



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

(11) PI 0418350-9 B1



(22) Data do Depósito: 04/11/2004

(45) Data de Concessão: 09/07/2019

(54) Título: MÉTODO E APARELHO PARA CONTROLAR A INSERÇÃO DE CAMPOS OU QUADROS ADICIONAIS EM UMA SEQUÊNCIA DE IMAGENS DE PRIMEIRO FORMATO PARA CONSTRUIR A PARTIR DA MESMA UMA SEQUÊNCIA DE IMAGENS DE SEGUNDO FORMATO

(51) Int.Cl.: H04N 7/01.

(30) Prioridade Unionista: 21/01/2004 EP 04090021.9.

(73) Titular(es): THOMSON LICENSING.

(72) Inventor(es): CARSTEN HERPEL; HEINZ WERNER KEESEN; ANDREJ SCHEWZOW; MARCO WINTER.

(86) Pedido PCT: PCT EP2004012483 de 04/11/2004

(87) Publicação PCT: WO 2005/071956 de 04/08/2005

(85) Data do Início da Fase Nacional: 04/07/2006

(57) Resumo: MÉTODO E APARELHO PARA CONTROLAR A INSERÇÃO DE CAMPOS OU QUADROS ADICIONAIS EM PRIMEIRA SEQUÊNCIA EM UMA PRIMEIRA SEQUÊNCIA DE IMAGEM DE FORMATO PARA PARTIR DA MESMA UMA SEGUNDA SEQUÊNCIA DE IMAGEM DE FORMATO. Os principais sistemas de TV no mundo utilizam varredura entrelaçada quer seja em frequência de campo de 50 Hz ou frequência de campo de 60 Hz. Contudo, os filmes são produzidos em frequência de quadro de 24 Hz e varredura progressiva, cujo formato será utilizado para futuros discos de vídeo digital e serem vendidos em países que utilizam 50 Hz. Em dispositivos de vídeo de 50Hz o conteúdo do disco é apresentado com intensidade de áudio original, mas com quadros ou campos de vídeo repetidos para se obter na média a velocidade de fonte de vídeo original. Contudo, a inserção de quadro ou campo é realizada em um padrão regular, mas de forma adaptativa para reduzir vibração de movimento visível.

“MÉTODO E APARELHO PARA CONTROLAR A INSERÇÃO DE CAMPOS OU QUADROS ADICIONAIS EM UMA SEQUÊNCIA DE IMAGENS DE PRIMEIRO FORMATO PARA CONSTRUIR A PARTIR DA MESMA UMA SEQUÊNCIA DE IMAGENS DE SEGUNDO FORMATO”

5 A invenção se refere a um método e a um aparelho para controlar a inserção de campos ou quadros adicionais em uma sequência de imagens de primeiro formato tendo, por exemplo, 24 quadros progressivos por segundo para construir a partir da mesma uma sequência de imagens de segundo formato tendo, por exemplo, 25 quadros por segundo.

Antecedentes

Os principais sistemas de TV no mundo utilizam varredura entrelaçada e frequência de campo de 50 Hz (por exemplo, na Europa e China para PAL e SECAM) ou frequência de campo de 60 Hz ou quase 60 Hz (por exemplo, nos Estados Unidos e Japão para NTSC), denotados 50i e 60i, respectivamente. Contudo, filmes são produzidos em varredura progressiva e frequência de quadro de 24 Hz, denotada 24p, cujo valor quando expresso em formato de entrelaçamento corresponderia a 48i.

Atualmente, conversão de filme de 24p para exibição entrelaçada de 60 Hz é manipulada por 'conversão (pull-down) 3:2' como mostrado na Figura 2, no qual um campo de conversão 3:2 é inserido por repetição de campo a cada cinco campos. Campos entrelaçados ILF são derivados de quadros de filme original ORGFF. A partir de um primeiro quadro de filme original OFR1 três campos de saída OF1 a OF3 são gerados, e a partir de um terceiro quadro de filme original OFR3 três

campos de saída OF6 a OF8 são gerados. A partir de um segundo quadro de filme original OFR2 dois campos de saída OF4 e OF5 são gerados, e a partir de um quarto quadro de filme original OFR4 dois campos de saída OF9 e OF10 são gerados, e
5 assim por diante.

É desejável que mídia de distribuição tenha uma trilha de vídeo e áudio de formato único que seja reproduzível mundialmente mais propriamente do que a situação atual onde existe pelo menos uma versão de 50 Hz e uma versão de
10 60 Hz de cada título de mídia comercial, por exemplo, DVD. Devido ao fato de muitas fontes consistirem em filme de 24 fps (quadros por segundo), esse formato de 24p é preferivelmente o formato desejado para tais trilhas de vídeo de formato único, cujo formato, portanto, precisa ser adaptado no
15 momento de reprodução para exibição de forma correta em dispositivos de exibição, tanto nos países de 50 Hz como nos de 60 Hz.

As soluções a seguir são conhecidas para conversão de 24p em 25p ou 50i ou, mais genericamente, para conversão
20 de 25 fps:

- Reproduzir em 4,2% mais rápido: isso muda a extensão do conteúdo e exige conversão de altura de áudio em tempo real dispendiosa e, portanto, não é aplicável aos produtos de consumidor. É verdadeiro que difusão de
25 filme atual e DVD aplicam essa solução para vídeo, mas a conversão de altura ou velocidade de áudio exigida já é tratada pelo lado do provedor de conteúdo de tal modo que pelo lado do consumidor nenhuma conversão de altura

de áudio é exigida. Discos de vídeo DVD vendidos em países que utilizam 50 Hz contêm fluxos de dados de áudio que já estão codificados de tal modo que o decodificador do aparelho de DVD automaticamente produz a velocidade ou altura correta do sinal de áudio.

- Aplicar um esquema de duplicação de campo/quadro regular: essa solução conduz à vibração de movimento regular inaceitável e, portanto, não é aplicada na prática.

- Aplicar conversão de taxa de quadro de movimento compensado: essa é uma solução genérica para tais problemas de conversão a qual é muito dispendiosa e, portanto, não é aplicável aos produtos de consumidor.

Invenção

Atualmente, conversão de fluxos de dados de áudio e vídeo de filme de formato de 24p original para vídeo entrelaçado de 50 Hz é realizada mediante reprodução do filme em uma velocidade superior em aproximadamente 4%. Isso significa, contudo, que em países que utilizam 50 Hz o conteúdo artístico do filme (sua duração, altura de vozes) é modificado. Esquemas de repetição de quadro/campo similares a 3:2 pull-down não são usados uma vez que eles mostram artefatos de vibração de movimento inaceitável quando aplicados de uma maneira regular, tal como inserção de um campo extra a cada 12 quadros.

Um problema a ser resolvido pela invenção é o de prover um esquema de inserção de campo ou quadro para conversão a partir do formato de 24p para o formato de 25 fps

de uma maneira aperfeiçoada desse modo minimizando os artefatos de vibração de movimento. Esse problema é resolvido pelo método revelado na reivindicação 1. Um aparelho que utiliza esse método é revelado na reivindicação 2.

