

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

(21) PI0617089-7 A2



* B R P I 0 6 1 7 0 8 9 A 2 *

(22) Data de Depósito: 01/09/2006
(43) Data da Publicação: 12/07/2011
(RPI 2114)

(51) Int.CI.:
G06F 17/30 2006.01
G06T 7/20 2006.01

(54) Título: MÉTODO PARA INTERROGAR OU
BUSCAR UM CORPO DE DADOS DIGITADOS
SEQÜENCIAIS, MÉTODO PARA AJUDAR NA
COMPUTAÇÃO DE WAVELETS PARA APLICAÇÕES
E MÉTODO PARA PROCESSAR UMA SEQÜÊNCIA
DE DADOS DIGITADOS

(30) Prioridade Unionista: 01/09/2005 US 60/712,810

(73) Titular(es): Astragroup AS

(72) Inventor(es): Bernard Jones

(74) Procurador(es): David do Nascimento Advogados
Associados

(86) Pedido Internacional: PCT GB06003243 de 01/09/2006

(87) Publicação Internacional: WO 2007/026162 de 08/03/2007

(57) Resumo: MÉTODO PARA INTERROGAR OU BUSCAR UM CORPO DE DADOS DIGITADOS SEQUENCIAIS, MÉTODO PARA AJUDAR NA COMPUTAÇÃO DE WAVELETS PARA APLICAÇÕES E MÉTODO PARA PROCESSAR UMA SEQÜÊNCIA DE DADOS DIGITADOS Quando são feitas gravações de dados digitais utilizando alguma forma de computador ou calculadora, os dados são inseridos em uma variedade de maneiras e são armazenados em alguma forma de meio eletrônico. Durante esse processo, os cálculos e as transformações são executados nos dados para otimizar os mesmos para o armazenamento. A presente invenção envolve projetar dos cálculos de tal maneira que eles incluem o que é necessário para cada um de muitos processos diferentes, tais como a compressão de dados, a detecção da atividade e o reconhecimento do objeto. Enquanto os dados de entrada são sujeitados a esses cálculos e armazenados, as informações sobre cada um dos processos são extraídas ao mesmo tempo. Os cálculos para os diferentes processos podem ser executados tanto em série em um único processador, ou em paralelo a múltiplos processadores distribuídos. O processo de extração é denominado como "decomposição sinótica", e as informações extraídas são denominadas como "dados sinóticos". O termo "dados sinóticos" não inclui normalmente o corpo principal de dados originais. Osdados sinóticos são criados sem nenhuma tendência prévia a interrogações específicas que podem ser feitas, de modo que é desnecessário inserir critérios de busca antes de fazer a gravação. Nem depende da natureza dos algoritmos/cálculos utilizados para fazer a decomposição sinótica. Os dados resultantes, que compreendem os dados originais (processados) juntamente com os dados sinóticos (processados), são então armazenados em um banco de dados relacional. Alternativamente,os dados sinóticos de uma forma simples podem ser armazenados como parte dos dados principais. Depois que a gravação é feita, os dados sinóticos podem ser analisados sem a necessidade de examinar o corpo principal dos dados. Essa análise pode ser feita muito rapidamente porque o volume dos cálculos necessários já foi executado no momento da gravação original. A análise dos dados sinóticos fornece marcadores que podem ser utilizados para acessar os dados relevantes da gravação dos dados principais caso requerido. O efeito global de fazer uma análise desta maneira é que uma grande quantidade de dados digitais gravados, que pode levar dias ou semanas para analisar por meios convencionais, pode ser analisada em segundos ou minutos. A presente invenção também se refere a um processo para a geração de famílias parameterizadas contínuas de wavelets. Muitas das wavelets podem ser expressas exatamente dentro de representações de 8 bit ou 16 bits. A presente invenção também se refere aos processos para a utilização de wavelets adaptáveis para extrair informações que são resistentes a variações em condições ambientais, e para executar a compressão de dados utilizando a quantização localmente adaptável e esquemas de limitação, e para executar a análise da pós-gravação.

MÉTODO PARA INTERROGAR OU BUSCAR UM CORPO DE DADOS DIGITADOS SEQÜENCIAIS, MÉTODO PARA AJUDAR NA COMPUTAÇÃO DE WAVELETS PARA APLICAÇÕES E MÉTODO PARA PROCESSAR UMA SEQÜÊNCIA DE DADOS DIGITADOS

5 ANTECEDENTES DA INVENÇÃO

CAMPO DA INVENÇÃO

Análise Pós-gravação

A presente invenção refere-se a um processo que permite que a análise muito rápida de dados digitais seja 10 realizada depois que os dados foram gravados.

Parametrização de Ondeletas

A presente invenção refere-se a um processo para a geração de famílias parametrizadas contínuas de ondeletas. Muitas das ondeletas podem ser expressas exatamente dentro 15 das representações de 8 bits ou de 16 bits.

Extração de Informações, Compressão de Dados e Análise Pós-gravação Utilizando Ondeletas

A presente invenção refere-se aos processos para a utilização de ondeletas adaptáveis para a extração de 20 informações que são resistentes a variações em condições ambientais, e para executar a compressão de dados utilizando a quantização localmente adaptável e esquemas de limitação e para executar a análise pós-gravação.

Uma vasta quantidade de dados digitais está sendo 25 gravada atualmente para aplicações em vigilância, meteorologia, geologia, medicina e muitas outras áreas.

A busca por esses dados para extrair informações relevantes é um processo maçante e demorado.

A menos que os marcadores específicos tenham sido 30 ajustados antes de fazer a gravação, a interrogação dos dados envolve a passagem através de toda a gravação de dados para buscar pelas informações desejadas.

Embora o processo de interrogação possa ser

automatizado, a necessidade de analisar todos os dados originais limita a velocidade em que a interrogação pode ser feita. As gravações de vídeo digitais, por exemplo, podem levar tanto tempo para serem exibidas quanto para serem 5 gravadas, e então a sua análise é um processo extremamente longo.

Quando uma situação de crise surge e as informações são requeridas imediatamente, o tamanho e o número de gravações podem tornar uma rápida extração de informações 10 simplesmente impossível.

Onde os marcadores específicos foram ajustados a priori, a interrogação subsequente dos dados gravados pode ser feita rapidamente, mas fica limitada às informações definidas por estes marcadores. A decisão sobre o que 15 procurar tem que ser feita antes que a gravação seja iniciada e pode envolver um processo de configuração complicado que tem que ser executado individualmente para cada gravação.

Uma característica chave da presente invenção é que os requisitos exatos da interrogação não têm que ser 20 especificados até depois que a gravação tenha sido feita. Uma gravação de dados simples padrão pode ser feita sem levar em consideração qualquer necessidade futura para a análise de dados.

Então, se uma análise posterior for necessária, o 25 processo permite que a interrogação seja feita de uma maneira extremamente rápida de modo que uma grande quantidade de dados possa ser analisada em um período de tempo curto.

Isto não somente propicia uma economia enorme em termos de potencial humano e custos, mas também fica possível 30 analisar uma vasta quantidade de informações digitais, em uma escala que, em termos práticos, seria anteriormente impossível.

O processo aplica-se a qualquer tipo de dados

digitais em fluxo, incluindo, mas sem ficar a eles limitados, dados de imagens, de áudio e dados sísmicos.

A análise pode ser de muitos tipos incluindo, mas sem ficar a elas limitados, as mudanças no comportamento 5 dinâmico dos dados e as mudanças na estrutura e distribuição espacial dos dados.

A análise pode ser geral (por exemplo, qualquer movimento não repetitivo ou qualquer objeto feito sob medida para o homem) ou pode ser detalhada (por exemplo, a 10 movimentação através de uma passagem ou uma similaridade específica a uma face específica).

Os exemplos do tipo de dados que estão sendo analisados geralmente incluem:

Gravações de vídeo digitais (para detectar tipos 15 particulares de atividade)

Gravações de vídeo digitais (para reconhecer determinados tipos de objetos, tais como faces ou placas numeradas)

Gravações de dados sísmicos (para detectar a 20 presença de minerais, etc.)

Gravações de dados sísmicos (para detectar a presença de ossos, vestígios arqueológicos, etc.)

Gravações de áudio (para detectar palavras chaves, sons especiais, padrões de voz, etc.)

25 Gravações de dados médicos (para detectar características particulares em cardiogramas, etc.)

Dados estatísticos (para monitorar fluxos de tráfego, tendências de consumo, etc.)

30 Dados ambientais (para analisar padrões meteorológicos, correntes marítimas, temperaturas, etc.).

Quando da análise das seqüências de vídeo, as ondeletas são freqüentemente utilizadas para a elaboração da decomposição de imagens. A utilização de ondeletas para essa

finalidade tem uma série de vantagens e as mesmas têm sido utilizadas em muitas aplicações.

Diversas classes de ondeletas foram definidas, as quais são particularmente bem adaptáveis a algumas 5 aplicações. Os exemplos incluem as ondeletas de Daubechie e de Coiflet. A presente invenção fornece uma maneira de expressar estas e todas as outras ondeletas de ponto uniformes de uma maneira parametrizada, utilizando uma variável contínua. Isto propicia uma maneira simples de 10 computar as ondeletas que podem ser automaticamente selecionadas para uma escala mais favorável e adaptadas desse modo ao teor dos dados.

A maior parte das ondeletas, incluindo as ondeletas de Daubechie e de Coiflet, envolve a computação de números 15 irracionais e elas devem ser calculadas utilizando a aritmética de ponto flutuante. A presente invenção fornece uma maneira de calcular as ondeletas que são arbitrariamente fechadas a qualquer ondeleta escolhida utilizando a aritmética de número inteiro. As computações de número 20 inteiro são exatas e reversíveis sem nenhum erro de arredondamento e podem ser executadas em microprocessadores utilizando menos energia e gerando menos calor do que seria requerido para a aritmética de ponto flutuante. Isto tem vantagens em muitas situações.

25 Os refinamentos nos métodos para filtrar o ruído e discriminar entre o movimento de fundo e o movimento intrusivo são úteis para otimizar o conteúdo das informações de dados sinóticos. A presente invenção apresenta métodos para a elaboração de uma série de tais refinamentos, 30 incluindo a utilização de uma pluralidade de moldes para determinar o fundo, a utilização da "substituição de kernel" também na determinação do fundo e um método de "pontuação de bloco" para estimar o significado das diferenças de pixel.

Na compressão das imagens de vídeo utilizando ondeletas, a utilização de ondeletas localmente adaptáveis fornece um mecanismo para proteger detalhes importantes nas imagens das consequências da forte compressão. Ao identificar 5 as áreas nas imagens que são de provável interesse especial, utilizando uma variedade de métodos para filtrar o ruído e determinar o fundo, máscaras podem ser construídas para excluir essas áreas da aplicação de algoritmos de forte compressão. Dessa maneira, as áreas de interesse especial 10 retêm níveis mais elevados de detalhes do que o restante da imagem, permitindo que os métodos de forte compressão sejam utilizados sem comprometer a qualidade das imagens.

A decomposição de ondeletas fornece um ambiente computacional natural para muitos dos processos envolvidos na 15 geração de dados sinóticos. As máscaras criadas ao identificar áreas especiais formam coletivamente um conjunto de dados que podem ser utilizados como dados sinóticos.

A invenção se aprofunda e sintetiza resultados de muitas especializações dentro do campo de processamento de 20 imagens. Em particular, a invenção explora uma pluralidade de decomposições piramidais dos dados de imagem com base em uma série de novas técnicas de análise de ondeletas. A utilização de uma pluralidade de representações de dados permite uma pluralidade de diferentes visualizações de dados que, quando 25 combinadas, fornecem indicações resistentes e confiáveis em relação ao que está acontecendo no nível de dados. Essas informações são codificadas como um conjunto de máscaras do atributo que se combina para criar os dados sinóticos que podem ser armazenados juntamente com os dados de imagem para 30 permitir a interrogação de alta velocidade e a correlação de vastas quantidades de dados.

DESCRIÇÃO DA TÉCNICA CORRELATA

A presente invenção refere-se aos métodos e

aparelhos de uma série de campos entre os quais se destacam: a extração de dados de vídeo, a detecção e classificação de movimento de vídeo, a segmentação de imagens, e a compressão de imagens de ondeletas. Um técnico no assunto será bem hábil 5 na técnica anterior que se refere a estes campos. Um dos principais problemas relacionados na presente invenção é o requisito para fazer este tipo de processamento de imagens em tempo real, um requisito que irá sempre impor limitações maiores em algoritmos tais como, por exemplo, a mudança de 10 gravação de televisão e vídeo para HDTV, e outros.

As variações na iluminação da cena constituem uma fonte principal de dificuldade na segmentação de fluxos de vídeo em tempo real. As comparações entre quadros sob tais circunstâncias são difíceis e dependentes de modelos, 15 particularmente quando as mudanças na iluminação são rápidas e episódicas. Aqui é introduzida uma maneira independente de modelo simples e eficaz se for manipulada em tempo real. O método que foi adotado também permite mover elementos no que seria de outra maneira o fundo da imagem (árvores balançando) 20 a ser manipulado com taxas muito baixas de detecções de falsos positivos.

