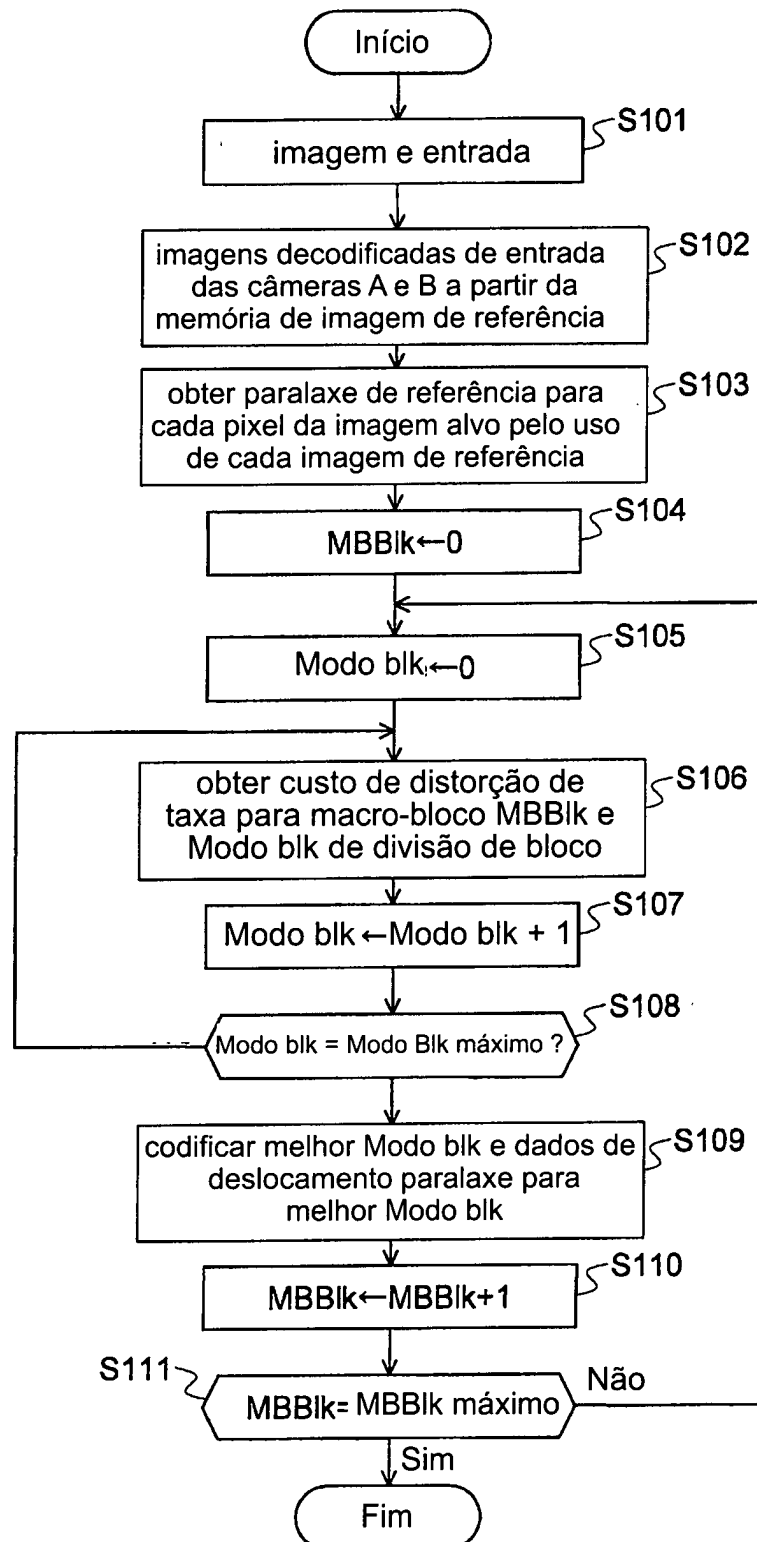


FIG. 5

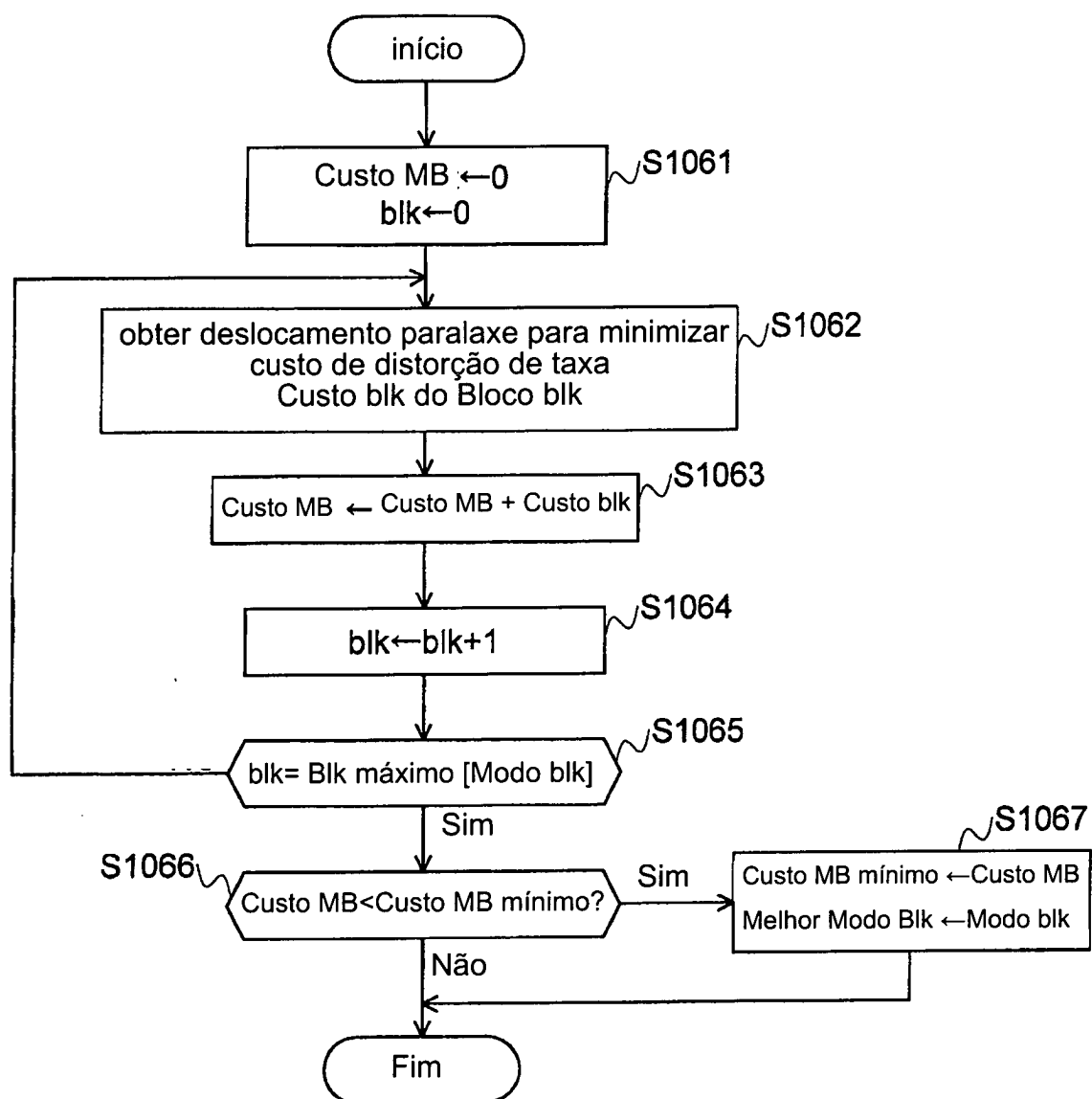