5 As características de uma cena de filme atual tal como movimento global, nível de brilho/intensidade e locais de mudança de cena são avaliados para aplicar quadros/campos duplicados ou repetidos em locais subjetivamente não-perturbadores. Em outras palavras, a invenção utiliza infor-
10 mação relativamente facilmente disponível sobre o material original a ser convertido a partir de 24p para 24 fps para inserir de forma adaptativa os quadros/campos repetidos em locais não-equidistantes onde os artefatos de inserção resultantes são mínimos.

15 Vantajosamente, a invenção pode ser usada para todos os problemas de conversão de taxa de quadro onde há uma pequena diferença entre taxa de quadro de origem e taxa de quadro de destino. Se essas taxas de quadro forem muito diferentes, tal como em conversão de 24 fps para 30 fps, há
20 pouca liberdade para se mudar em tempo os quadros ou campos a serem repetidos.

 A invenção facilita a conversão computacionalmente não-dispendiosa a partir de sequências de imagem de formato de 24 fps para 25 fps (valores exemplares) com vibração de
25 movimento minimizada.

 Em princípio, o método inventivo é adequado para controlar a inserção de campos ou quadros adicionais em uma sequência de imagens de primeiro formato para construir a

partir da mesma uma sequência de imagens de segundo formato cuja frequência de quadro é constante e é maior do que aquela da sequência de imagens de primeiro formato, o método incluindo as etapas:

- 5 - determinar locais de campos ou quadros na sequência de imagens de primeiro formato em cujos locais a inserção de um campo ou quadro adicional correspondente causa uma vibração de movimento visível mínima na sequência de imagens de segundo formato;
- 10 - inserir na sequência de imagens de primeiro formato um campo ou um quadro em alguns dos locais em distâncias de inserção de campo ou quadro não-regulares de tal modo que no total a distância média entre quaisquer quadros adjacentes corresponde àquela da sequência de
- 15 imagens de segundo formato;
- apresentar a sequência de imagens de primeiro formato em conjunto com os campos e/ou quadros não-regularmente inseridos no formato da sequência de imagens de segundo formato.

20 Em princípio o aparelho inventivo é adequado para controlar a inserção de campos ou quadros adicionais em uma sequência de imagens de primeiro formato para construir a partir da mesma uma sequência de imagens de segundo formato cuja frequência de quadro é constante e é maior do que aque-

25 la da sequência de imagens de primeiro formato, o aparelho incluindo meios que são adaptados:

 para determinar locais de campos ou quadros na sequência de imagens de primeiro formato em cujos locais a in-

serção de um campo ou quadro adicional correspondente causa uma vibração de movimento visível mínima na sequência de imagens de segundo formato,

e para inserir na sequência de imagens de primeiro formato um campo e um quadro em alguns dos locais em distâncias de inserção de campo ou quadro não-regulares de tal modo que no total a distância média entre quaisquer quadros adjacentes corresponde àquela da sequência de imagens de segundo formato,

e para apresentar a sequência de imagens de primeiro formato em conjunto com os campos e/ou quadros inseridos não-regularmente no formato da sequência de imagens de segundo formato.

Modalidades adicionais vantajosas da invenção são reveladas nas reivindicações dependentes respectivas.

Desenhos

Modalidades exemplares da invenção são descritas com referência aos desenhos anexos, os quais mostram na:

Figura 1. Diagrama simplificado de blocos de um aparelho tocador de disco inventivo;

Figura 2. Aplicação de 3:2 pull-down em uma sequência de imagem original de 24p para prover uma sequência de imagem de 60i;

Figura 3. Padrão regular de quadros repetidos;

Figura 4. Padrão regular de campos repetidos;

Figura 5. Linha de tempo para repetição de quadro regular de acordo com a Figura 3;

Figura 6. Valores de tolerância de vibração de mo-

vimento exemplar de uma sequência de vídeo;

Figura 7. Locais temporais irregulares exemplares para repetição de campo ou quadro e o retardo de apresentação variável resultante;

5 Figura 8. Distância de repetição de quadro ou campo expressa como uma função de retardo de vídeo e tolerância de vibração de movimento;

Figura 9. A função de distância de repetição de quadro ou campo, da Figura 8, pelo que os retardos de vídeo, máximo e mínimo, dependem do grau exigido de sincronia labial;

Figura 10. Quadros de formato de 24 fps incluindo um quadro repetido sem compensação de movimento;

Figura 11. Saída de quadro de formato de 25 fps relacionado à Figura 10;

15 Figura 12. Quadros de formato de 24 fps incluindo um quadro repetido com compensação de movimento;

Figura 13. Saída de quadro de formato de 25 fps relacionado à Figura 12.

Modalidades Exemplares

20 Na Figura 1 uma unidade de disco incluindo um captador e um PEC de estágio de correção de erro lê um sinal de vídeo e áudio codificado de formato de 24p a partir de um disco D. O sinal de saída passa através de um armazenador de trilha e estágio demultiplexador TBM para um decodificador

25 de vídeo VDEC e um decodificador de áudio ADEC, respectivamente. Um controlador CTRL pode controlar PEC, TBM, VDEC e ADEC. Uma interface com o usuário UI e/ou uma interface IF entre um receptor de TV ou um vídeo (não ilustrado) e o apa-

relho tocador de disco é/são utilizadas para comutar a saída do aparelho quer seja para o modo de 24 fps ou para o modo de 25 fps. A interface IF pode verificar automaticamente qual modo ou modos o receptor de TV ou um vídeo pode processar e apresentar. A informação do modo de reprodução é derivada automaticamente a partir dos dados de recurso (isto é, dados sobre qual modo de exibição está disponível no receptor de TV ou no vídeo) recebido pela interface IF que é conectada por fio, mediante ondas de rádio ou opticamente ao receptor de TV ou ao dispositivo de vídeo. Os dados de recurso podem ser recebidos regularmente pela interface IF, ou mediante envio de uma solicitação correspondente para o receptor de TV ou para um dispositivo de vídeo. Como alternativa, a informação de modo de reprodução é introduzida pela interface com o usuário UI mediante exibição de uma solicitação correspondente para um usuário. No caso de saída de 25 fps a partir do decodificador de vídeo VDEC, do controlador CTRL, ou do próprio decodificador de vídeo VDEC, determina a partir das características do sinal de vídeo decodificado em quais locais temporais um campo ou um quadro deve ser repetido pelo decodificador de vídeo. Em algumas modalidades da invenção esses locais temporais também são controlados pelo sinal de áudio ou sinais provenientes do decodificador de áudio ADEC como explicado abaixo.

Em vez de um tocador de disco, a invenção também pode ser usada em outros tipos de dispositivo, por exemplo, um aparelho decodificador digital ou um receptor de TV digital, em cujo caso o dispositivo auxiliar incluindo a unidade

de disco e o armazenador de trilha é substituído por um sintonizador para sinais digitais.

A Figura 3 mostra um padrão regular de quadros repetidos em que um quadro é repetido a cada 24 quadros, isto é, em t_n , t_{n+1} , t_{n+2} , t_{n+3} , etc. segundos, para se obter uma conversão conhecida de 24p para 25 fps.