Segmentação de imagens. O agora clássico documento de Toyoma, K.; Krumm, J.; Brumitt, B.; e Meyers, B. 1999. Wallflower: Principles and practice of background 25 maintenance. In International Conference on Computer Vision, 255-261 e as páginas da Web relacionadas à Microsoft Corporation (<http://research.microsoft.com/~jckrumm/WallFlower/TestImages.htm>) são recursos para o "sistema Wallflower" que é o tema 30 de uma vasta literatura. Métodos de segmentação baseados em equações diferenciais parciais (conforme exemplificado por Caselles et al. 1997 IEEE Trans Patt. Anal. Machine Intel., 19, 394) são interessantes, mas ainda não são realistas para

as aplicações em tempo real. Entre outros procedimentos foi encontrado Filtração de Kalman, Mistura de Modelos Gaussianos e de Modelos de Markov Escondidos.

Filtração de ruído das imagens. Este é um tema com 5 um histórico longo e venerável. Há uma abundância de métodos para identificar o componente do ruído que varia da limitação uniforme fácil aos métodos de estilo de entropia máxima faminta por recursos. O mundo das ondeletas foi dominado pelo trabalho surpreendente de Donoho e seus colaboradores (por 10 exemplo: o pioneirismo de D. L. Donoho e I. M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, páginas 425-455, 1994) e tudo que se seguiu. Há também uma riqueza de abordagens para a remoção de ruído com preservação de características com base em filtros não-lineares exemplificados por trabalhos anteriores tais como G. 15 Ramponi, "Detail-preserving filter for noisy images", *Electronics Letters*, 1995, 31, 865. Os filtros à base de filtros medianos pesados e outras estatísticas da ordem se voltam discutivelmente a J. W. Tukey em "Nonlinear methods 20 for smoothing data", *Conf. Rec. Eascom* (174) p673. "

Classificação e Busca. Parte do caráter do trabalho atual pode ser acompanhada por projetos de uma década atrás: projeto VISION (Video Indexing for Searching Over Networks), DVLS (Digital Video Library System) e QBIC (Query by Image and Video Content). Vide, por exemplo: M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, P. Yanker, *Query by Image and Video Content: The QBIC System*, Computer, v. 28 nº. 9, páginas 23-32, setembro de 1995 e "The VISION Digital 25 Video Library Project", S. Gauch, J. M. Gauch e K. M. Pua, *The Encyclopedia of Library and Information Science*. Vol. 68, suplemento 31, 2000, páginas 366-381, 2000. Desde esse tempo 30 tem havido muito desenvolvimento nesta área de automatização

das buscas de dados de vídeo.

5 Representações de Resolução Múltipla e Ondeletas na Geração de Imagens. A utilização de transformações de ondeletas hierárquicas (de resolução múltipla) para a manipulação de imagens tem uma vasta literatura que engloba uma faixa de tópicos que inclui a remoção de ruído, a procura por características e a compressão de dados. Os argumentos freqüentemente se voltaram à questão a respeito da maneira na qual a ondeleta funciona melhor e por que, 10 sendo que as ondeletas com finalidades especiais são produzidas para cada aplicação.

15 Outras Tarefas do Processamento de Imagens. Até mesmo dentro dos confins estreitos da indústria de segurança e de vigilância pode-se observar que as aplicações da geração de imagens cobrem aspectos da aquisição da imagem tais como a agitação da câmera e aspectos do processamento em seqüência da imagem tais como a correspondência de região, a detecção do movimento e o acompanhamento do alvo. Grande parte dessa tecnologia foi inserida em produtos comerciais. A eliminação 20 do movimento aleatório da câmera e o acompanhamento do movimento sistêmico foram ressaltados por vários pesquisadores. Aqui serão citados alguns trabalhos do programa de óptica adaptável (OA) da comunidade de astronomia. Entre uma série de métodos testados, o método de 25 Correlação Quad é muito simples e eficaz em uma situação em tempo real. Herriot et al. (2000) Proc SPIE, 115, 4007 é a fonte original. Vide Thomas et al. (2006) Mon. Not. R Astr. Soc. 371, 323 para uma revisão recente, também no contexto de estabilização de imagem astronômica.

30 DESCRICAÇÃO RESUMIDA DA INVENÇÃO

Ao fazer gravações de dados digitais utilizando alguma forma de computador ou calculadora, os dados são inseridos em uma variedade de maneiras e são armazenados em

alguma forma de mídia eletrônica. Durante este processo, os cálculos e as transformações são executados nos dados para sua a otimização para o armazenamento.

A presente invenção envolve o projeto dos cálculos de uma maneira tal que incluam o que é necessário para cada processo diferente, tais como a compressão de dados, a detecção da atividade e o reconhecimento do objeto.

Enquanto os dados de entrada são submetidos a esses cálculos e armazenados, as informações sobre cada um dos processos são extraídas ao mesmo tempo.

Os cálculos para processos diferentes podem ser executados em série em um único processador ou em paralelo em processadores múltiplos distribuídos.

É feita referência ao processo de extração como a "decomposição sinótica" e às informações extraídas como "dados sinóticos". O termo "dados sinóticos" não inclui normalmente o corpo principal dos dados originais.

Os dados sinóticos são criados sem nenhuma inclinação prévia a interrogações específicas que podem ser feitas, e então é desnecessário inserir os critérios de busca antes de fazer a gravação. Nem depende da natureza dos algoritmos/cálculos utilizados para fazer a decomposição sinótica.

Os dados resultantes, que compreendem os dados originais (processados) junto com os dados sinóticos (processados), são então armazenados em um banco de dados relacional. Alternativamente, os dados sinóticos podem ser armazenados de uma forma simples como parte dos dados principais.

Depois que a gravação é feita, os dados sinóticos podem ser analisados sem a necessidade de examinar o corpo principal dos dados.

Essa análise pode ser feita muito rapidamente

porque o grosso dos cálculos necessários já foi feito no momento da gravação original.

A análise dos dados sinóticos fornece os marcadores que podem ser utilizados para acessar os dados relevantes a 5 partir da gravação principal dos dados, caso requerido.

O resultado final de se fazer uma análise dessa maneira é que uma grande quantidade de dados digitais gravados, que pode levar dias ou semanas para ser analisada por meios convencionais, pode ser analisada em segundos ou em 10 minutos.

Não há nenhuma limitação no estilo da interface do usuário necessária para executar a análise.

Em uma realização, a presente invenção é baseada em um processamento de imagens em tempo real através do qual as 15 imagens adquiridas são analisadas e segmentadas de tal maneira a identificar de modo confiável todos os alvos móveis na cena sem danos com relação ao tamanho, cor, formato, posição, padrão de movimento ou qualquer outro tal atributo que se possa ter em um conjunto de dados de fluxo. Tal 20 identificação será, tanto quanto possível dentro dos recursos disponíveis, independente do movimento sistêmico ou aleatório da câmera e independente das variações na iluminação da cena.

BREVE DESCRIÇÃO DOS DESENHOS

A Figura 1 é um diagrama de blocos do processo em 25 uma forma geral.

A Figura 2 mostra uma hierarquia da transformação de ondeletas. Transformações diferentes ocorrem entre níveis diferentes.

A Figura 3 mostra um processo de geração de 30 famílias de ondeletas com ondeletas de quatro pontos.

A Figura 4 mostra um processo de geração de famílias de ondeletas generalizado para ondeletas de seis pontos e de pontos uniformes de ordem superior.

A Figura 5 descreve os estágios separados da realização da presente invenção.

A Figura 6 descreve as etapas que são tomadas a partir do ponto de aquisição dos dados ao ponto em que os 5 dados são refinados suficientemente para análise detalhada e produção de dados sinóticos. As etapas envolvem a remoção de artefatos que surgem da movimentação da câmera e do ruído da imagem, e então a resolução das imagens em fundos estáticos e estacionários e em um componente em primeiro plano dinâmico.

10 A Figura 7 descreve o processo de agrupar temporal e espacialmente os pixels dos primeiros planos dinâmicos em uma série de máscaras do objeto que serão transformados nos dados sinóticos.

15 A Figura 8 descreve o processo de armazenamento dos dados em que a representação de ondeletas dos dados de imagem e dos dados sinóticos é comprimida.

A Figura 9 descreve o processo de interrogação de dados e recuperação.

20 A Figura 10 mostra os processos que ocorrem após a seleção do evento.

A Figura 11 mostra os processos que ocorrem no primeiro circuito através da análise da imagem recentemente adquirida.

25 A Figura 12 é a transformação piramidal: cada nível da pirâmide contém uma versão dos dados originais menor e de resolução inferior.

A Figura 13 mostra como a hierarquia é gerada primeiramente através da aplicação de uma ondeleta W_1 e então com uma ondeleta W_2 . O painel inferior mostra a maneira por 30 meio da qual os dados são armazenados.

A Figura 14 mostra o processo de substituição do kernel de ondeletas.

A Figura 15 mostra um conjunto de máscaras digitais

extraídas de uma seqüência de imagens. Essas máscaras irão se transformar posteriormente em parte dos dados sinóticos.

A Figura 16 mostra uma série de padrões de 3x3, com as pontuações atribuídas ao pixel central (painéis superiores), juntamente com a ilustração das pontuações de pixels desviantes totais em alguns blocos de 3x3 particulares (painéis inferiores).

A Figura 17 resume os elementos do processo de compressão de dados.

A Figura 18 mostra como há uma correspondência de um-para-um entre os dados de imagem sinóticos e os dados comprimidos por ondeletas.

A Figura 19 mostra as etapas de recuperação de dados e o ciclo de análise.

A Figura 20 descreve como os dados são adquiridos, processados, armazenados e recuperados.

DESCRÍÇÃO DETALHADA

SEÇÃO 1: ANÁLISE PÓS-GRAVAÇÃO

A Figura 1 é um diagrama de blocos do processo em uma forma geral. Os blocos 1 a 8 compreendem o "gravador" e os blocos 9 a 15 compreendem o "analisador". Cada um dos blocos individuais representa um processo ou conjunto menor de processos que pode ser novo ou conhecido. Os dados digitados seqüenciais são inseridos no gravador e são submetidos a uma ou mais decomposições piramidais (Bloco 1). Um exemplo de tal decomposição é uma transformação de ondeletas, mas qualquer decomposição piramidal será possível. Os dados decompostos são "peneirados" através de uma ou mais "peneiras" (Bloco 2) que separam tipos diferentes de índices de informações. Um exemplo é um filtro de ruído ou um detector de movimento. As peneiras podem ser aplicadas uma vez ou várias vezes de uma maneira repetitiva. Os resultados dos processos de peneiração são separados em três categorias

que dependem da finalidade de aplicação:

(a) dados "não desejados" (Bloco 3), que são tipicamente o ruído, mas esta categoria pode ser nula se um tratamento sem perda ou uma compressão de dados sem perda for 5 requerida;

(b) dados "principais" (Bloco 4) que contêm quaisquer informações com exceção de (a);

(c) dados "sinóticos" (Bloco 5) que consistem nos resultados de um número selecionado de processos de 10 peneiração, dependendo da finalidade da aplicação.

A propriedade chave dos dados sinóticos é que são dados peneirados em que os processos de peneiração extraíram informações de uma natureza geral e não identificaram simplesmente características ou eventos particulares em 15 posições particulares nos dados.

Em etapas opcionais, os dados principais separados são então comprimidos (Bloco 6) e os dados sinóticos separados também podem ser comprimidos (Bloco 7). Se os processos de peneiração fossem aplicados aos dados no ápice 20 da decomposição piramidal, o tamanho dos dados sinóticos seria geralmente significativamente menor do que o tamanho dos dados principais.

Os dados principais e os dados sinóticos são então armazenados em um banco de dados (Bloco 8) e posicionados em 25 seqüência. O índice liga os dados principais aos dados sinóticos correspondentes. Isto completa o estágio de gravação do processo.

O estágio de análise começa com o ajuste de um processo de interrogação (Bloco 9) que pode assumir a forma 30 de interrogações específicas sobre os dados, por exemplo, sobre a ocorrência de eventos particulares, a presença de objetos particulares que têm propriedades particulares ou a presença de tendências texturais na seqüência de dados. A

interface do usuário para esse processo pode assumir qualquer forma, mas as perguntas devem ser compatíveis com o formato e âmbito dos dados sinóticos.