FIG. 6

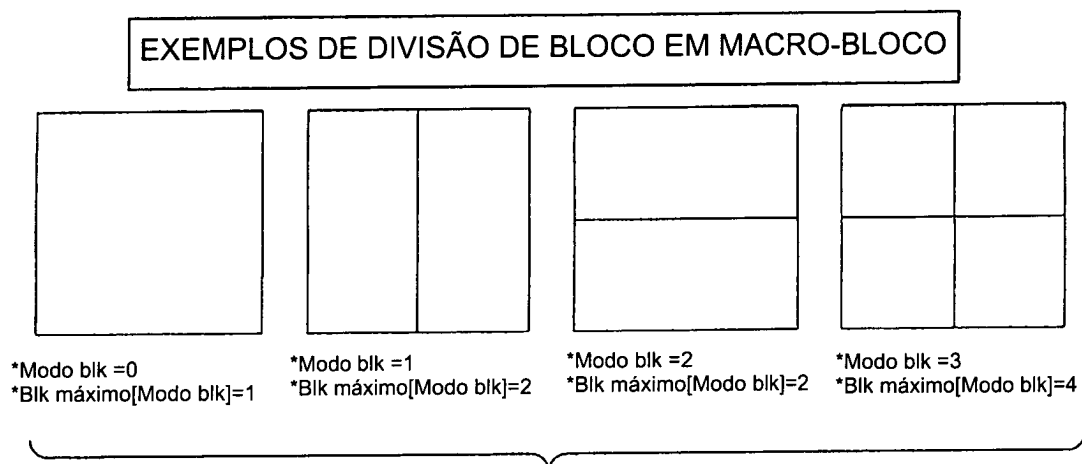