A Figura 4 mostra um padrão regular de campos repetidos em que um campo é repetido a cada 24 quadros, isto é, em t_n , $t_{n+0,5}$, t_{n+1} , $t_{n+1,5}$, t_{n+2} , etc. segundos, para obter uma conversão conhecida de 24p para 25 fps. Esse tipo de processamento é aplicável se o dispositivo de exibição tiver uma saída entrelaçada. O número de locais no eixo de tempo onde ocorre vibração é dobrado, mas a intensidade de cada "instância de vibração" é dividida em partes iguais em comparação com a repetição de quadro. Campos superiores são derivados a partir da primeira, terceira, quinta, etc. linha do quadro indicado da sequência de origem e os campos inferiores são derivados a partir da segunda, quarta, sexta, etc. linha do quadro indicado da sequência de origem.

A Figura 5 mostra uma linha de tempo para repetição regular de quadro de acordo com a Figura 3, com marcadores nos locais temporais t_n , t_{n+1} , t_{n+2} , t_{n+3} , etc. segundos, onde ocorre a repetição de quadro.

Para realizar a inserção adaptativa inventiva de campos ou quadros repetidos em locais não-equidistantes (ou irregulares) é exigida informação de controle correspondente. Informação de conteúdo e características de sinal de imagem sobre o material de origem torna-se disponível tão

logo a sequência de imagem seja compactada por um esquema tal como Vídeo MPEG-2, Vídeo MPEG-4 ou parte 10 de Vídeo MPEG-4, que supostamente será usado não somente para mídia comercial e de difusão de geração atual tal como DVD, mas
5 também para mídia futura tal como discos baseados em tecnologia de laser azul.

Características ou informação de sinal de imagem que são úteis no contexto desta invenção são:

- 10 - os vetores de movimento gerados e/ou transmitidos,
- informação de mudança de cena gerada por um codificador,
- informação de intensidade ou brilho médio, que pode ser derivada a partir de análise de coeficientes de transformação DC,
15
- informação de intensidade de textura média, que pode ser derivada a partir de análise de coeficientes de transformação AC.

Tais características de sinal de imagem podem ser
20 transferidas a partir do codificador por intermédio de um disco ou por intermédio de difusão para o decodificador como dados de usuário MPEG ou dados privados. Alternativamente, o decodificador de vídeo pode coletar ou calcular e prover tal informação.

25 Para explorar informação de vetor de movimento, o conjunto de vetores de movimento MV para cada quadro é coletado e processado de tal modo que pode ser determinado se um quadro atual tem grandes áreas visivelmente móveis, uma vez

que tais áreas sofrem mais a partir de vibração de movimento quando os quadros ou campos são duplicados. Para determinar a presença de tais áreas a extensão de vetor absoluta média AvgMV_i pode ser calculada para um quadro como uma indicação
5 para um movimento de efeito panorâmico:

$$AvgMV_i = \frac{1}{VX \cdot VY} \sum_{x=0}^{VX-1} \sum_{y=0}^{VY-1} |MV_{x,y}|$$

com "i" denotando o número de quadros, "VX" e "VY" sendo o número de vetores de movimento na direção x (horizontal) e na direção y (vertical) da imagem. Portanto, VX e VY são obtidos tipicamente mediante divisão do tamanho da imagem na
10 direção respectiva pelo tamanho de bloco para estimação de movimento.

Se os vetores de movimento dentro de um quadro apontar para diferentes quadros de referência em diferente distância temporal para o quadro atual, um fator de normalização RDist_{x,y} para essa distância é exigido adicionalmente:
15

$$AvgMV_i = \frac{1}{VX \cdot VY} \sum_{x=0}^{VX-1} \sum_{y=0}^{VY-1} \frac{|MV_{x,y}|}{RDist_{x,y}} .$$

Em uma outra modalidade utilizando processamento mais complexo, uma segmentação de movimento de cada imagem é calculada, isto é, um ou mais agrupamentos de blocos adjacentes tendo vetores de movimento com comprimento e direção, simi-

lares, são determinados, para detectar múltiplas áreas móveis grandes o suficiente com diferentes direções de movimento. Em tal caso o vetor de movimento médio pode ser calculado, por exemplo, por:

$$AvgMV_i = \frac{\sum_{c=1}^{nClusters} AvgMV_c \cdot ClusterSize_c}{\sum_{c=1}^{nClusters} ClusterSize_c},$$

5 em que AvgMVC é o comprimento de vetor de movimento médio para o agrupamento identificado "c".

Vantajosamente essa abordagem elimina o efeito de vetores de movimento para objetos pequenos móveis aleatoriamente dentro de uma imagem que não são membros de qualquer movimento de agrupamento de bloco identificado e que não contribuem significativamente para visibilidade de vibração de movimento.

O processamento pode considerar como fatores de ponderação para AvgMVi se as áreas móveis são intensamente texturizadas ou se têm bordas acentuadas, uma vez que isso também aumenta a visibilidade de vibração de movimento. Informação sobre intensidade de textura pode ser derivada mais convenientemente a partir de uma análise estatística de coeficientes de transformação AC transmitidos ou recebidos ou reproduzidos para o erro de predição. Em princípio, intensidade de textura deve ser determinada a partir de análise de um bloco de imagem original, contudo, em muitos casos tais blocos intensamente texturizados após codificação utilizando predição de movimento compensado também terão mais energia

de erro de predição em seus coeficientes AC do que os blocos menos texturizados. A tolerância de vibração de movimento MJT em um local temporal específico da sequência de vídeo pode, portanto, ser expressa como:

5 MJT = f(AvgMV, intensidade de textura, intensidade de borda) com as seguintes características gerais:

- Dados os valores fixos de intensidade de textura e intensidade de borda, MJT é proporcional a $1/\text{AvgMV}$;
- 10 - Dados os valores fixos de AvgMV e intensidade de borda, MJT é proporcional a $1/(\text{intensidade de textura})$;
- Dados os valores fixos de AvgMV e intensidade de textura, MJT é proporcional a $1/(\text{intensidade de}$
- 15 $\text{borda})$.

A Figura 6 mostra valores de tolerância de vibração de movimento exemplares MJT(t) em relação a uma sequência de origem.

Preferivelmente o tamanho atual do valor de tolerância de vibração de movimento influencia a distribuição, como ilustrado na Figura 7a, de quadros ou campos repetidos inseridos na sequência resultante de 25 fps, isto é, a distância de repetição de quadro ou campo FRD. Inserção antecipada ou retardada de quadros repetidos causa um retardo negativo ou positivo da trilha de áudio em relação à trilha de vídeo como indicado na Figura 7b, isto é um retardo de apresentação variável para vídeo. Um retardo de vídeo máximo tolerável em relação ao áudio em ambas as direções é conside-

rado ao se aplicar o mapeamento a partir da tolerância de vibração de movimento MJT à distância de repetição de quadro ou campo FRD.

Uma solução possível para esse problema de controle é ilustrada na Figura 8. A distância de repetição de quadro ou campo FRD é expressa como uma função do retardo de vídeo VD e da tolerância de vibração de movimento MJT:

$$FRD = f(VD, MJT),$$

com as seguintes características gerais:

- Dado um valor fixo de VD, FRD é proporcional a $1/MJT$;
- Dado um valor fixo de MJT, FRD é proporcional a $1/VD$.