Os subconjuntos seqüenciais relevantes dos dados 5 são determinados pelas perguntas, por exemplo, as perguntas podem limitar a interrogação a um dado intervalo de tempo e os dados sinóticos correspondentes são recuperados a partir do banco de dados e, caso necessário, são descomprimidos (Bloco 10). Os dados sinóticos recuperados são então 10 interrogados (Bloco 11). O processo de interrogação compreende a conclusão dos processos de peneiração que foram executados no Bloco 2, carregando os mesmos a um estágio conclusivo que identifica as características ou os eventos particulares em posições particulares - espacial ou 15 temporalmente - dentro dos dados. Os detalhes necessários para a extração dessas informações específicas são fornecidos no estágio de interrogação (Bloco 9), isto é, depois que a gravação foi feita. O resultado da interrogação é um conjunto de posições específicas dentro dos dados onde as condições da 20 pergunta são satisfeitas (Bloco 12). Os resultados são limitados pela quantidade de informações contida nos dados sinóticos. Se resultados mais detalhados forem necessários, os subconjuntos dos dados principais que correspondem às 25 posições identificadas devem ser recuperados do banco de dados (Bloco 13) e, caso necessário, descomprimidos. Uma peneiração mais detalhada é então aplicada a estes subconjuntos para responder às perguntas detalhadas (Bloco 14).

Para visualizar os dados correspondentes que 30 resultam dos Blocos 13 ou 14, uma interface gráfica do usuário apropriada ou outro programa de apresentação podem ser utilizados. Isto pode assumir qualquer forma. Se a descompressão dos dados principais for requerida para

peneiração ou visualização posterior (Blocos 13 ou 14), a decomposição piramidal original deve ser invertível.

A quantidade de computação necessária para a extração das informações dos dados sinóticos é menor do que a 5 quantidade de computação necessária para a extração das informações e a execução da peneiração adicional dos subconjuntos dos dados principais, mas ambos os processos requerem menos computação do que a peneiração dos dados principais gravados sem as informações fornecidas pelos dados 10 sinóticos.

Uma realização detalhada do processo é fornecida na Seção 3.

SEÇÃO 2: ONDELETAS E DECOMPOSIÇÃO DE ONDELETAS

Ondeletas em uma Dimensão. A transformação de 15 ondeletas de um conjunto de dados unidimensionais é uma operação matemática em um estiramento de dados por meio do qual os dados são divididos pela transformação em duas partes. Uma parte é simplesmente uma versão encolhida de metade do tamanho dos dados originais. Se isto for expandido 20 simplesmente por um fator de dois, claramente não irá reconstruir os dados originais a partir dos quais foi gerado: as informações foram perdidas no processo de encolhimento. O que é positivo sobre a transformação de ondeletas é que ela gera não somente a versão encolhida dos dados, mas também uma 25 porção dos dados que são requeridos para reconstruir os dados originais na expansão.

Somas e Diferenças. Referente a Figura 2. Os dados transformados são do mesmo tamanho que o original, mas consistem em duas partes: uma parte que são os dados 30 encolhidos e outra que se parece com todas as características que têm que ser adicionadas novamente na expansão. As mesmas são denominadas Soma, S, e Diferença, D, partes da transformação de ondeletas.

Um Exemplo Trivial. Um exemplo totalmente trivial consiste em considerar um conjunto de dados que consiste nos dois números a e b . A soma é $S = (a+b)/2$, sendo que a diferença é $D = (a-b)/2$. Os dados originais são reconstruídos 5 ao simplesmente fazer $a = S+D$, $b = S-D$. Esta é a base da mais elementar de todas as ondeletas: a Ondeleta de Haar.

Há uma quantidade enorme de ondeletas que fazem isto enquanto agem em qualquer número de pontos ao mesmo tempo. Todas têm propriedades um tanto diferentes e fazem 10 coisas diferentes aos dados. Desse modo, a pergunta proeminente é sempre sobre quais destas são melhores de serem utilizadas e sob quais circunstâncias.

Níveis. A parte da soma das ondeletas pode ser ela mesma transformada por ondeletas, para produzir uma peça 15 quatro vezes mais curta do que os dados originais. Isto deve ser considerado como o segundo nível de transformação de ondeletas. Os dados originais são desse modo do Nível 0, sendo que a primeira transformação de ondeletas é então de Nível 1.

20 É possível continuar até que os dados encolhidos sejam simplesmente um ponto (na prática, isto requer que o comprimento dos dados originais seja uma força de 2).

FILTROS DE ONDELETAS DE QUATRO PONTOS

Filtros de ondeletas de quatro pontos. Os filtros 25 de ondeletas de N pontos se tornaram importantes há aproximadamente uma década (vide I. Daubechies, 1992, Ten Lectures on Wavelets, SIAM, Philadelphia, PA) e o histórico da transformação de ondeletas data de muito tempo antes disso. Há numerosas revisões e várias abordagens sobre o 30 tema, todas descritas em diversos livros e artigos.

Aqui, o ponto de interesse são as famílias de ondeletas e, para fins de simplicidade, será focada a atenção nos filtros de quatro pontos. Os resultados se generalizam a

seis pontos e a um número uniforme superior de pontos.

O filtro de quatro pontos. O filtro de ondeletas de quatro pontos tem quatro coeficientes, que serão denotados por $\{\alpha_0, \alpha_1, \alpha_2, \alpha_3\}$. Dados os valores (h_0, h_1, h_2, h_3) de alguma função a quatro pontos igualmente espaçados em uma linha é possível calcular dois números s_0 e d_0 :

$$s_0 = \alpha_0 h_0 + \alpha_1 h_1 + \alpha_2 h_2 + \alpha_3 h_3 \quad 1.1$$

$$d_0 = \alpha_3 h_0 - \alpha_2 h_1 + \alpha_1 h_2 - \alpha_0 h_3$$

Se for deslocado o filtro $\{\alpha_0, \alpha_1, \alpha_2, \alpha_3\}$ ao longo de uma linha de pontos de dados $2N$, nas etapas de dois pontos, é possível calcular pares de números N (s_i, d_i). Desse modo:

$$\{h_0, h_1, h_2, \dots, h_{2N}\} \rightarrow \{s_0, \dots, s_N\} \{d_0, \dots, d_N\} \quad 1.2$$

no rearranjo dos coeficientes.

O requisito chave é que essa transformação seja reversível. Isto impõe as condições

$$\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 1 \quad 1.3$$

$$\alpha_0 \alpha_2 + \alpha_1 \alpha_3 = 0$$

Também temos

$$\alpha_0 - \alpha_1 + \alpha_2 - \alpha_3 = 0 \quad 1.4$$

$$\alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 = \sqrt{2}$$

Condições adicionais podem ser impostas nos coeficientes de modo que os dados transformados tenham propriedades desejáveis específicas, tal como um número particular de momentos que desaparecem.

Uma interpretação geométrica

As duas relações 1.3 admitem uma interpretação geométrica simples e elegante que permite que classifiquemos estas ondeletas de quatro pontos e encontremos conjuntos de coeficientes interessantes que têm valores de números inteiros exatos.

Referente a Figura 2. Toma-se um conjunto de eixos retangulares $\{Ox, Oy\}$ com origem O e desenha-se uma linha OC a 45° . Colocar o ponto C a uma distância de unidade de O e desenhar um círculo de diâmetro de unidade com o centro C .

5 Será útil identificar o ponto L onde o círculo cruza Ox e o ponto M onde o círculo cruza Oy . A linha OC se estende para se encontrar com o círculo em I , então OI é um diâmetro e tem o comprimento de unidade.

Considera-se agora dois pontos P e Q no círculo de modo que o ângulo POQ seja um ângulo reto. Então PQ é uma diagonal do círculo. Identifica-se ψ como o ângulo que OP forma com o eixo Oy . Então, pela construção, ψ é o ângulo no sentido horário que OQ forma com o eixo Ox . Finalmente, atribui-se coordenadas a P e Q :

15 $P = P(\alpha_0, \alpha_3)$ 2.1

$$Q = Q(\alpha_2, \alpha_1)$$

e se tem tudo o que é necessário.

Os fatos que o círculo tem diâmetro de unidade e que PQ é um diâmetro dizem que $OP^2 + OQ^2 = 1$. Em termos das 20 coordenadas atribuídas dos pontos, isto mostra que

$$\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 1 \quad 2.2$$

A ortogonalidade dos vetores OP e OQ fornece

$$\alpha_1\alpha_3 + \alpha_0\alpha_2 = 0 \quad 2.3$$

que são precisamente a equação 1.3. Também se observa que uma 25 vez que $OL = OM = 1/\sqrt{2}$:

$$\alpha_0 - \alpha_1 + \alpha_2 - \alpha_3 = 0 \quad 2.4$$

$$\alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 = \sqrt{2}$$

que é 1.4.

Deve ser observado que há uma liberdade para 30 permutar as entradas fornecidas. As permutações deixam as relações 2.2, 2.3 e 2.4 inalteradas. Isto corresponde à

transformação

$$\psi \rightarrow \psi' = -\left(\frac{\pi}{4} + \psi\right) \quad 2.2$$

A família de ondeletas de quatro pontos. O ângulo ψ que OP faz com o eixo Oy determina uma família de ondeletas. É a família completa de ondeletas de quatro pontos uma vez que as equações 1.3 são condições necessárias e suficientes em coeficientes de ondeletas de quatro pontos. Sem perda de generalidade, foi escolhida a faixa de ψ para que seja de $-45^\circ < \psi < +45^\circ$.

10 As ondeletas mais famosas da família são relacionadas na tabela:

Nome	ψ
Daubechies 4	-15°
Haar	0°
Coiflet 4	15°

Há uma boa simetria, previamente despercebida, entre as ondeletas de Daubechies 4 e de Coiflet 4.

15 O ângulo ψ fornece uma maneira de dizer o quanto próximas estão duas ondeletas da família.

Uma parametrização alternativa. Podemos introduzir dois números, p e q , de modo que

$$\tan \psi = \frac{\alpha_3}{\alpha_0} = \frac{\alpha_1}{\alpha_2} = \frac{-p}{q} \quad 3.1$$

Uma vez que

$$20 \quad \frac{IP}{OP} = \tan \left\{ \frac{\pi}{4} - \psi \right\} = \frac{p - q}{p + q} \quad 3.2$$

temos

$$\alpha_0 = OP \operatorname{sen} \psi = -q \frac{p - q}{p + q} \quad 3.3$$

$$\alpha_3 = OP \cos \psi = p \frac{p - q}{p + q}$$

Por esse motivo, os coeficientes de ondeletas são

$$\{\alpha_0, \alpha_1, \alpha_2, \alpha_3\} =$$

$$\frac{1}{p+q} \{-q(p-q), q(p+q), p(p+q), p(p-q)\} \quad 3.4$$

Voltando ao fator normalizador correto, temos

5 $\{\alpha_0, \alpha_1, \alpha_2, \alpha_3\} =$

$$\frac{1}{\sqrt{2}(p^2+q^2)} \{-q(p-q), q(p+q), p(p+q), p(p-q)\} \quad 3.5$$

Se p e q forem números inteiros, temos, além do termo normalização, números inteiros por todo lado.

10 Aproximações de números inteiros. Se for observado que $\sqrt{3} \approx 7/4$, então os números irracionais que aparecem nas expressões familiares para a ondeleta daub4 são $3 + \sqrt{3} \approx 19/4$ e $3 - \sqrt{3} \approx 5/4$, de modo que $p = 19$ e $q = 5$, conduzindo à aproximação de números inteiros não normalizada

$$W_{\text{daub4}} \approx \{-35, 60, 228, 133\} \quad 4.1$$

15 Isto corresponde a $\psi = -14^\circ.744$, comparado com o valor real $\psi_{\text{daub4}} = -15^\circ$.

Há uma outra ondeleta de número inteiro de quatro pontos que fica bem perto disso com coeficientes não normalizados

20 $W_A \approx \{-3, 5, 20, 12\} \quad 4.2$

Isto tem $\psi = -14^\circ.03$.

Também deve ser observado que os mesmos coeficientes podem ser permutados para se obter uma outra ondeleta

25 $W_B \approx \{-3, 12, 20, 5\} \quad 4.3$

Isto resulta em $p = 5$ e $q = 3$, que, como esperado, tem $\psi = -30^\circ.96$. W_A e W_B têm larguras de banda eficazes diferentes.

A ondeleta mais simples é

$$W_X \approx \{-1, 2, 6, 3\}$$

4.3

$$W_Y \approx \{-1, 3, 6, 2\}$$

WX é conhecida como a ondeleta de quatro pontos com a largura de banda eficaz mais ampla.

5 Um denso conjunto de aproximações de números inteiros. Perto de qualquer número irracional há um número infinito de números racionais que formam um conjunto que se aproxima ainda mais de perto dos irracionais. Desse modo, há ondeletas não normalizadas com coeficientes de números 10 inteiros que se encontram arbitrariamente perto de qualquer ondeleta.

15 Ondeletas de seis pontos e ordens superiores. Com referência à Figura 3, pode-se observar como o processo acima é generalizado para ondeletas de seis pontos e ondeletas de 20 pontos de uma ordem superior. O painel superior da Figura 3 é uma versão atualizada da Figura 4: As coordenadas de P foram novamente etiquetadas para P(A, B), um novo círculo foi adicionado tendo OP como o diâmetro e um retângulo ORPS foi inscrito desenhado no círculo novo. Desse modo, os triângulos OSP e ORP são de ângulos retos e o ângulo SOR é um ângulo reto; em outras palavras, OS e OR são ortogonais. O painel inferior da Figura 3 extrai o retângulo ARPS e o triângulo OQP do painel superior: isso é tudo o que se faz necessário.