FIG. 7

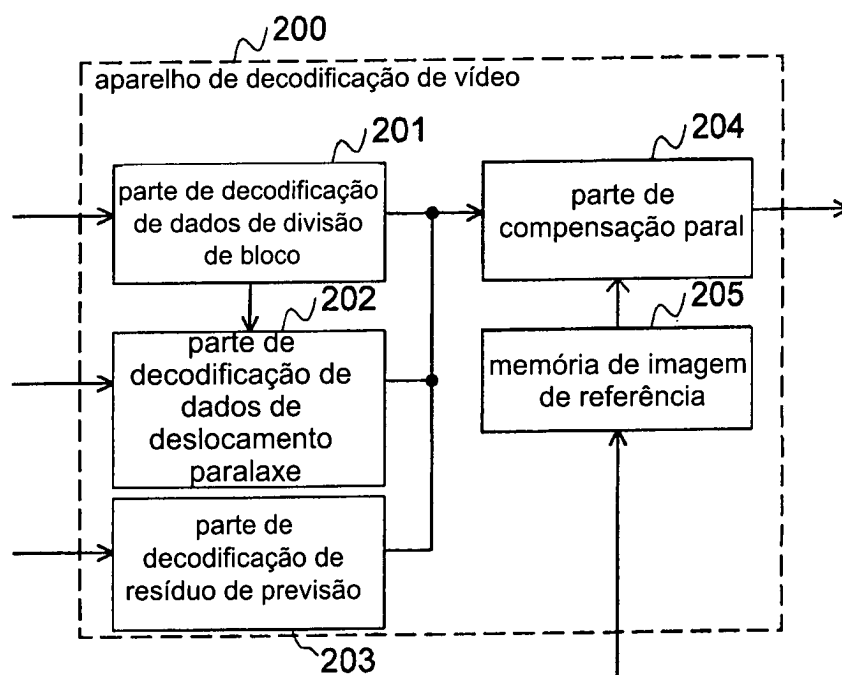


FIG. 8