Essa relação pode ser expressa em uma característica de $FRD = f(VD)$ que muda dependendo do valor de tolerância de vibração de movimento, como é o caso na Figura 8, favorecendo aberturas maiores do que as consideradas ótimas entre quadros repetidos inseridos no caso de baixa tolerância de vibração de movimento (por exemplo, alto grau de movimento) e favorecendo aberturas inferiores àquelas consideradas ótimas no caso de tolerância de vibração de movimento elevada (por exemplo, grau de movimento inferior à média). A distância de repetição de campo ou quadro, ótima, é mostrada como FRD_{opt} . O retardo de vídeo máximo permissível é mostrado como VD_{max} . O retardo de vídeo máximo permissível na direção negativa é mostrado como VD_{min} .

Uma vez que um efeito de congelamento de quadro curto em locais de mudança de cena não é considerado como

irritante, informação de mudança de cena gerada por um codificador de vídeo (ou por um decodificador de vídeo) pode ser usada para inserir um ou mais campos ou quadros repetidos em tais locais, o número de repetições dependendo do grau atual
5 de retardo de vídeo. Pela mesma razão, campos ou quadros repetidos podem ser inseridos após uma sequência de desvanecimento/preto, uma sequência de desvanecimento/branco ou um desvanecimento para qualquer cor. Todos os tais locais singulares têm um valor MJT muito elevado.

10 Notavelmente os quadros repetidos poderiam ser usados em tais locais mesmo se em outros campos de conteúdo de imagem somente seriam repetidos para reduzir a intensidade de vibração de movimento em locais individuais. Geralmente, quadros repetidos ou campos repetidos podem coexistir em
15 uma sequência de imagem convertida.

Limites de retardo tipicamente aceitos para sincronia labial percebida precisam ser observados somente se pelo menos um orador estiver efetivamente visível dentro da cena. Portanto o retardo entre apresentação de áudio e vídeo
20 pode se tornar maior do que os limites mencionados acima enquanto nenhum orador estiver visível. Esse é tipicamente o caso durante cenas de movimento rápido. Portanto um controle adicional pode ser realizado como mostrado na Figura 9, em que os limites de retardo de vídeo VD_{min} e VD_{max} são comutados
25 ou suavemente mudados entre:

- valores aceitáveis de sincronia labial $VD_{minLipSync}$ e $VD_{maxLipSync}$ se picos curtos de som ou fala (que são causados por eventos especiais como uma porta que bate) forem detec-

tados e se uma cena de movimento lento ou estática for detectada;

- Valores VD maiores VD_{\min} e VD_{\max} caso contrário.

Uma detecção de fala pode ser derivada, por exemplo, no caso do áudio de multicanal mais usado mediante avaliação do canal central em relação aos canais esquerdo e direito, uma vez que a fala nos filmes é na maior parte codificada para o canal central. Se o canal central mostrar uma distribuição de energia em rajada com o passar do tempo isto é significativamente diferente da distribuição de energia nos canais esquerdo e direito, então, é grande a probabilidade de fala estar presente.

Todos os controles acima para determinar de forma adaptativa a distância de repetição de quadro local funcionam para uma passagem única através da sequência de vídeo. Contudo, o controle inventivo se beneficia de um processamento de codificação de duas passagens como realizado em muitos codificadores MPEG-2 profissionais. Nesse caso a primeira passagem é utilizada para coletar a curva de intensidade de movimento, locais de corte de cena e contagem, número, local e extensão das cenas que exigem sincronia labial justa, quadros pretos, etc. Então um esquema de controle modificado pode ser aplicado o qual não somente considera a informação disponível para o quadro atualmente processado e seu antecessor, como também uma proximidade de quadros passados e futuros:

$$FRD(i) = f(VD, MJT(i-k) \dots MJT(i+k)), \quad (5)$$

em que "i" denota o número de quadro atual e "k" denota um

número corrente referente a quadros adjacentes. Uma característica geral de cada tal função é que FRD aumenta se $MJT(i)$ for menor do que os valores MJT circundantes e diminui se $MJT(i)$ for maior do que os valores MJT circundantes. Características relacionadas de sinal de imagem podem ser transferidas como dados de usuário MPEG ou dados privados a partir do codificador por intermédio de um disco ou por intermédio de sinal de difusão para o decodificador.

Em uma outra modalidade da invenção, sob circunstâncias específicas a interpolação de movimento compensado de quadros, mais propriamente do que a repetição de quadros pode ser aplicada sem custo computacional. Tal interpolação de movimento compensado pode fazer uso dos vetores de movimento transmitidos para o quadro atual. Em geral, esses vetores de movimento não são adequados para interpolação de quadro de movimento compensado uma vez que eles são aperfeiçoados para ganho ótimo de predição mais propriamente do que indicar o movimento verdadeiro de uma cena. Contudo, se uma análise de decodificador de vetores de movimento, recebidos, mostrar que ocorre um efeito panorâmico homogêneo da cena, um quadro altamente exato pode ser interpolado entre o quadro atual e o quadro anterior. Efeito panorâmico significa que todos os vetores de movimento dentro de um quadro são idênticos ou quase idênticos em extensão e orientação. Portanto um quadro interpolado pode ser gerado mediante translação do quadro anterior em metade da distância indicada pelo vetor de movimento médio para o quadro atual. Supõe-se que o quadro anterior seja o quadro de referência para a

predição de movimento compensado do quadro atual e que o quadro interpolado esteja posicionado equidistantemente entre o quadro anterior e o quadro atual. Se o quadro de predição não for o quadro anterior, redimensionamento adequado do vetor de movimento médio deve ser aplicado.

As considerações correspondentes são verdadeiras para o caso onde um zoom pode ser determinado a partir dos vetores de movimento recebidos. Um zoom é caracterizado por vetores de movimento zero no centro de zoom e comprimento crescente de vetores de movimento centro-(oposto)-dirigido em torno desse centro de zoom, o comprimento de vetor de movimento aumentando em relação à distância a partir do centro de zoom.

Vantajosamente esse tipo de interpolação de movimento compensado produz um comportamento de vibração de movimento aperfeiçoado em comparação com a repetição de um quadro, como ilustrado nas Figuras 10 a 13. A Figura 10 no formato de 24 fps e a Figura 11 após conversão de formato de 25 fps mostra quadros (indicados como barras verticais) com uma trajetória de movimento para um objeto verticalmente móvel e uma instância de repetição de quadro, que resulta em um "quadro congelado". A Figura 12 mostra inserção de quadro de movimento interpolado que quando apresentada na taxa de quadros alvo aumentada de 25 fps, como ilustrado na Figura 13, conduz a um "quadro lentamente móvel" mais propriamente do que um "quadro congelado".

Os controles revelados acima para repetição de quadro e/ou campo e interpolação para conversão de taxa de

quadro podem ser aplicados tanto no codificador como no decodificador de um sistema de compactação MPEG-2 (ou similar) uma vez que a maior parte da informação relacionada está disponível em ambos os lados, possivelmente exceto indicação
5 de mudança de cena confiável.