25 Agora fica fácil verificar que as seguintes relações estão satisfeitas:

$$\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \alpha_4^2 + \alpha_5^2 = 1 \quad 5.2$$

$$\alpha_0\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_4 + \alpha_3\alpha_5 = 0$$

$$\alpha_0\alpha_3 + \alpha_1\alpha_4 + \alpha_2\alpha_5 = 0 \quad 5.3$$

$$\alpha_0\alpha_4 + \alpha_1\alpha_5 = 0$$

$$30 \quad \alpha_0 - \alpha_1 + \alpha_2 - \alpha_3 + \alpha_4 - \alpha_5 = 0 \quad 5.4$$

$$\alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 = \sqrt{2}$$

e, desse modo, com esta construção

W{ $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_2, \alpha_3$ }

5.4

É uma ondeleta de seis pontos construída na de quatro pontos $\{\alpha_0, \alpha_1, \alpha_2, \alpha_3\}$. Certamente, o ciclo de geração das ondeletas de quatro pontos e das ondeletas de seis pontos 5 começa com a construção de uma ondeleta de quatro pontos com base em $Q = Q(\alpha_2, \alpha_1)$ (o círculo conduz a P automaticamente, dado Q).

O estágio seguinte, que gera um conjunto de ondeletas de seis pontos, começa com o desenho de um outro 10 círculo com OP como o diâmetro e o desenho de um retângulo inscrito ORPS e então a utilização de OS para continuar o processo.

Famílias de ondeletas. O estágio seguinte, que gera um conjunto de ondeletas de seis pontos começa com o desenho 15 de um outro círculo com OP como o diâmetro e o desenho de um retângulo inscrito ORPS e então a utilização de OS para continuar o processo. Isto fornece um mecanismo para aumentar o número de pontos na ondeleta por 2 cada vez. A família inteira é relacionada ao primeiro ponto Q e desse modo ao 20 ângulo ψ .

SEÇÃO 3: EXTRAÇÃO DE INFORMAÇÕES, COMPRESSÃO DE DADOS E ANÁLISE PÓS-GRAVAÇÃO Utilizando ONDELETAS

A presente invenção compreende uma série de processos individuais, sendo que alguns ou todos podem ser 25 aplicados ao utilizar ondeletas para extrair as informações dos dados digitados multidimensionais e para comprimir os dados. A invenção também apresenta um contexto natural para a execução da análise pós-gravação tal como descrito na Seção 1.

30 Os dados podem assumir a forma de qualquer conjunto de dados digitados de pelo menos duas dimensões. Tipicamente, uma das dimensões é o tempo, fazendo um conjunto de dados

seqüencial. Os processos são especialmente apropriados para o tratamento de imagens de vídeo digitadas, que compreendem uma seqüência de pixels de imagem que têm duas dimensões espaciais e planos adicionais de cor e de intensidade de 5 informações.

Na descrição a seguir, referência será feita a essa realização preferida, mas os processos podem ser aplicados de um modo equivalente a qualquer conjunto de dados digitados multidimensionais.

10 Entre os processos que são particularmente relevantes estão incluídos os seguintes:

a. Substituição do kernel (parágrafos da página 26, linha 24 e página 47, linha 9)

15 b. Representação de ondeletas adaptáveis de imagens (parágrafos da página 24, linha 7 e página 45, linha 32)

c. Auto-thresholding de diferenças de imagem (parágrafos da página 28, linha 31 e página 57, linha 28])

20 d. Utilização de moldes feitos sob medida para permitir múltiplos métodos de comparação (parágrafos da página 51, linha 23 e página 53, linha 11)

e. Conjuntos específicos de ondeletas ajustáveis (parágrafo da página 19, linha 6)

25 f. Método de pontuação de bloco para a discriminação do nível mais elevado e a classificação de eventos detectados (parágrafo da página 64, linha 3)

g. Utilização do limite localizado e níveis de quantização junto com a difusão de erro controlada para melhorar a qualidade percebida de imagens comprimidas (parágrafos da página 35, linha 27; da página 47, linha 18; e 30 parágrafos da página 70, linha 4 a página 73, linha 6)

Referência será feita agora em detalhes a uma realização da invenção, cujo exemplo é ilustrado nos desenhos em anexo. O exemplo descreve um sistema no qual uma seqüência

de imagens de vídeo é adquirida e processada para extrair as informações na forma de dados sinóticos, comprimidos, armazenados, recuperados, interrogados e os resultados indicados. Uma visão geral é apresentada na Figura 5.

5 Sempre que possível, os mesmos números de referência serão utilizados através dos desenhos e da descrição para se referir as partes idênticas ou similares.

10 Cada quadro de imagem em seqüência é submetido à decomposição de ondeletas. Na realização preferida, é feita a utilização de ondeletas parametrizadas tal como descrito na Seção 2, as quais ajudam na computação dos processos. No entanto, qualquer representação de ondeletas apropriada pode ser utilizada.

15 Em seguida, a menos que esteja indicado de alguma outra maneira, as indicações ao efeito com que uma "imagem" ou "quadro" é processada referem-se à hierarquia de ondeletas inteira e não simplesmente à imagem original.

20 A Figura 5 descreve o processo de aquisição inteiro (Bloco 12), através do processamento (Bloco 13) e da classificação (Bloco 14) para o armazenamento (Bloco 15) e a recuperação com perguntas (Bloco 16).

25 No bloco 12, em uma realização, as seqüências temporais das imagens de vídeo 11 são recebidas de uma ou mais fontes de vídeo e, caso requerido, são traduzidas a um formato digital apropriado às etapas posteriores. Os dados de qualquer fonte de vídeo podem ser censurados a uma taxa de quadro requerida. Os dados de uma série de fontes podem ser manipulados em paralelo e uma referência cruzada pode ser feita para acesso posterior aos fluxos múltiplos.

30 No bloco 13 as imagens são submetidas à análise de baixo nível à medida que são adquiridas. A análise é feita em termos de uma série de transformações piramidais (de resolução múltipla) dos dados de imagem, culminando em uma

transformação de ondeleta adaptável que é um precursor à compressão da imagem.

A análise identifica e remove o ruído não desejado e identifica qualquer movimento sistêmico ou aleatório da 5 câmera. É importante lidar com qualquer ruído nos componentes de cor das imagens, uma vez que este é o lugar onde as câmeras de CCTV de baixo custo são as mais fracas. Uma série de processos, a ser descrita, identifica então quais partes da imagem constituem fundos estáticos ou estacionários e 10 quais partes são componentes dinâmicos da cena. Isto é feito independentemente do movimento da câmera e independentemente das mudanças na iluminação. Os detalhes são mostrados na Figura 6 e descritos nos parágrafos da página 27, linha 17 e página 33, linha 23.

15 As máscaras digitais constituem uma parte importante do presente processo. As máscaras são codificadas e armazenadas temporariamente como planos de bits de um nível ou de múltiplos níveis. Um conjunto de máscaras de imagem digital é produzido ao delinear as regiões da imagem que têm 20 atributos diferentes. Em uma máscara de um bit, os dados em um ponto têm ou não têm o atributo particular. Uma máscara codificada com mais bits pode armazenar valores para os atributos. As máscaras são utilizadas para proteger partes particulares de uma imagem dos processos que poderiam 25 destruir as mesmas se elas não fossem mascaradas, ou para modificar partes dos dados seletivamente.

No bloco 14 os resultados da análise do bloco 13 são avaliados quantitativamente e uma análise mais profunda das partes dinâmicas da cena é efetuada. Os resultados são 30 expressos como um conjunto de máscaras digitais que serão transformadas mais tarde nos dados sinóticos. Os detalhes são mostrados na Figura 7 e descritos nos parágrafos da página 11, linha 13 e página 35, linha 15 e os exemplos de tais

máscaras são apresentados na Figura 15.

No bloco 15 a saída dos processos é descrita no bloco 14. As representações adaptáveis de ondeletas da cena original e de seus dados sinóticos associados são comprimidas 5 e armazenadas no disco para recuperação posterior. Os detalhes são mostrados na Figura 8 e descritos nos parágrafos da página 35, linha 23 e da página 36 linha 20.

No bloco 16 os dados sinóticos armazenados no bloco 15 são inquiridos e todas as respostas positivas à pergunta 10 são recuperadas a partir dos dados comprimidos da seqüência de imagem e indicadas como eventos. Um "evento" neste sentido é uma seqüência contínua de quadros de vídeo durante o qual o comportamento perguntado persiste junto com uma pluralidade de quadros relacionados de outras fontes de vídeo. Os 15 detalhes são mostrados nas Figuras AE e AF e descritos nos parágrafos página 36, linha 32 e página 38, linha 24.

A Figura 6 ilustra um circuito longo que consiste em diversos "nós de processamento" (Blocos 22 - 31) que constituem a primeira fase de resolução das seqüências de 20 vídeo 21 em componentes de acordo com a presente invenção.

Há uma série de características importantes deste circuito. (1): Ele pode ser executado qualquer número de vezes contanto que os recursos para que isso seja feito estejam disponíveis. (2): A execução do processo em qualquer 25 nó é opcional, dependendo do tempo, dos recursos e da estratégia algorítmica total. (3): O processamento pode levar imagens precedentes em consideração, outra vez dependendo da disponibilidade dos recursos. Este processo repetitivo pode ser expresso como

$$30 \quad S_j = S_{j-1} + I_j, \quad S_{-1} = 0$$

6.1

onde S_{j-1} é o estado de conhecimento na extremidade do circuito $j-1$, e I_j é a informação que será adicionada para produzir um novo estado S_j no circuito j .

A finalidade deste circuito é dividir os dados em uma série de componentes: (1) Ruído, (2): Dados limpos para a análise que será comprimida eventualmente, (3): Componentes estáticos, estacionários e dinâmicos dos dados. As definições 5 para estes termos são fornecidas no Glossário e há uma discussão mais detalhada desta divisão de componentes nos parágrafos página 39, linha 1 e página 40 linha 4.

No bloco 21 é recebida uma série de quadros de vídeo.

10 No bloco 22 cada quadro 21 é transformado em uma representação de ondeletas utilizando alguma ondeleta apropriada. Em uma realização, por razões de eficiência computacional, uma ondeleta de números inteiros de 4 casas que tem coeficientes de números inteiros pequenos é 15 utilizada. Isto permite uma análise dos dados de primeira passagem computacionalmente eficiente.

No bloco 23 a diferença entre as transformações de ondeletas computadas no bloco 22 do quadro de vídeo atual e seu predecessor é calculada e armazenada. Em uma realização 20 deste processo, é computada uma diferença simples de pontos-de-dados-por-pontos-de-dados. Isto permite uma análise dos dados de primeira passagem computacionalmente eficiente. Em uma outra realização do processo, uma diferença mais sofisticada entre os quadros é calculada utilizando o 25 processo de "Substituição de Kernel de Ondeleta" descrito em detalhes no parágrafo da página 48, linha 20. A vantagem da substituição de kernel de ondeletas é que é eficaz na eliminação das diferenças devido às mudanças na iluminação sem a necessidade de um modelo de fundo explícito.

30 No bloco 24 quadros sucessivos são verificados para ver se há movimento sistêmico da câmera. Em uma realização, isto é feito ao correlacionar as características principais da transformação de ondeleta de primeiro nível da diferença

do quadro calculada no bloco 23. O parágrafo da página 41, linha 26, se expande sobre outras realizações deste processo. O deslocamento computado é registrado para predizer o movimento subsequente da câmera através de um processo de 5 extração. Uma máscara digital é computada ao gravar essas partes da imagem atual que se sobrepõem ao seu predecessor, e a transformação entre as regiões da sobreposição é computada e armazenada.

No bloco 25 todos os resíduos do movimento 10 sistêmico da câmera são tratados como sendo devidos ao movimento irregular da câmera: agitação da câmera. A agitação da câmera torna a imagem visível não somente difícil de olhar, mas ela também descorrelaciona quadros sucessivos tornando a identificação do objeto mais difícil. A correção 15 para a agitação da câmera é geralmente um processo repetitivo: a primeira abordagem pode ser melhorada uma vez que se sabe o que é o fundo estático do campo da imagem (vide o parágrafo). Por sua natureza, os componentes estáticos da imagem permanecem fixos e então é facilmente possível 20 construir rapidamente um molde de fundo especial para essa finalidade. O isolamento das principais características desse molde torna a correção para a agitação da câmera relativamente objetiva. Vide o parágrafo da página 41, linha 26, para maiores detalhes.

25 No bloco 26 essas partes da imagem atual que diferem em menos de algum limite determinado (automaticamente) são utilizadas para criar uma máscara que define essas regiões onde a imagem não mudou em relação ao seu predecessor. Na primeira passagem até o bloco 26, o 30 limite é computado, em uma realização do processo, a partir do histograma truncado de valor extremo da diferença de imagem e, em uma outra realização, a partir da estatística mediana das diferenças de pixel. A máscara é reajustada em

cada passagem. Vide o parágrafo da página 42, linha 3, para maiores detalhes técnicos.