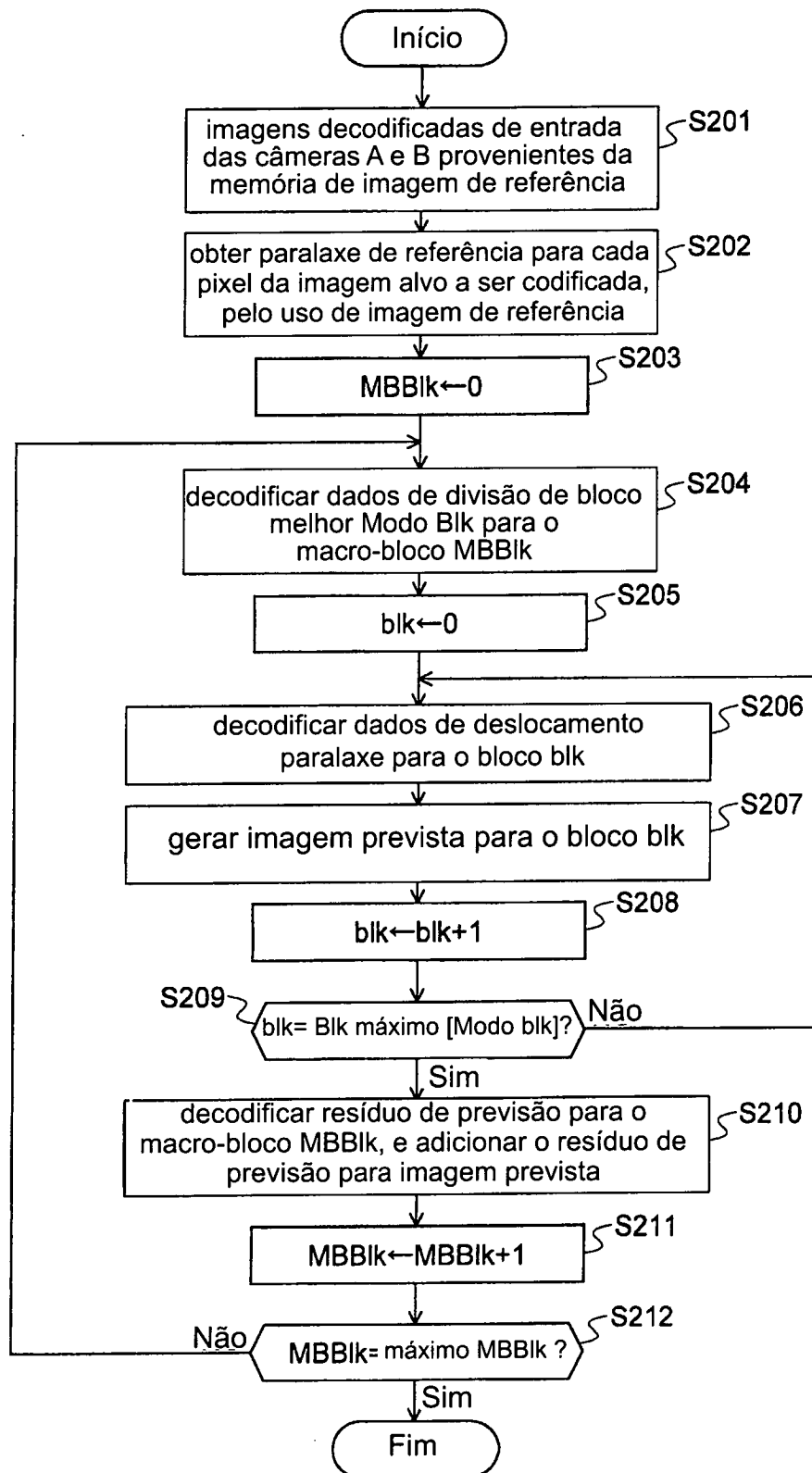


FIG. 9

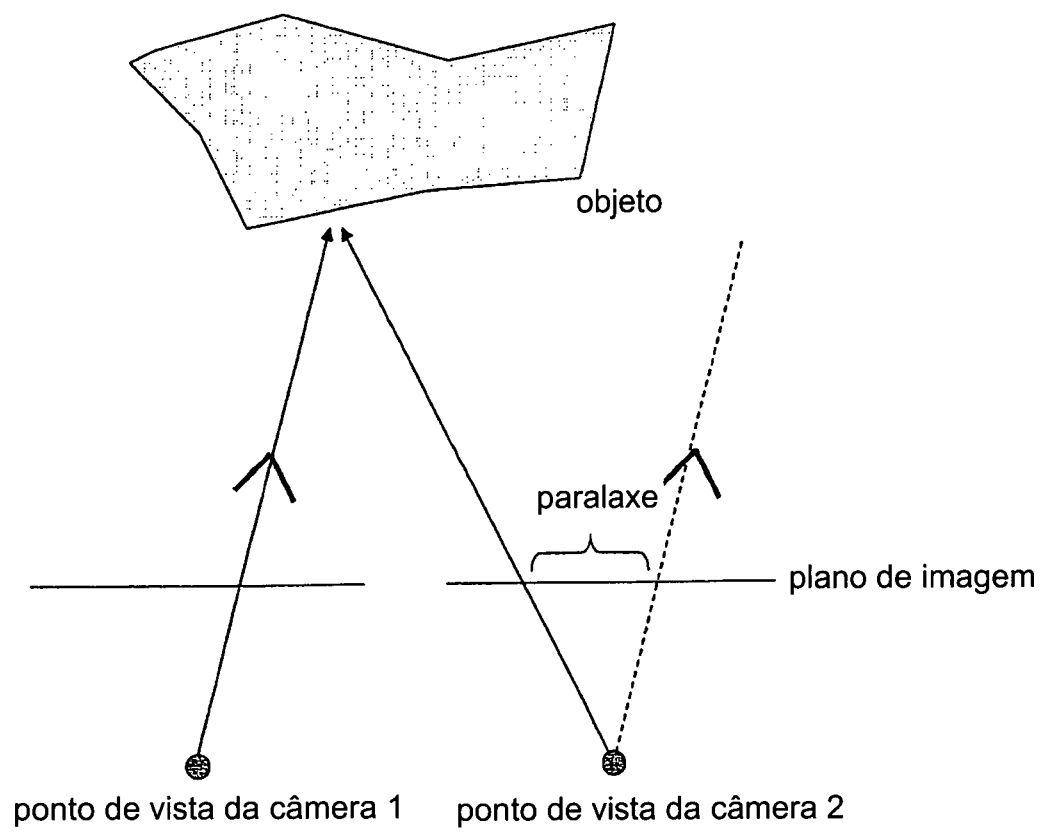