Contudo, para explorar o conhecimento das características de sequência de imagem superior do codificador, os locais para os campos ou quadros serem repetidos ou interpolados podem ser transportados no sinal de vídeo compactado
10 de 24 fps (MPEG-2 ou de outro modo). Indicadores para indicar a ordem temporal dos campos (top_field_first) e repetição do primeiro campo para exibição (repeat_first_field) já existem na sintaxe de MPEG-2. Se for exigido sinalizar o padrão para conversão de 24 fps para 30 fps e de 24 fps para
15 25 fps para o mesmo sinal de vídeo, uma das duas séries de indicadores pode ser transportada em um campo de dados de usuário adequado para cada imagem.

Os valores de 24 fps e 25 fps e os outros números mencionados acima são valores exemplares os quais podem ser
20 adaptados correspondentemente a outras aplicações da invenção.

A invenção pode ser aplicada em:

- mídia comercial (DVD, discos de laser azul, etc.),
- mídia transferida inclusive vídeo sob demanda,
25 quase-vídeo sob demanda, etc.,
- mídia de difusão.

A invenção pode ser aplicada em um tocador de disco ótico ou em um gravador de disco ótico, ou em um gravador

de disco rígido, por exemplo, um gravador de HDD ou um PC, ou em um aparelho decodificador, ou em um receptor de TV.

REIVINDICAÇÕES

1. Método para controlar (CTRL, VDEC) a inserção de campos ou quadros adicionais em uma sequência de imagens de primeiro formato (24p) tendo uma primeira frequência de quadro a fim de construir a partir da mesma uma sequência de imagens de segundo formato (25 fps) cuja frequência de quadro é constante e maior do que a primeira frequência de quadro, o método incluindo as etapas de:

determinar (CTRL, VDEC, ADEC) locais de campos ou quadros na sequência de imagens de primeiro formato em cujos locais a inserção de um campo ou quadro adicional correspondente causa uma vibração de movimento visível mínima (MJT) na sequência de imagens de segundo formato;

inserir (CTRL, VDEC) na sequência de imagens de primeiro formato um campo ou um quadro em alguns dos ditos locais em distâncias de inserção de campo ou quadro não-regulares (FRD) de tal modo que no total a distância média entre quaisquer quadros adjacentes corresponde àquela da sequência de imagens de segundo formato;

apresentar a sequência de imagens de primeiro formato em conjunto com os campos e/ou quadros inseridos não regularmente no formato da sequência de imagens de segundo formato,

CARACTERIZADO pelo fato de que:

uma quantidade máxima de retardo de apresentação (VD_{max} , VD_{min}) da faixa de vídeo em relação à faixa de áudio é determinada, e

os locais de inserção de campo ou quadro na sequência de imagens de primeiro formato são controlados de

modo que, a fim de se obter sincronia labial percebida, na sequência de imagens de segundo formato, uma menor quantidade máxima de retardo de apresentação ($VD_{\max\text{LipSync}}$, $VD_{\min\text{LipSync}}$) da faixa de vídeo em relação à faixa de áudio, causada por irregularidade de inserção, é usada caso sejam cumpridas as seguintes condições:

ser detectada uma cena de movimento lento ou estática, e

serem detectados picos curtos de som ou fala nas informações de áudio atribuídas à sequência de imagens de primeiro formato.

2. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que fala na informação de áudio atribuída à sequência de imagens de primeiro formato é detectada mediante avaliação, em áudio de multicanal, se o canal central em relação aos canais esquerdo e direito mostra uma distribuição de energia em rajada em relação ao ponto que é diferente a partir da distribuição de energia nos canais esquerdo e direito.

3. Método, de acordo com a reivindicação 1 ou 2, **CARACTERIZADO** pelo fato de que a sequência de imagens de primeiro formato (24p) é armazenada ou gravada em um meio físico de armazenamento (D), ou é difundida ou transferida como um sinal de TV digital.

4. Método, de acordo com qualquer uma das reivindicações 1 a 3, **CARACTERIZADO** pelo fato de que os locais de inserção de campo ou quadro na sequência de imagens de primeiro formato são quadros ou campos que não contêm grandes áreas de conteúdo de imagem móvel, o movimento sendo

determinado mediante avaliação de vetores de movimento.

5. Método, de acordo com qualquer uma das reivindicações 1 a 4, **CARACTERIZADO** pelo fato de que os locais de inserção de campo ou quadro na sequência de imagens de primeiro formato são quadros ou campos nos quais ocorrem mudanças de cena ou um desvanecimento para preto ou um desvanecimento para branco ou um desvanecimento para qualquer cor.

6. Método, de acordo com qualquer uma das reivindicações 1 a 5, **CARACTERIZADO** pelo fato de que os campos ou quadros inseridos são de movimento compensado antes de serem produzidos na sequência de imagens de segundo formato.

7. Método, de acordo com qualquer uma das reivindicações 1 a 6, **CARACTERIZADO** pelo fato de que a sequência de imagens de primeiro formato é uma sequência de imagem MPEG-2 e em que a inserção (CTRL, VDEC) de campos ou quadros na sequência de imagens de primeiro formato é controlada mediante avaliação de indicadores para indicar ordem temporal de campos ou para indicar repetição do primeiro campo para exibição, cujos indicadores são transportados na sequência de imagens de primeiro formato em um campo de dados de usuário para cada imagem.

8. Método, de acordo com qualquer uma das reivindicações 1 a 7, **CARACTERIZADO** pelo fato de que a primeira frequência de quadro é 24 Hz e a frequência de quadro da sequência de imagens de segundo formato é 25Hz.

9. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que os locais de sinalização de

indicadores para campo ou quadro para serem repetidos ou interpolados são transportados em um campo de dados de usuário da sequência de imagens de primeiro formato (24p).

10. Aparelho para controlar (CTRL, VDEC) a inserção de campos ou quadros adicionais em uma sequência de imagens de primeiro formato (24p) tendo a primeira frequência de quadro para construir a partir da mesma uma sequência de imagens de segundo formato (25 fps), a frequência de quadro da qual é constante e é maior do que a primeira sequência de quadro, o aparelho incluindo meios (CTRL, VDEC, ADEC) que são adaptados para determinar locais de campos ou quadros na sequência de imagens de primeiro formato em cujos locais a inserção de um campo ou quadro adicional correspondente causa uma vibração mínima de movimento visível (MJT) na sequência de imagens de segundo formato; e

para inserir na sequência de imagens de primeiro formato um campo ou um quadro em alguns dos ditos locais em distâncias de inserção de campo ou quadro não-regulares (FRD) de modo que no total a distância média entre quaisquer quadros adjacentes corresponde àquela da sequência de imagens de segundo formato; e

para apresentar a sequência de imagens de primeiro formato em conjunto com os campos e/ou quadros não regularmente inseridos no formato da sequência de imagens de segundo formato,

CARACTERIZADO pelo fato de que:

uma quantidade máxima de retardo de apresentação (VD_{\max} , VD_{\min}) da faixa de vídeo em relação à faixa de áudio é determinada, e

os locais de inserção de campo ou quadro na sequência de imagens de primeiro formato são controlados de modo que, a fim de se obter sincronia labial percebida, na sequência de imagens de segundo formato, uma menor quantidade máxima de retardo de apresentação ($VD_{\max\text{LipSync}}$, $VD_{\min\text{LipSync}}$) da faixa de vídeo em relação à faixa de áudio, causada por irregularidade de inserção, é usada caso sejam cumpridas as seguintes condições:

ser detectada uma cena de movimento lento ou estática, e

serem detectados picos curtos de som ou fala nas informações de áudio atribuídas à sequência de imagens de primeiro formato.