No bloco 27 a máscara calculada no bloco 26 é utilizada para refinar os parâmetros estatísticos da 5 distribuição do ruído da imagem. Esses parâmetros são utilizados para separar a imagem em um componente de ruído e em um componente limpo.

Em uma realização repetitiva, o processo retorna ao bloco 23 a fim de refinar as estimativas de movimento e ruído 10 da câmera.

Ao utilizar câmeras dé CCTV de baixo custo é importante tratar corretamente o ruído nos componentes de cor de sinal, uma vez que este é freqüentemente completamente substancial. As bordas afiadas nas imagens são 15 particularmente suscetíveis ao ruído de cor.

No bloco 28 a imagem limpa atual do bloco 27 é submetida à decomposição piramidal utilizando uma nova Transformação de Ondeletas Adaptável. Em tal decomposição piramidal dos dados, cada nível da pirâmide é construído 20 utilizando uma ondeleta cujas características são adaptadas às características da imagem nesse nível. Em uma realização, as ondeletas utilizadas nos níveis de alta resolução (superiores) da pirâmide são as ondeletas de alta resolução, sendo que aquelas utilizadas nos níveis inferiores são as 25 ondeletas de definição inferior da mesma família parametrizada. O processo é ilustrado adicionalmente no parágrafo da página 43, linha 20, e discutido nos parágrafos da página 19, linha 6 e da página 22, linha 17, onde várias famílias de ondeletas apropriadas são apresentadas.

30 Os coeficientes numéricos que representam essa decomposição de ondeletas adaptável da imagem podem ser censurados, quantizados e comprimidos. Em qualquer nível da decomposição, a censura e a quantização podem variar

dependendo de (a) onde há características descobertas na transformação de ondeleta e (b) onde o movimento foi detectado (a partir das máscaras de movimento do bloco 26 ou do bloco 30 se o processo foi repetido).

5 No bloco 29 uma nova versão da imagem atual é criada utilizando as informações de baixa resolução da transformação de ondeletas da imagem precedente. Essa nova versão da imagem atual tem a mesma iluminação total que seu predecessor. Este novo processo, "substituição de kernel de
10 ondeleta", é utilizado para compensar as mudanças entre os quadros na iluminação. Esse processo é elucidado em maiores detalhes no parágrafo da página 48, linha 20.

No bloco 30 as diferenças entre a imagem atual modificada pelo kernel do bloco 29 e a imagem precedente são
15 devidas ao movimento dentro da cena, sendo que a substituição do kernel eliminou em grande parte os efeitos devido às mudanças na iluminação. Uma máscara digital pode ser criada ao definir as áreas onde o movimento foi detectado.

O mesmo princípio que aquele do parágrafo da página
20 30, linha 19, é aplicado a uma série de imagens e moldes precedentes que já foram armazenados. Várias estratégias de armazenamento do molde estão disponíveis. Em uma realização desse processo, uma variedade de moldes diferentes é armazenada, os quais têm um quadro com um ano de idade (isto
25 é, o quadro de dados precedente), quadros com dois anos de idade, quadros com quatro anos de idade, e assim por diante em uma progressão geométrica. A limitação para isto se deve ao armazenamento dos dados e aos recursos computacionais adicionais requeridos para verificar um número maior de
30 moldes. Há uma discussão mais detalhada sobre os moldes no parágrafo da página 51, linha 24.

Os moldes são criados em uma variedade de maneiras de transformação de ondeletas dos dados. O molde mais simples

é a transformação de ondeletas de uma imagem precedente. Em uma realização, a média das imagens precedentes de ondeletas é armazenada como um molde adicional. Em uma outra realização, é armazenada uma média pesada pelo tempo sobre imagens de ondeletas passadas. Isto é computacionalmente eficiente se a seguinte fórmula for utilizada para atualizar o molde T_{j-1} para T_j que utiliza a imagem mais recente é I_j :

$$T_j = (1-\alpha) T_{j-1} + \alpha I_j \quad 7.1$$

onde α é a contribuição fracionária da imagem atual ao molde. Com este tipo de fórmula, o molde tem uma memória na ordem dos quadros α^{-1} e os objetos móveis do primeiro plano são borrados e desvanezem eventualmente. Fundos estacionários tais como árvores com folhas balançando podem ser manipulados por este efeito de suavização: a detecção do movimento não ocorre mais contra um fundo de atividade pronunciada. (Vide o parágrafo da página 40, linha 4). A obtenção de tais moldes requer um período de "aquecimento" pelo menos dos quadros α^{-1} .

Em uma outra realização deste processo, os moldes são armazenados para uma pluralidade de valores α . Em algumas realizações, α depende de quanto a imagem I_j difere de seu predecessor, I_{j-1} : uma imagem altamente dissimilar iria poluir o molde, a menos que α se tornasse menor para esse quadro.

Diversas máscaras de histórico do molde são criadas, refletindo o nível da atividade passada na imagem filtrada de ruído. O comprimento do histórico armazenado depende da quantidade de memória atribuída a cada pixel de cada máscara e da quantidade de energia computacional disponível para atualizar continuamente as máscaras. As máscaras não precisam ser mantidas para todos os níveis da transformação de ondeletas.

Em uma realização, essas máscaras são de oito bits.

A "máscara de histórico recente" codifica a atividade de cada pixel durante os oito quadros precedentes como de 0 bit ou como de 1 bit. Duas "máscaras do nível de atividade" codificam a taxa média de transições entre os estados '0' e '1' e o comprimento percorrido consecutivo para o número '1' consecutivo em relação ao histórico passado. Em outras realizações, outras estatísticas do estado serão utilizadas - certamente não há nenhuma falta de possibilidades. Isto resulta em um dispositivo para codificar o nível de atividade em todos os pontos da imagem antes da segmentação em movimento em primeiro plano e movimentos de fundo.

Uma ou mais das máscaras do nível de atividade podem ser armazenadas como parte dos dados sinóticos. No entanto, geralmente elas não comprimem muito bem e desse modo, em uma realização, somente as máscaras de resolução inferior são armazenadas a intervalos dependentes das taxas de atualização do molde, α .

A imagem atual e a sua representação piramidal são armazenadas como moldes para comparações possíveis com dados futuros. Os moldes mais antigos podem ser reprovados se o armazenamento for um problema. Vide o parágrafo da página 51, linha 23, para maiores informações sobre moldes.

Em uma realização repetitiva, o processo retorna ao bloco 27 a fim de refinar as estimativas do ruído e os efeitos das variações na iluminação. Há uma série de características importantes desse circuito: (1): Ele pode ser executado qualquer número de vezes contanto que os recursos para que isso seja feito estejam disponíveis; (2): A execução do processo em qualquer nó é opcional, dependendo do tempo, dos recursos e da estratégia algorítmica total; (3): O processamento pode levar imagens precedentes em consideração, outra vez dependendo da disponibilidade dos recursos. Se a repetição for utilizada, nem todos os estágios precisam ser

executados no primeiro circuito.

No bloco 31 a análise do movimento é executada de uma maneira tal que são levados em consideração os fundos estacionários onde há um movimento limitado (ao contrário dos fundos estáticos que estão livres de movimento de qualquer tipo). Os limiares de decisão são definidos dinamicamente, dessensibilizando eficazmente as áreas onde há um movimento de fundo e as comparações são feitas com moldes de históricos múltiplos. A perda de sensibilidade que isto pode gerar pode ser compensada utilizando os moldes que são integrados durante períodos de tempo, desse modo borrando os movimentos localizados (vide o parágrafo da página 31, linha 4 e as discussões dos parágrafos da página 40, linha 4 e da página 51, linha 23).

O resultado é uma identificação temporária dos lugares na imagem transformada de ondeletas onde há uma atividade em primeiro plano. Isto será refinado quando as considerações sobre correlações espaciais e temporais forem apoiadas (vide o parágrafo seguinte e o parágrafo da página 63, linha 12).

No bloco 32 os lugares da imagem onde o movimento foi detectado no bloco 31 são reavaliados à luz de correlações espaciais entre as detecções e de correlações temporais que descrevem o histórico dessa região da imagem. Essa avaliação é feita em todos os níveis da hierarquia de ondeletas de resolução múltipla. Vide o parágrafo da página 63, linha 3, para maiores informações sobre isto.

A Figura 7 descreve um processo para agrupar temporal e espacialmente os pixels de primeiros planos dinâmicos em uma série de máscaras do objeto que serão transformadas nos dados sinóticos. Para continuidade, o bloco 32 é levado para esse diagrama da Figura 6.

No bloco 43 os dados dinâmicos do primeiro plano

revelados no bloco 31 são analisados espacial e temporalmente. Essa avaliação é feita em todos os níveis da hierarquia de ondeletas de resolução múltipla.

5 Em uma realização, a análise espacial é eficazmente uma análise da correlação: cada elemento do primeiro plano dinâmico revelado no bloco 31 é marcado de acordo com a proximidade de seus vizinhos entre esse conjunto (Bloco 44). Isto favorece os agrupamentos de pixels coerentes em todas as escalas e desfavorece os pixels dispersos e isolados.

10 Em uma realização, a análise temporal é feita ao comparar os elementos do primeiro plano dinâmico com os elementos correspondentes em quadros precedentes e com os dados sinóticos que já foram gerados para quadros precedentes (Bloco 44). Nessa realização, referências temporais 15 armazenadas são mantidas 1, 2, 4, 8... quadros no passado. A única limitação neste histórico é a disponibilidade de armazenamento rápido.

20 No bloco 45, os resultados da pontuação de correlação espacial e temporal são interpretados. Em uma realização, isto é feito de acordo com uma tabela pré-atribuída de padrões espaciais e temporais. Estes são denominados como peneiras espaciais e temporais (Blocos 46 e 47).

25 No bloco 48 os vários padrões espaciais e temporais são classificados em objetos e deslocamentos de cena. Para os objetos, os vetores de movimento podem ser calculados por qualquer um de uma variedade de dispositivos (vide o parágrafo da página 65, linha 18) e os 'thumbnails' podem ser armazenados, caso desejado, utilizando componentes de baixa 30 resolução da transformação de ondeletas. Para as mudanças de cena, caso desejado, uma seqüência de imagens passadas relevantes pode ser recolhida a partir dos componentes de baixa resolução da transformação de ondeletas para formar um

trecho de filme que possa ser examinado para referência futura. Em uma realização, um exame dos processos e parâmetros que geraram essas máscaras também é mantido.

No bloco 49 as máscaras de imagem são geradas para 5 cada um dos atributos do fluxo de dados descoberta no bloco 48, delineando onde nos dados de imagem o atributo está localizado. Realizações diferentes irão apresentar conjuntos de máscaras que descrevem categorias diferentes. ¶Essas máscaras formam a base dos dados sinóticos. A Figura 15 10 ilustra três máscaras que descrevem os componentes em mudança principais de uma cena.

No bloco 50 a versão final dos dados codificados por ondeletas livres de ruído está disponível para o estágio seguinte: compressão. A compressão dos coeficientes de 15 ondeletas será dependente do local.

A Figura 8 descreve os processos envolvidos na compressão, na criptografia e no armazenamento de dados para pergunta e recuperação posteriores. Os blocos 49 e 50 são tirados da Figura 7 para continuidade.

20 No bloco 61, os dados sinóticos gerados no bloco 49 são comprimidos sem perda com somas de controle dos dados e são então criptografados no caso da criptografia ser desejada.

No bloco 62 os dados de ondeletas codificados 25 adaptavelmente são comprimidos primeiramente por um processo de limite localmente adaptável e quantização para reduzir a taxa de bits, e então uma codificação dos coeficientes resultantes para o armazenamento eficiente. Em uma realização, pelo menos duas posições são determinadas e 30 codificadas com uma única máscara: os lugares na representação de ondeletas onde há um movimento dinâmico do primeiro plano e os lugares onde não há nenhum. Em uma outra realização, esses lugares na representação de ondeletas onde

há um fundo estacionário, mas não um fundo estático (por exemplo: as folhas se movendo) são codificados com uma máscara e recebem seu próprio limite e quantização.

As máscaras são codificadas e armazenadas para a 5 recuperação e reconstrução e os códigos de validação de imagem são criados para finalidades legais. Em uma realização, os dados comprimidos resultantes são criptografados e providos com somas de controle.

No bloco 63 os dados dos blocos 61 e 62 são 10 colocados em uma estrutura do banco de dados. Em uma realização, esta é uma utilização simples do sistema de arquivos de computador, em uma outra realização este é um banco de dados relacional. No caso de sincronização de tempo de fluxo de dados de entrada múltiplas, as informações são 15 vitais, especialmente onde os dados ultrapassam os limites de fusos horários.