FIG. 10

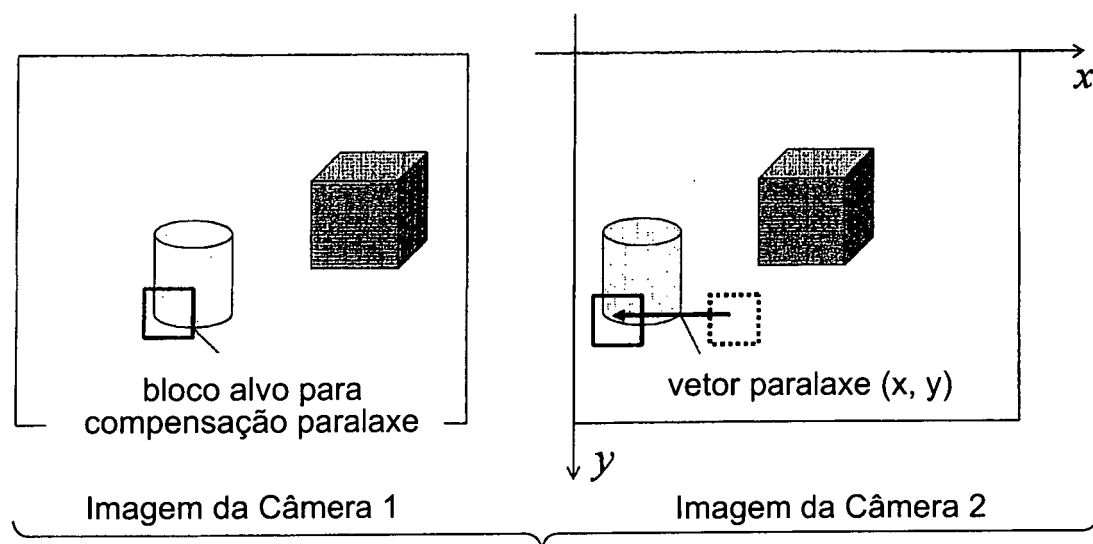


FIG. 11