11. Aparelho, de acordo com a reivindicação 10, **CARACTERIZADO** por ser um tocador de disco ótico, ou um gravador de disco ótico, ou um gravador de disco rígido.

12. Aparelho, de acordo com a reivindicação 10 ou 11, **CARACTERIZADO** por ser um tocador de disco ótico ou um gravador de disco ótico, ou um gravador de disco rígido ou um aparelho decodificador, em que o aparelho produz a sequência de imagens de primeiro formato (24p) original ou a sequência de imagens de segundo formato (25 fps), cuja escolha é controlada pela informação de modo de reprodução recebida automaticamente a partir de uma interface (IF) que é conectada a um dispositivo incluindo um dispositivo de exibição, ou recebida a partir de uma interface com o usuário (UI).

13. Aparelho, de acordo com qualquer uma das reivindicações 10 a 12, **CARACTERIZADO** pelo fato de que fala

na informação de áudio atribuída à sequência de imagens de primeiro formato é detectada mediante avaliação, em áudio de multicanal, se o canal central em relação aos canais esquerdo e direito mostra uma distribuição de energia em rajada em relação ao ponto que é diferente a partir da distribuição de energia nos canais esquerdo e direito.

14. Aparelho, de acordo com qualquer uma das reivindicações 10 a 13, **CARACTERIZADO** pelo fato de que a sequência de imagens de primeiro formato (24p) é armazenada ou gravada em um meio físico de armazenamento (D), ou é difundida ou transferida como um sinal de TV digital.

15. Aparelho, de acordo com qualquer uma das reivindicações 10 a 14, **CARACTERIZADO** pelo fato de que os locais de inserção de campo ou quadro na sequência de imagens de primeiro formato são quadros ou campos que não contêm grandes áreas de conteúdo de imagem móvel, o movimento sendo determinado mediante avaliação de vetores de movimento.

16. Aparelho, de acordo com qualquer uma das reivindicações 10 a 15, **CARACTERIZADO** pelo fato de que os locais de inserção de campo ou quadro na sequência de imagens de primeiro formato são quadros ou campos nos quais ocorrem mudanças de cena ou um desvanecimento para preto ou um desvanecimento para branco ou um desvanecimento para qualquer cor.

17. Aparelho, de acordo com qualquer uma das reivindicações 10 a 16, **CARACTERIZADO** pelo fato de que os campos ou quadros inseridos são de movimento compensado antes de serem produzidos na sequência de imagens de segundo formato.

18. Aparelho, de acordo com qualquer uma das reivindicações 10 a 17, **CARACTERIZADO** pelo fato de que a sequência de imagens de primeiro formato é uma sequência de imagem MPEG-2 e em que a inserção (CTRL, VDEC) de campos ou quadros na sequência de imagens de primeiro formato é controlada mediante avaliação de indicadores para indicar ordem temporal de campos ou para indicar repetição do primeiro campo para exibição, cujos indicadores são transportados na sequência de imagens de primeiro formato em um campo de dados de usuário para cada imagem.

19. Aparelho, de acordo com qualquer uma das reivindicações 10 a 18, **CARACTERIZADO** pelo fato de que a primeira frequência de quadro é 24 Hz e a frequência de quadro da sequência de imagens de segundo formato é 25Hz.