No bloco 64 todos os dados são armazenados em sistemas de armazenamento locais ou em rede. Os dados podem ser adicionados e recuperados simultaneamente. Em uma 20 realização, os dados são armazenados em um dispositivo de armazenamento óptico (por exemplo: DVD). Uma trilha do exame validada é gravada ao lado dos dados.

A Figura 9 mostra o processo de Recuperação de Dados em que as perguntas são dirigidas sobre os Dados 25 sinóticos e, em resposta, uma lista é gerada dos eventos gravados que satisfazem essa pergunta. A pergunta pode ser refinada até que uma seleção final de eventos seja atingida. O bloco 64 é tirado da Figura 8 para fins de clareza.

No bloco 71 os dados ficam disponíveis para a 30 pergunta do bloco 72. A pergunta do bloco 72 pode ser lançada no computador local que mantém o banco de dados, ou através de uma estação remota em uma rede de computador. A pergunta pode envolver uma ou mais fluxo de dados para as quais há

dados sinóticos, e as fluxo relacionadas que não têm tais dados. A pergunta pode ser endereçada aos dados sinóticos distribuídos dentro dos bancos de dados diferentes em uma pluralidade de localizações e pode acessar os dados a partir 5 de uma pluralidade diferente de bancos de dados em uma pluralidade de localizações diferente.

No bloco 73 os Dados Sinóticos são procurados por correspondências à pergunta. Uma lista de quadros que corresponde à pergunta é gerada. Isto é referido como 10 "quadros chave". No bloco 74 uma lista de eventos é construída na base dos quadros chave descobertos.

Há uma distinção importante entre um evento e os quadros dos dados (quadros chave) a partir dos quais este é construído. Um evento pode consistir em um quadro único ou de 15 uma pluralidade de quadros de uma pluralidade de fluxo de dados de entrada. Onde é feita referência a uma pluralidade de fluxo de dados, os eventos definidos nos diferentes fluxos não precisam ser nem co-temporais nem do mesmo banco de dados que o quadro chave descoberto pela pergunta. Isto permite que 20 os dados sejam utilizados para finalidades investigativas em larga escala. Essa correspondência distribuída é conseguida no bloco 75. A construção dos eventos em torno dos quadros chave é explicada no parágrafo da página 81, linha 6.

No bloco 76 os dados associados com a pluralidade 25 de eventos gerada nos blocos 74 e 75 são recuperados a partir dos dados codificados por ondeletas associadas (Bloco 77) e a partir dos dados externos relevantes e disponíveis (Bloco 78) e descomprimidos conforme necessário. Os Quadros dos Dados dos blocos 77 e 78 são agrupados em eventos (Bloco 79) e 30 exibidos (Bloco 80).

No bloco 81 há uma avaliação dos resultados da busca com a possibilidade de refinar a busca (Bloco 82). O término da busca resulta em uma lista de eventos selecionados

(Bloco 83).

A Figura 10 mostra os processos que ocorrem após a seleção do evento (Bloco 81, que é aqui repetido para fins de clareza).

5 No bloco 91, os dados do evento são convertidos em um formato apropriado. Em uma realização, o formato é a mesma compressão de ondeletas adaptáveis conforme utilizado no armazenamento dos dados originais. Em uma outra realização, o formato pode ser um formato de terceiros para os quais há
10 visores de dados disponíveis (por exemplo: dados de áudio no formato Ogg-Vorbis).

No bloco 92 os dados são anotados, conforme pode ser requerido para finalidades de referência ou de exames futuros. Tal anotação pode ser armazenada como texto em um
15 banco de dados local simples ou em alguma ferramenta de terceiros projetada para acessar tais dados (por exemplo: uma ferramenta à base de SGML). No bloco 93 uma trilha de exame que descreve como essa busca de dados foi formulada e executada e um código de validação que assegura a integridade
20 dos dados é adicionado ao pacote.

No bloco 94 a lista inteira de eventos que resultam da pergunta e que compreendem os dados do evento (Bloco 79) e todas as anotações (Bloco 92) é acondicionada para armazenamento em um banco de dados ou é colocada onde o
25 pacote pode ser recuperado. No bloco 95 os resultados da busca são exportados para outras mídias; em uma realização, essa mídia é removível ou de armazenamento óptico (por exemplo: um dispositivo de memória removível ou um DVD).

Componentes dos dados

30 O ruído (N) é essa parte dos dados de imagem que não representa exatamente qualquer parte da cena. Geralmente ele surge dos efeitos instrumentais e serve para diminuir uma apreciação clara dos dados de imagem. Geralmente, refere-se

no componente de ruído como não tendo correlação com os dados de imagem (por exemplo, "neve" do vídeo superposta). Este não é necessariamente o caso, uma vez que o ruído pode depender diretamente da natureza local da imagem.

5 Os fundos estáticos (S) consistem nos elementos da cena que são fixos e que mudam somente em virtude das mudanças na resposta, na iluminação ou na oclusão da câmera pela movimentação de objetos. Um fundo estático pode existir até mesmo quando uma câmera está girando em movimento 10 panorâmico, inclinando ou fazendo zoom. A repetição de uma cena em épocas diferentes irá mostrar os mesmos elementos de fundo estático. Os edifícios e as estradas são exemplos de elementos que constituem o fundo estático. As folhas que caem de uma árvore durante períodos de dias devem entrar nessa 15 categoria: é meramente uma questão de faixas de tempo.

O fundo estacionário (M) consiste em elementos da cena que são fixados no sentido de que a repetição de uma cena em épocas diferentes irá mostrar os mesmos elementos em formas ligeiramente deslocadas. Os ramos e as folhas que se 20 movem em uma árvore são exemplos de componentes de fundo estacionário. O movimento é localizado e limitado e a sua variação de tempo pode ser episódica. As reflexões em uma janela devem entrar nesta categoria. O componente de fundo estacionário pode freqüentemente ser modelado como um 25 processo aleatório estacionário limitado.

O primeiro plano dinâmico (D) consiste nas características na cena que entram ou saem de cena ou que executam movimentos substanciais, durante o período da aquisição de dados. Um objetivo desse projeto é a 30 identificação dos eventos que ocorrem no primeiro plano ao apresentar muito poucas detecções falso-positivas e nenhuma falso-negativa.

Essas distinções entre componentes (parágrafos da

página 39, linha 1 e da página 39, linha 30) são distinções práticas que permitem que o implementador do processo tome decisões sobre a manipulação de vários aspectos da separação de componentes. Considere uma pessoa entrando em cena, 5 movendo uma cadeira e então saindo de cena. A cadeira é uma parte estática da cena antes de ser movida e depois de ter sido colocada para baixo. Quando em movimento, a cadeira é uma parte dinâmica da cena, bem como a pessoa que move a mesma. Isto enfatiza o fato que a separação em componentes 10 varia com o tempo e a implementação da separação deve levar isso em consideração.

Há algumas advertências ao fazer estas distinções. A distinção entre os fundos "estáticos" e "estacionários" é uma questão de selecionar uma escala de tempo relativamente à 15 qual o julgamento de valor é feito. Os ramos da árvore serão agitados ao vento em faixas de tempo de segundos, ao passo que a mesma árvore irá soltar as suas folhas durante períodos de semanas. Os ramos da árvore que se movem compreendem o componente "móvel" do fundo, sendo que, na ausência de tal 20 movimento, a perda das folhas é corretamente vista como parte do fundo estático (ainda que seja um componente que varia lentamente). À medida que começa a escurecer, a aparência da árvore muda, mas isto é mais bem considerado como um aspecto de decomposição estática.

25 Matematicamente, isto se resume à representação dos dados de imagem G como a soma de uma série de componentes dependentes do tempo:

$$G(\underline{x}, t) = G^S(\underline{x}) + G^M(\underline{x}, \varepsilon t) + G^D(\underline{x}, t) \quad 8.1$$

O primeiro componente é verdadeiramente estático; o 30 segundo se move lentamente no sentido descrito acima, sendo que o terceiro é o componente dinâmico que tem que ser classificado em seu primeiro plano e em uma contribuição de fundo. Observe que para as presentes finalidades, o caso das

câmeras que se movem sistematicamente é agregado em G^S . Uma definição mais precisa deve requerer que sejam mostradas explicitamente as transformações na coordenada espacial \underline{x} que resultam do movimento da câmera.

5 A base para classificar G^D em seu primeiro plano G^{DF} e os componentes de fundo G^{DB} , alega que G^{DB} , o componente de fundo dinâmico, é eficazmente estacionário:

$$\int_{-\infty}^T G^{DB}(\underline{x}, t) dt \rightarrow \hat{G}^S(\underline{x}) \quad 8.2$$

10 para alguns fundos estáticos $\hat{G}^S(\underline{x})$ (que representa onde as árvores deveriam estar se não estivessem balançando ao vento). A utilização de um molde pesado pelo tempo consegue isto e permite a separação dos componentes dinâmicos do primeiro plano (vide o parágrafo da página 51, linha 23).

15 O parâmetro ε determina o que se pretende dizer com uma taxa de mudança lenta. Idealmente, ε será pelo menos uma ordem de magnitude menor do que a taxa de aquisição de vídeo. Pode haver diversos componentes moventes, cada um com sua própria taxa ε :

$$G(\underline{x}, t) = G^S(\underline{x}) + \sum_i G_i^M(\underline{x}, \varepsilon_i t) + G^D(\underline{x}, t) \quad 8.3$$

20 O mais lento destes pode ser agregado no componente estático contanto que algo seja feito para esclarecer mudanças "adiabáticas" do componente estático.

25 A correção do movimento da câmera e da agitação da câmera em particular é uma arte com um longo histórico: há muitas abordagens. Em uma realização, o método de Correlação Quad de Herriot et al. (2000) Proc SPIE, 115, 4007 é utilizado. Vide Thomas et al. (2006) Mon. Not. R Astr. Soc. 371, 323 para uma revisão recente sobre o contexto de estabilização de imagem astronômica.

30 Filtro de Ruído de Primeiro Nível

O primeiro estimador do componente de ruído é obtido ao diferenciar dois quadros sucessivos da mesma cena e olhar a distribuição estatística dessas partes da imagem que são classificadas como "fundo estático", isto é, a versão 5 mascarada da diferença. A variância do ruído pode ser estimada de um modo consistente a partir de

$$\sigma_n = 1.483 \text{ Mediana } (M_n - M_{n-1}) \quad 9.1$$

onde

$$M_n = M (F_n - F_{n-1}) \quad 9.2$$

10 é a versão mascarada da diferença entre os quadros brutos.

Na primeira passagem, a máscara está vazia, ($M=I$, a identidade), uma vez que nada foi determinado ainda sobre o quadro F_n .

A mediana das diferenças é utilizada para estimar a 15 variação, uma vez que esta é mais estável para valores da parte separada do corpo principal (conforme seria causado por diferenças perceptíveis entre os quadros). Isto é particularmente vantajoso se, no campo de velocidade computacional, a variância for estimada a partir de uma sub-20 amostra aleatória de pixels da imagem.

Duas correções serão requeridas para esta estimativa de variação do ruído: (1) Correção para flutuações totais de intensidade da luz entre as cenas e (2) Correção para os elementos da imagem que não fazem parte do fundo 25 estático. A primeira dessas correções é feita através do processo de "Substituição de Kernel de Ondeletas" (parágrafo da página 48, linha 20). A segunda dessas correções é feita através do componente "VMD" da análise: visualizar em que partes da imagem houve mudanças significativas.

30 Se a máscara estiver vazia ($M=I$), a limpeza é alcançada ao ajustar para zero todos os pixels na diferença de imagem que tem valores menores do que algum fator vezes a variação e então nova escala do histograma das diferenças de

modo que a diferença mínima seja igual a zero ("Encolhimento de ondeletas" e suas variantes).

Se a máscara não estiver vazia, o valor da variação será utilizado para filtrar espacialmente o quadro F_n , 5 levando em consideração as áreas onde houve mudanças na imagem e os lugares onde a filtração pode ser prejudicial à aparência da imagem (tais como bordas importantes).

Há diversas técnicas possíveis para a filtração espacial dependente de características entre as quais são 10 incluídas (1) Filtragem do tipo Weiner dependente de fase e (2) Filtragem sensível a características não-lineares (por exemplo, os Filtros estilo Teager).

Observe que a remoção do ruído é a última coisa que é feita antes que a transformação de ondeleta das imagens 15 seja executada: a remoção do ruído é benéfica à compressão.

A Figura 11 sintetiza os processos que vão no primeiro circuito através da análise da figura recentemente adquirida. A Figura descreve um conjunto de quadros F_0 , F_{-1} , F_{-2} , F_{-3} ... que já foram adquiridos e utilizados para construir 20 uma série de moldes T_0 , T_{-1} , T_{-2} , T_{-3} ... e imagens de características de borda E_0 , E_{-1} , E_{-2} , E_{-3} ... Essas imagens E_i serão utilizadas para a detecção e o monitoramento da agitação da câmera. F_0 e T_0 irão se transformar nas imagens de referência para a nova imagem F_1 .