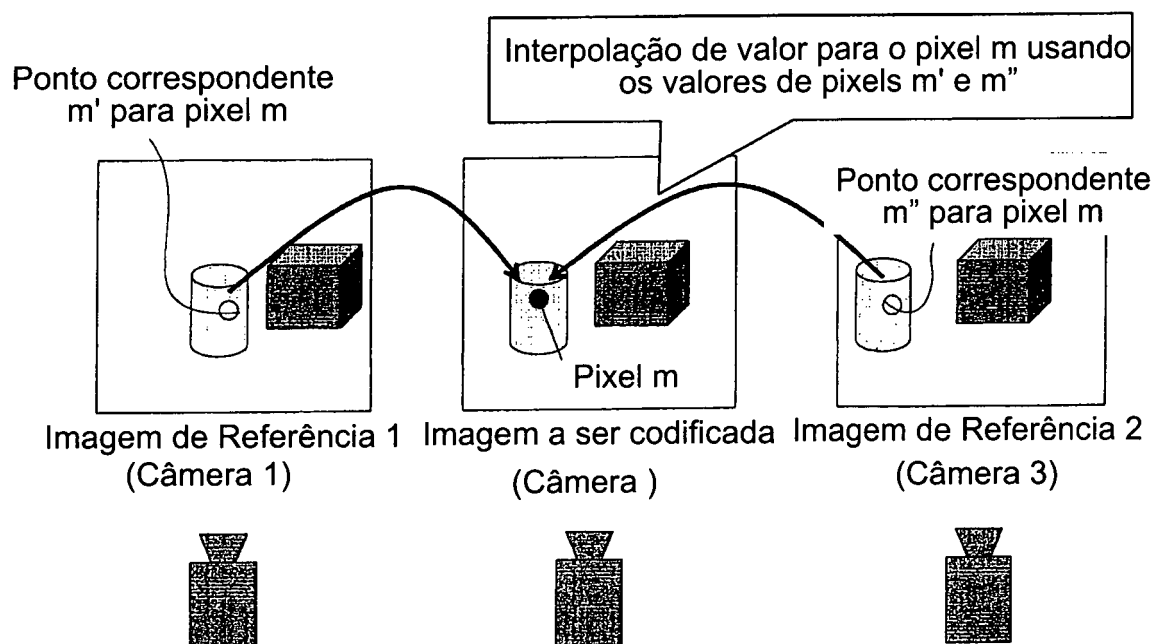
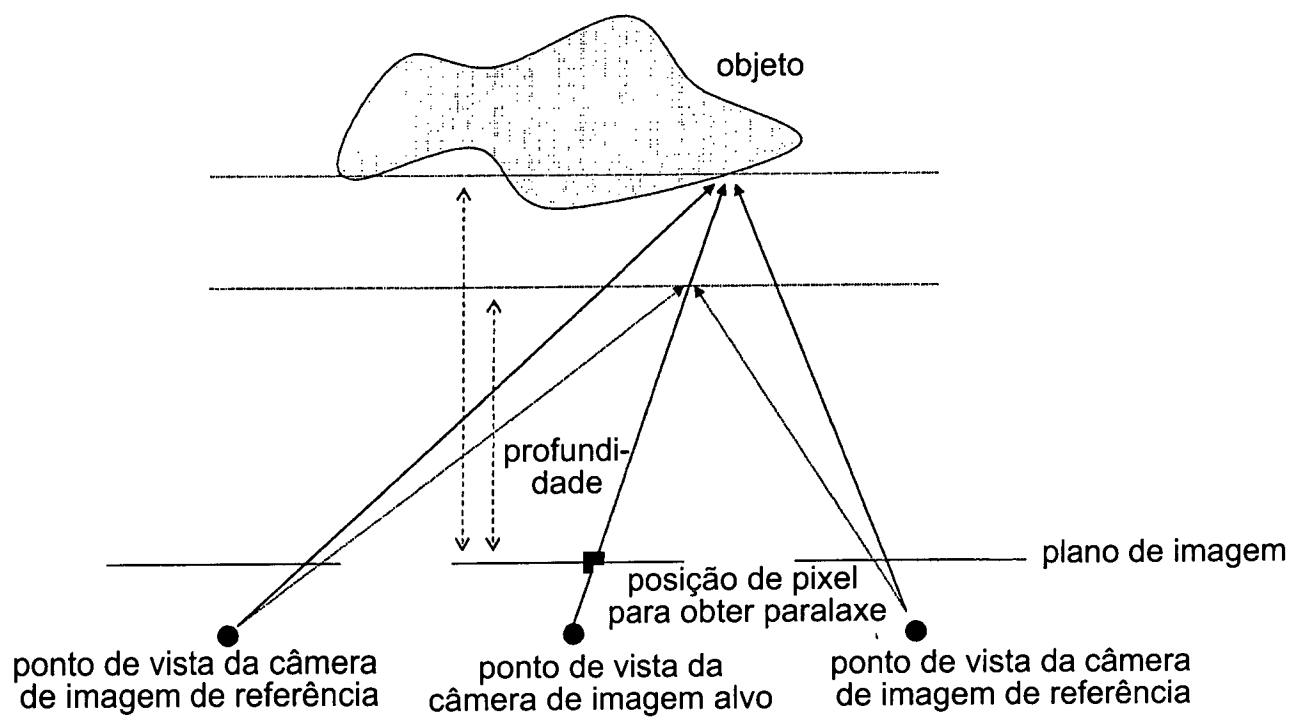