P 104 18350

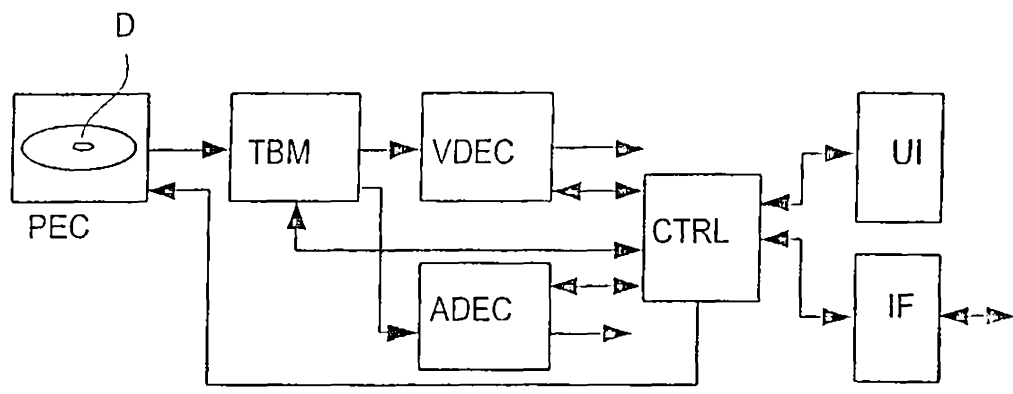


Fig.1

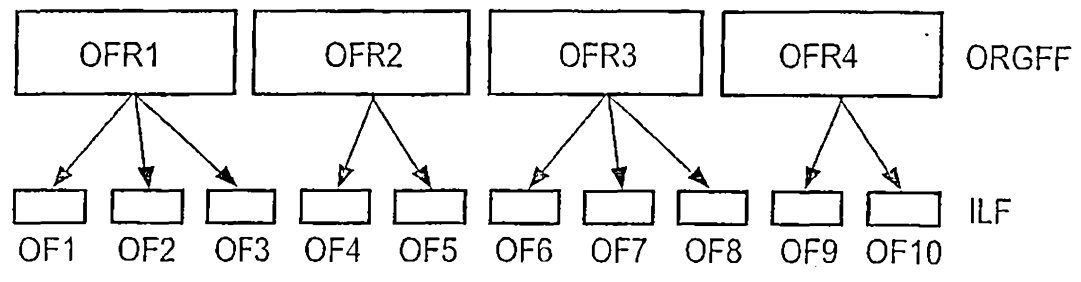
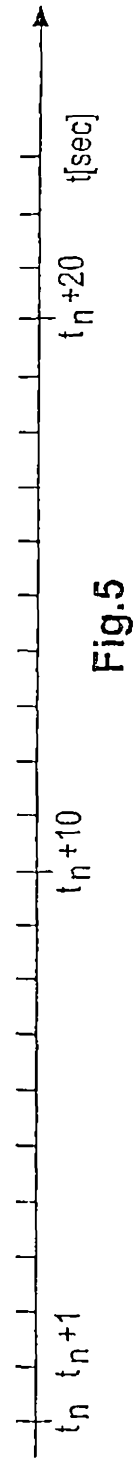
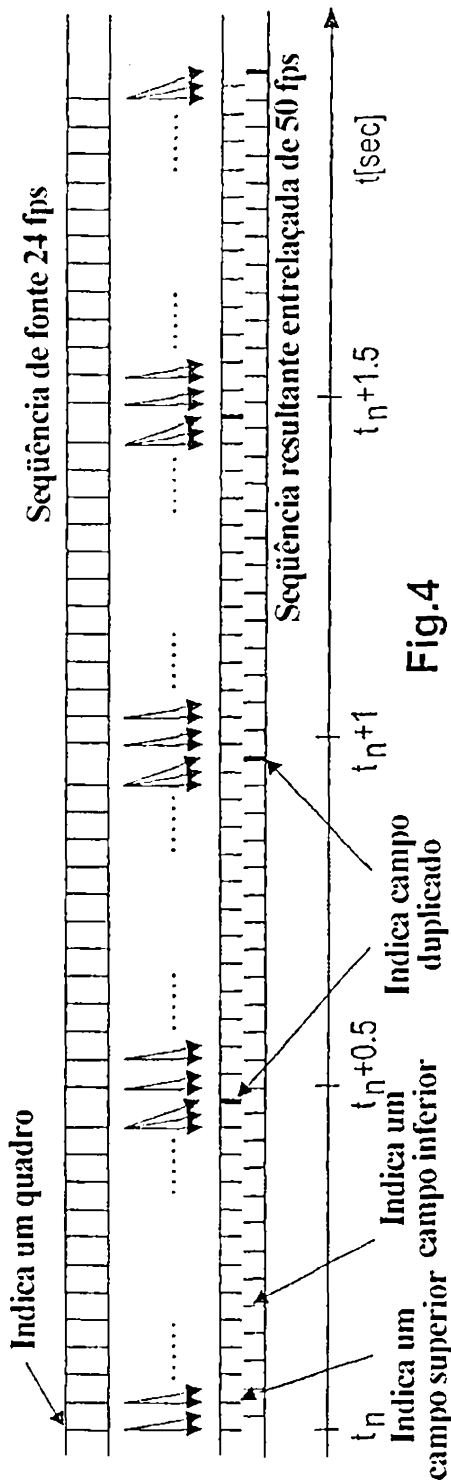
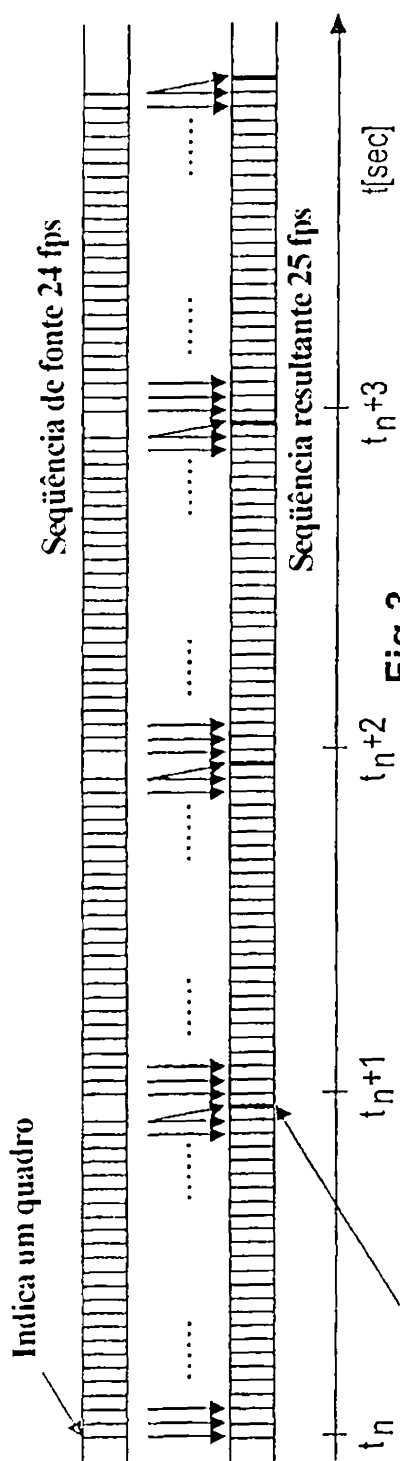
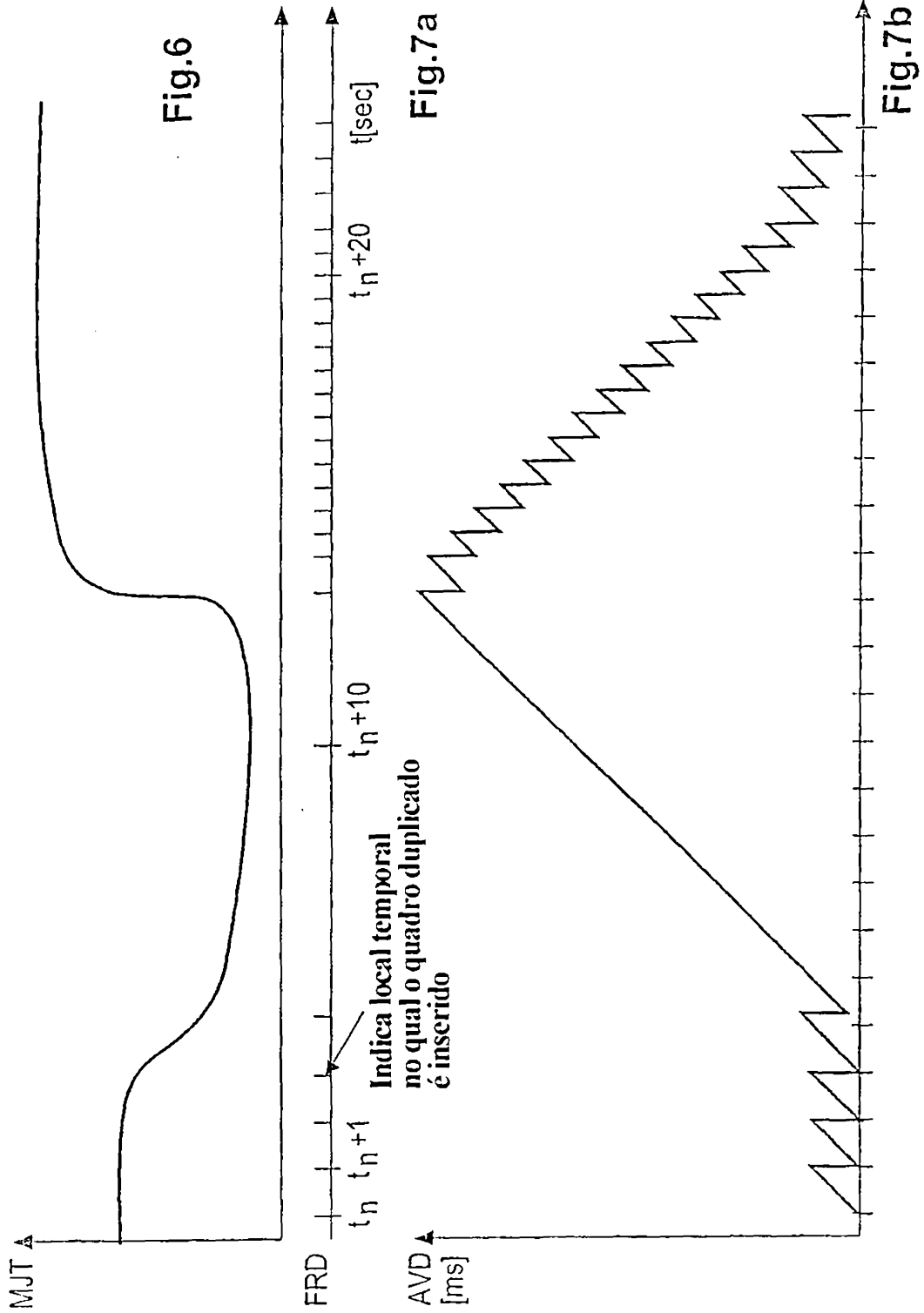