25 Se a agitação da câmera for detectada, esta é corrigida nesse momento (vide o parágrafo da página 41, linha 24). A correção pode necessitar um refinamento em uma repetição posterior.

30 F_1 (com agitação possivelmente corrigida) é comparado agora com o quadro precedente, F_0 e com o molde atual T_0 . Os mapas da diferença são computados e enviados a um detector de VMD, onde há duas possibilidades: há ou não há nenhuma mudança detectada em ambos os mapas da diferença.

Isto é tratado no parágrafo da página 42, linha 3.

Se não houver nenhuma mudança detectada, as características do ruído podem ser diretamente estimadas a partir do retrato da diferença $F1-F0$: todas as diferenças 5 devem ser devidas ao ruído. $F1-F0$ pode ser limpo e adicionado novamente à versão previamente limpa $f0$ de $F0$. Isto cria uma versão limpa $f1$ de $F1$, que está disponível para utilização na repetição seguinte.

Se houver uma diferença, então a correção para o 10 ruído tem que ser feita diretamente no quadro $F1$. A máscara que descreve onde há diferenças entre $F1$ e $F0$ ou $F1$ e $T0$ é utilizada para proteger as partes $F1-F0$ e $F1-T0$ onde houve uma mudança detectada neste nível. A limpeza dessas 15 diferenças permite uma versão $f1$ de $F1$ que foi limpa em toda parte a não ser onde havia uma mudança detectada. Essas regiões dentro da máscara, onde a mudança foi detectada, podem ser limpas utilizando uma borda de limpeza não-linear simples que preserva o filtro de ruído como o filtro de Teager ou uma de suas generalizações.

20 Representação de Dados em Termos de Transformações Piramidais

As transformações de ondeletas e outras transformações piramidais são exemplos de análise de resolução múltipla. Tal análise permite que os dados sejam visualizados em uma hierarquia de escalas e se tornou lugar 25 comum na ciência e na engenharia. O processo é ilustrado na Figura 12. Cada nível da pirâmide contém uma versão dos dados originais de resolução menor, inferior, juntamente com um conjunto de dados que representam as informações que têm que ser adicionadas novamente para reconstruir o original. 30 Geralmente, mas não sempre, os níveis da pirâmide escalam novamente os dados por um fator de dois em cada dimensão.

Há muitas maneiras de fazer isto: a maneira que é utilizada aqui é denominada representação de resolução

múltipla de Mallat por causa do matemático que a descobriu. O painel superior da Figura 13 mostra como a hierarquia é gerada primeiramente através da aplicação de uma ondeleta W_1 e então com uma ondeleta W_2 . O painel inferior mostra a maneira na qual os dados são armazenados.

A transformação de ondeletas de um conjunto de dados unidimensionais é um processo de duas partes que envolve somas e diferenças de grupos de dados vizinhos. As somas produzem médias desses dados vizinhos e são utilizadas para produzir encolhimentos. Versão dos dados de resolução inferior. A diferenciação reflete os desvios a partir das médias criadas pela parte de soma da transformação e isto é o necessário para reconstruir os dados. As partes da soma são denotadas por S e as partes da diferença por D . Os dados bidimensionais são processados primeiramente em cada fileira horizontalmente e então em cada coluna verticalmente. Isto gera as quatro peças descritas como $\{SS, SD, DS, DD\}$ mostradas na Figura 13.

A Hierarquia de Ondeletas. É usual a utilização da hierarquia de dados gerada por uma única ondeleta específica, escolhida a partir da abundância de ondeletas que são conhecidas. Desse modo, em termos da Figura 13, $W_1 = W_2$. As escolhas comuns para as ondeletas neste contexto são vários indivíduos da família CDF, a variante CDF(2,2) (também conhecida como "ondeleta 5-3") que é particularmente popular, em grande parte por causa de sua facilidade de execução.

Hierarquias de Ondeletas Adaptáveis. No processo descrito na presente invenção, uma hierarquia especial de transformações de ondeletas é utilizada, na qual os elementos da hierarquia são selecionados a partir de um conjunto contínuo de ondeletas parametrizadas por um ou mais valores. As ondeletas de quatro pontos dessa família requerem somente um parâmetro, sendo que os elementos de seis pontos requerem

dois, e assim por diante. Para um conjunto distinto de valores de parâmetro, os elementos de quatro pontos têm coeficientes que são números racionais: estes são computacionalmente eficientes e exatos.

5 A ondeleta utilizada em níveis diferentes é alterada de um nível a outro mediante a escolha de valores diferentes deste parâmetro. Isto é denominado uma Transformação de Ondeleta Adaptável. Em uma realização desse processo, uma ondeleta que tem alta resolução é utilizada no 10 nível de resolução mais elevado, enquanto que as ondeletas de resolução sucessivamente inferior são utilizadas à medida que se move para níveis de resolução inferiores.

Para qualquer ondeleta distinta, as larguras de banda eficazes do filtro podem ser definidas em termos da 15 transformação de Fourier do filtro de ondeletas. Algumas têm larguras de passagem mais largas do que outras: são utilizadas ondeletas de largura de passagem estreita nos níveis superiores (de alta resolução) e ondeletas de largura de passagem ampla nos níveis inferiores (de baixa resolução). 20 Em uma realização desse processo, as ondeletas que são utilizadas foram organizadas em um conjunto parametrizado requisitado pela largura de banda.

Nos níveis inferiores (pelos quais se deve entender aqueles níveis onde a transformação está sendo operada em uma 25 imagem que é quase o tamanho da imagem original), se está interessado em preservar detalhes e manter um bom fundo a fim de otimizar a compressão desses níveis. Nos níveis mais elevados (pelos quais se deve entender aqueles níveis que têm as menores imagens) está sendo mapeada a estrutura em grande 30 escala na imagem que é desprovida de características importantes. Além disso, a exatidão aqui é importante uma vez que todos os erros serão propagados através dos níveis inferiores onde serão altamente visíveis como artefatos de

bloco.

5 Thresholding. O thresholding nas partes SD, DS e DD da transformação de ondeletas elimina os valores de pixels que podem ser considerados como sendo ignoráveis do ponto de vista da compressão de dados de imagem. A identificação desses lugares onde o limiar pode ser maior é uma maneira importante de se conseguir uma compressão maior. A identificação de onde isto pode ser impróprio também é importante, uma vez que minimiza a degradação percebida da 10 imagem. A detecção da característica e a detecção do evento apontam para as localidades (espaciais e temporais) onde uma forte limitação forte deve ser evitada.

15 Quantização. A quantização refere-se ao processo em que uma faixa de números é representada por números menores de um conjunto, permitindo desse modo (embora aproximadamente) uma representação aproximada mais compacta dos dados. A quantização é feita após o thresholding e também pode depender do índice da imagem local (espacial e temporal). Os lugares onde o thresholding deve ser 20 conservador são também os lugares onde a quantização deve ser conservadora.

25 Empréstimo de bits. A utilização de um conjunto muito pequeno de números para representar os valores dos dados tem muitos inconvenientes e pode ser seriamente nociva à qualidade da imagem reconstruída. A situação pode ser auxiliada consideravelmente por qualquer técnica de uma variedade de técnicas conhecidas. Em uma realização desse processo, os erros a partir da quantização de um ponto de dados são levados a se difundir através dos pontos de dados 30 vizinhos, conservando desse modo tanto quanto possível o índice de informações total da área local. A redistribuição uniforme do restante ajuda a suprimir o contorno nas áreas com iluminação uniforme. Além disso, a redistribuição

criteriosa desse resto onde há características irá ajudar a suprimir os danos aos detalhes da imagem e irá produzir dessa maneira resultados visuais consideravelmente melhores. Isto reduz o contorno e outros tais artefatos. Isto é denominado 5 "emprestimo de bits".

O mecanismo para a distribuição do resto na técnica de empréstimo de bits é simplificado na análise de ondeletas uma vez que tal análise delineia facilmente as características da imagem a partir de áreas de dados 10 relativamente planas. As partes SD e DS da transformação em cada nível determinam a pesagem unida à redistribuição do restante. Isto torna o processo de empréstimo de bits computacionalmente eficiente.

KERNELS, MOLDES E LIMIARES DE ONDELETAS

15 Substituição do kernel de ondeletas. Este é o processo por meio do qual as características em ampla escala (baixa resolução) de uma imagem precedente podem ser feitas para substituir essas mesmas características na imagem atual. Uma vez que a iluminação é geralmente um atributo em ampla 20 escala, esse processo pinta essencialmente a luz de uma imagem em outra e desse modo tem a virtude de permitir que a detecção do movimento (entre outras coisas) seja feita na face de variações claras bastante fortes e rápidas. A técnica é mais eficaz, uma vez que, na representação de ondeletas, os 25 componentes SD, DS e DD em cada nível têm então somente um componente DC bem pequeno.

Em uma realização desse processo, é utilizada a substituição do kernel para melhorar o VMD de primeiro nível que é feito como uma parte do ciclo de pré-processamento da 30 imagem. Isto ajuda a eliminar mudanças na iluminação e desse modo melhora a descoberta das mudanças no primeiro plano da imagem.

O processo de substituição do kernel de ondeletas é

esboçado na Figura 14 onde é visto o componente T3 do kernel do molde atual sendo colocado no lugar do componente F3 do kernel da imagem atual para produzir uma nova versão da imagem atual cujos componentes de ondeletas são J_i $\{J_0, J_1, J_2, T_3\}$. Esse dados novos podem ser utilizados no lugar da imagem original I_i $\{F_0, F_1, F_2, F_3\}$ para estimar o ruído e computar as várias máscaras.

Formalmente, o processo pode ser descrito tal como segue. Deixar que as imagens capturadas sejam denominadas como $\{I_i\}$. É possível derivar deste um conjunto de imagens, através da transformação de ondeletas, chamada $\{J_i\}$, em que as variações espaciais em grande escala na iluminação foram removidas utilizando o kernel da transformação da imagem precedente.

Se houver duas imagens $\{I_i\}$ e $\{I_j\}$ da mesma seqüência com uma transformação de ondeleta que tem hierarquias do componente SS

$$\{I_i\} = \{^1SS(i), ^2SS(i), ^3SS(i), \dots, ^kSS(i)\} \quad 10.1$$

$$\{I_j\} = \{^1SS(j), ^2SS(j), ^3SS(j), \dots, ^kSS(j)\} \quad 10.2$$

é criada a nova imagem

$$\{J_j\} = \{^1\bar{SS}(j), ^2\bar{SS}(j), ^3\bar{SS}(j), \dots, ^k\bar{SS}(j)\} \quad 10.3$$

utilizando o kernel da imagem i para a imagem j .

Observe as sobre-barras nas partes SS das novas ondeletas - estas são modificadas pelo fato de que reconstruímos a imagem j utilizando o kernel de ondeletas i^{th} . Observe também que não foram modificadas as partes SD, DS ou DD da transformação: elas são utilizadas diretamente na reconstrução de $\{J_j\}$ a partir de $^kSS(i)$.

Então é possível calcular a diferença corrigida da luz ambiente entre a imagem $i=j-m$ e j :

$$\delta_{j, (m)} = J_j - I_{j-m} \quad 10.4$$

Essa diferença de imagem representa as mudanças na

imagem, uma vez que a imagem de m quadros antes foi tirada, sobre e acima de todas as mudanças devido à iluminação ambiente.

5 Há uma questão sobre a atualização do kernel da imagem j com aquela de $j-m$ ou vice-versa. Na prática, eficiência computacional faz com que a substituição seja feita tal como descrito, uma vez que sempre temos a transformação de ondeleta inteira da imagem atual armazenada na memória.

10 Mudanças Relativas. Na prática, é possível olhar somente para as mudanças em um único nível p da transformação de ondeletas:

$$\delta_{j,(m)}^p = {}^p\overline{SS}(j) - {}^pSS(i), m = j - i \quad 11.1$$

15 Isto descreve a diferença entre a parte SS do nível p^{th} do kernel substituída da transformação de ondeletas da imagem atual, j , pela parte correspondente da transformação de ondeletas da imagem i . O valor de retardamento m depende simplesmente da taxa de quadros e na prática se torna um comprimento fixo de tempo durante o qual as mudanças de 20 movimento são perceptíveis. No entanto, isto causa a perda de discriminação de tamanho que vem naturalmente com a análise de resolução múltipla e é sempre melhor utilizar a transformação inteira se possível.

25 Imagen atual. É comum pensar sobre a imagem atual como sendo simplesmente uma única imagem que se deseja avaliar em relação a suas predecessoras. Este é geralmente o caso. No entanto, há realizações deste processo em que pode ser útil substituir a única imagem atual com uma média de uma seleção de imagens precedentes.

30 Eliminação de transientes. Na aplicação do monitoramento ambiental, não é útil que se tenha as imagens poluídas por fenômenos transientes tais como animais, pessoas

e veículos. A utilização de dados que são uma média apropriadamente pesada do tempo para um conjunto recentemente passado de imagens irá eliminar transientes. É possível se referir a esses dados como a "imagem transiente eliminada 5 atual".