FIG. 12



RESUMO

Patente de Invenção: **"MÉTODO DE CODIFICAÇÃO E DECODIFICAÇÃO DE VÍDEO, APARELHOS PARA OS MESMOS, PROGRAMAS PARA OS MESMOS E MEIOS DE ARMAZENAMENTO PARA ARMAZENAR OS PROGRAMAS"**.

- 5 A presente invenção refere-se ao uso de compensação paralaxe que realiza previsão pelo uso de paralaxe entre as imagens de vídeo, as imagens de vídeo são codificadas como uma imagem de vídeo única. É ajustada paralaxe de referência para uma imagem-alvo a ser codificada, em que
- 10 a paralaxe de referência é estimada usando uma imagem de referência; é ajustada divisão de área em uma estrutura de imagem; é ajustado o deslocamento paralaxe para cada área dividida, em que o deslocamento paralaxe é a diferença entre a paralaxe de diferença e a paralaxe para a compensação paralaxe; são codificados os dados de divisão de área; e são codificados
- 15 os dados para indicar o deslocamento paralaxe. Durante a decodificação, é ajustada a paralaxe de referência para uma imagem-alvo a ser decodificada, em que é estimada usando uma imagem de referência; são decodificados dados para indicar divisão de área, que estão incluídos em dados codificados; e são decodificados os dados de deslocamento paralaxe, que estão
- 20 incluídos nos dados codificados, para cada área indicada pelos dados de divisão de área.