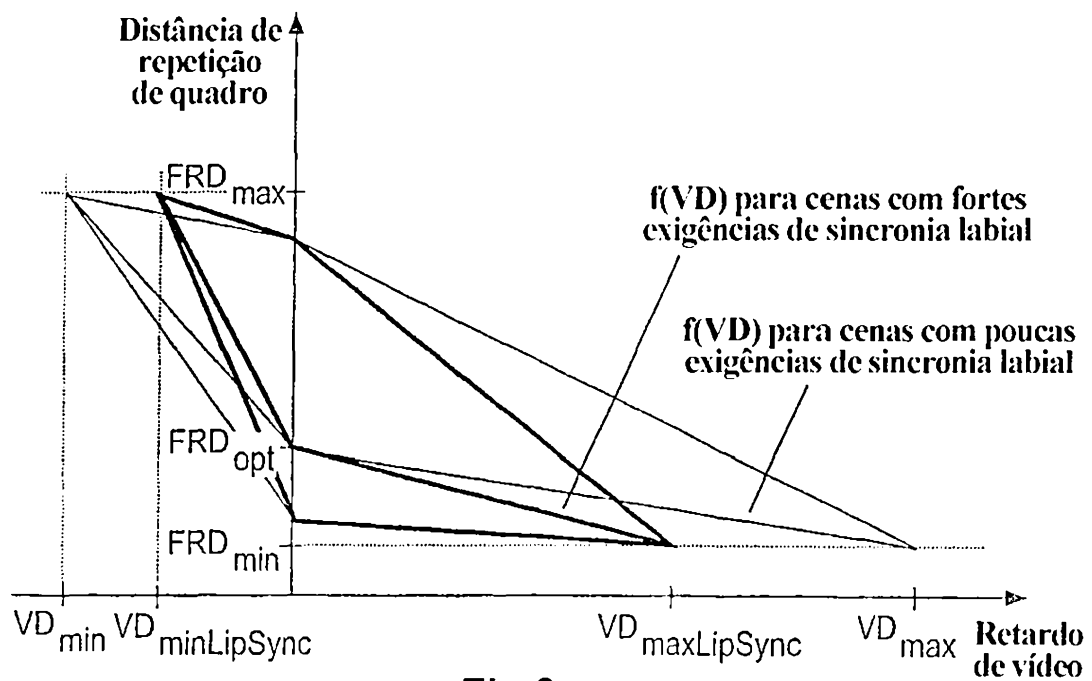
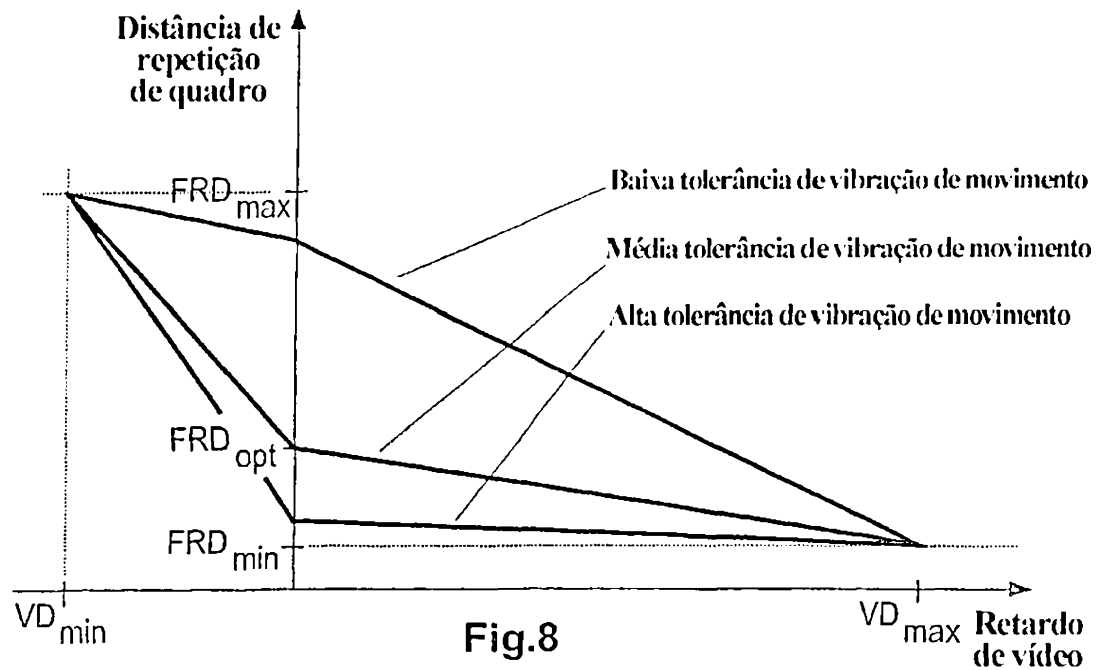
Fig.2



P. 104.16380



P 104 18350



P 104 18300

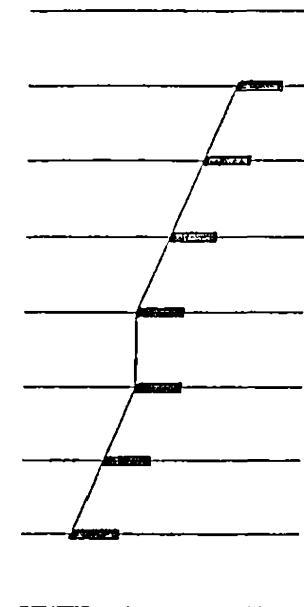


Fig.10



Fig.11

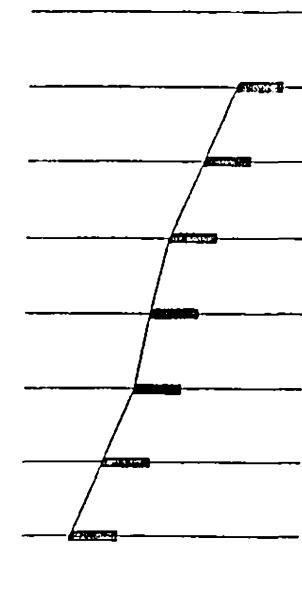


Fig.12



Fig.13