Em uma realização desse processo que foi adaptada a tal situação, a seguinte fórmula é utilizada para definir e atualizar a "imagem eliminada transiente atual" C_{j-1} a C_j que utiliza a única imagem mais recente é I_j :

10 $C_j = (1-\tau) C_{j-1} + \tau I_j$ 12.1

onde τ é a contribuição fracionária da imagem atual ao molde. Com esse tipo de fórmula, a imagem retém as informações na ordem dos quadros τ^{-1} . Nessa aplicação, os moldes seriam armazenados durante um período de tempo 15 significativamente mais longo do que os quadros τ^{-1} (dias ou até mesmo semanas, ao invés de minutos).

MOLDES E MÁSCARAS

Moldes. Em todos os processos descritos na presente invenção, uma variedade do que pode ser chamada de "moldes de 20 imagem" é armazenada em uma base provisória. Geralmente, os moldes são registros históricos dos próprios dados de imagem (ou sua transformação piramidal) e fornecem uma base para fazer as comparações entre a imagem atual e as imagens precedentes, exclusivamente ou em combinações. Tais moldes 25 são geralmente, mas não sempre, construídos ao co-adicionar grupos de imagens precedentes com fatores de pesos apropriados (vide o parágrafo da página 53, linha 11).

Um molde também pode ser uma variante na imagem atual: uma versão suavizada da imagem atual pode, por 30 exemplo, ser mantida para o processo de "unsharp masking" ou algum outro processo de imagem única.

Máscaras. As máscaras, tal como os moldes, também

são imagens, mas são criadas para delinear eficientemente aspectos particulares da imagem. Desse modo, uma máscara pode mostrar onde na imagem, ou em sua transformação piramidal, há um movimento acima de algum limiar ou onde alguma textura particular deve ser encontrada. A máscara é, portanto, um mapa junto com uma lista de atributos e seus valores que definem o índice das informações do mapa. Se o valor do atributo é "verdadeiro ou falso" ou "sim ou não", as informações podem ser codificadas como um mapa de um bit. Se o atributo for uma textura, o mapa pode codificar a dimensão local fractal como um número inteiro de 4 bits, e assim por diante.

Quando uma máscara é aplicada à imagem a partir da qual foi derivada, as áreas da imagem que compartilham de valores particulares do atributo da máscara são delineadas. Quando duas máscaras que têm os mesmos atributos são aplicadas a um par de imagens, a diferença entre as máscaras mostra a diferença entre as imagens a respeito desse atributo.

As informações sobre uma ou mais máscaras vão em direção aos Dados Sinóticos em construção para o fluxo de dados. A sinopse reflete os atributos que definiram os vários mapas de que são construídos. A Figura 15 ilustra três máscaras de nível-0 que correspondem ao primeiro plano dinâmico e os componentes de fundo estático e estacionário que devem ser colocados no fluxo de dados sinóticos.

Nessa figura, a Máscara de VMD revela uma porta de abertura e uma pessoa que sai dessa porta. A máscara de fundo móvel indica a posição das folhas e dos arbustos que se movem. A máscara de iluminação mostra onde há variações na iluminação devido às sombras das árvores que se movem. (Este último componente não aparece como parte do fundo móvel, uma

vez que é eliminado em grande parte pela substituição do kernel de ondeletas).

5 Moldes Específicos. Os moldes são as imagens de referência contra as quais é avaliado o teor da imagem atual ou de alguma variante na imagem atual (parágrafos da página 50, linha 30 e da página 51, linha 4). O molde mais simples é justamente a imagem precedente:

$$T_j = I_{j-1}$$

13.1

10 Ligeiramente mais sofisticada, é uma média das imagens m passadas:

$$\bar{T}_j = \frac{1}{m} \sum_{k=1}^m I_{j-k}$$

13.2

que têm a virtude de produzir um molde que reduz o ruído. Mais útil é a média pesada de tempo sobre as imagens passadas:

$$15 \quad \tilde{T}_j = \alpha I_j + (1-\alpha) \tilde{T}_{j-1}$$

13.3

onde α é a contribuição fracionária da imagem atual ao molde. Esta última equação pode ser alternativamente resolvida como

$$\tilde{T}_n = \alpha \sum_{r=0}^n (1-\alpha)^r I_{n-r}$$

13.4

20 mostrando \tilde{T}_n como uma soma pesada de quadros passados com as imagens do quadro r que tinham previamente o fator de pesagem $\alpha(1-\alpha)^r$. Com este tipo de fórmula, o molde tem uma memória na ordem dos quadros α^{-1} e assim a obtenção deste molde requer um período de "aquecimento" pelo menos dos quadros α^{-1} .

25 Na prática, α pode depender de quanto a imagem I_j difere de seu predecessor, I_{j-1} : uma imagem altamente dissimilar polui o molde a menos que α seja feito menor para esse quadro. A flexibilidade em escolher α é utilizada quando uma oclusão dinâmica de primeiro plano mudaria

significativamente o molde (vide parágrafo da página 60, linha 1).

5 Máscara de histórico recente. A "máscara de histórico recente" codifica a atividade de cada pixel durante os oito quadros precedentes como 0-bit ou 1-bit.

10 Máscaras do nível de atividade. Duas "máscaras do nível de atividade" codificam a média e a variância do número de 'uns' consecutivos sobre o histórico passado e uma terceira máscara recente da atividade codifica o comprimento do funcionamento atual dos 'uns'.

Outros moldes: Observe que não se fica restringido aos predecessores de I_j ao construir moldes. É para algumas finalidades úteis que se consideram os moldes com base nas imagens futuras tais como

$$15 \quad \dot{T}_j = I_{j+1} - I_{j-1} \quad 14.1$$

ou ainda

$$\ddot{T}_j = \frac{1}{2}(I_{j+1} - 2I_j + I_{j-1}) \quad 14.2$$

20 Conforme a notação sugere, estes são estimadores dos derivados da primeira e segunda vez do fluxo de imagem no momento em que a imagem I_j é adquirida. A utilização de tais moldes envolve a introdução de uma retardaçāo de tempo para proteger a análise do fluxo enquanto as imagens "futuras" são capturadas.

25 Há numerosas outras possibilidades. O Molde da imagem suavizado

$$S_j = Suave(I_j) \quad 14.3$$

onde "suavizado" representa qualquer um de uma série de operadores de suavização possíveis aplicados à imagem I_j . O Molde mascarado de imagem

$$30 \quad \hat{T}_j = Mask(T_j) \quad 14.4$$

onde o operador "Máscara" aplica uma máscara apropriadamente definida de imagem à imagem do molde T_i . A lista é obviamente longe de ser exaustiva, mas meramente ilustrativa.

Máscara de histórico recente. As "máscaras de histórico recente" codificam alguma medição da atividade de cada pixel na cena durante os quadros precedentes. Uma medição da atividade é se uma diferença de pixel entre dois quadros sucessivos ou entre um quadro e o molde então presente estava acima do limiar definido no parágrafo da página 60, linha 16.

Em uma realização, isto é armazenado como uma máscara de 8 bits do tamanho dos dados de imagem, de modo que a atividade é gravada para os oito quadros passados como um '0' ou um '1'. Cada vez que a diferença de pixel é avaliada, esta máscara é atualizada ao mudar o plano de bits apropriado.

Máscaras de histórico de período mais longo. Tal como as Máscaras de histórico recente, estas codificam dados históricos das cenas precedentes. A diferença é que tais máscaras podem armazenar os dados da atividade em instantes fiduciais no passado. Os pontos uniformemente espaçados são fáceis de atualizar, mas não tão úteis quanto os pontos geometricamente espaçados que são mais difíceis de atualizar. Tais máscaras facilitam a avaliação do comportamento a longo prazo a respeito da atividade da cena.

Máscaras do nível de atividade. Duas "máscaras do nível de atividade" apresentam um sumário estatístico da atividade em um pixel dado conforme apresentado na Máscara de histórico recente. As entradas na primeira dessas máscaras gravam o número ou a taxa das mudanças de estado que esse pixel foi submetido. Isto é mais fácil de manter como uma média de funcionamento de modo que se a taxa for R_{j-1} e a mudança seguinte for $e_j = 0$ ou 1, então o estimador da taxa R

é atualizado para

$$R_j = \epsilon R_{j-1} + (1-\epsilon) e_j \quad 15.1$$

O número ϵ reflete a extensão dos dados sobre a qual é calculada a média dessa taxa.

5 A segunda máscara mantém um registro do comprimento médio das execuções onde $e_j = 1$: o "comprimento de execução da atividade". Isto deve ser calculado da mesma maneira que o estimador de taxa, desse modo, se a taxa for uma média ϵ tal como acima, assim deve ser o comprimento de execução da
10 atividade.

15 Essas máscaras da atividade são bem caras de manter e, desse modo, em algumas realizações, pode ser conveniente restringir a máscara a um nível menor da pirâmide de dados e aqueles níveis ainda menores acima dele. Tipicamente, foi verificado que a manutenção de um máximo de uma metade da resolução da imagem principal é perfeitamente adequado; este é o nível 1 ou o nível 2 na Figura 12.

20 Máscara da mudança de fundo - detecção de não-movimento. Há duas questões importantes que podem ser feitas sobre o fundo estático (que não deve, por definição, mudar). Há algo que é considerado normalmente como parte do fundo estático que não está mais lá? Por outro lado, há agora algo que é parte do fundo estático que não estava lá antes? Claramente, esse tipo de mudança deveria requerer que 25 houvesse algum movimento na cena para causar a mudança. No entanto, a pergunta é mais complexa do que meramente pedir para verificar uma mudança. A pergunta é se o fundo estático é sempre restaurado e, em caso afirmativo, quando?

30 As máscaras que registram o movimento do primeiro plano não podem suportar isto, então uma mudança de fundo especial deve ser utilizada, a qual permite a identificação das características no fundo estático através de comparação

ou correlação. Essa máscara irá permanecer constante se o componente de fundo estático não mudar, a não ser naqueles lugares obstruídos por objetos de primeiro plano dinâmicos. Desse modo, as diferenças entre as máscaras de fundo 5 estático, idealmente, serão iguais a zero e não custarão nada para armazenar.

Uma máscara ideal para essa finalidade é a soma das partes SD e DS do nível 1 da pirâmide de ondeletas (vide Figura 12) uma vez que ela mapeia as características na cena 10 com resolução relativamente alta. A diferenciação de duas tais máscaras sucessivas construídas a partir de suas representações substituídas pelo kernel de ondeletas permite que essa comparação seja feita contanto que também se tenha acesso às máscaras do componente dinâmico correspondente. Com 15 estas últimas, é possível eliminar as características que correspondem às partes móveis da cena.

A mudança da máscara de fundo resultante pode ser comprimida e armazenada como parte dos dados sinóticos.

DIFERENÇAS ENTRE IMAGENS

20 Diferença de Imagens. Para as finalidades desta seção, a palavra "imagem" será considerada como se referindo a qualquer uma das seguintes. (1) Uma imagem que foi capturada de um fluxo de dados, (2) Uma imagem que foi capturada de um fluxo de dados e processada subsequentemente. 25 Desse modo, são incluídas até mesmo as transformações da imagem tais como uma versão encolhida da imagem ou sua Transformação de Ondeletas. (3) Parte de uma imagem ou uma de suas transformações.

30 Em outras palavras, é levada em consideração a comparação de uma disposição de dados tirada de um fluxo de tais disposições com seus predecessores.

Será denotado o j^{th} como tal disposição no fluxo pelo símbolo I_j e o objeto relativo ao qual é feita

comparação (o "molde") pelo símbolo T_j . T_j pode ser qualquer um dos vários moldes que podem ser definidos a partir de outros elementos do fluxo I_j .

É considerada a avaliação das diferenças entre uma 5 imagem e qualquer um destes vários moldes. Considere a diferença de imagem

$$\delta_j = I_j - T_j \quad 16.1$$

A média dos pixels que constituem δ_j não precisa ser igual a zero, a menos que todas as imagens que constituem 10 o molde T_j e a imagem I_j sejam idênticas. Este é um ponto importante quando se considera as estatísticas dos valores de pixels de δ_j .

Em média, os valores dos pixels na imagem δ_j são 15 iguais a zero se as mudanças de luz ambiente forem de tal modo que a substituição do kernel (parágrafos da página 48, linha 20 a página 49, linha 5) é eficaz. Quando os pixels não são iguais a zero, é preciso avaliar se correspondem a mudanças reais na imagem ou se isso é devido a flutuações estatísticas.

20 Pixels desviantes. Aqui se concentra no acompanhamento, como uma função do tempo, dos valores dos pixels nas diferenças de imagens. Os critérios desenvolvidos utilizam o histórico da série de tempo das variações em cada pixel sem levar em consideração a posição do pixel ou o que 25 seus vizinhos espaciais estão fazendo. Isto tem a vantagem que o ruído não-uniforme pode ser manipulado sem fazer suposições sobre a distribuição espacial do ruído. A distribuição espacial dessa variação será considerada posteriormente (vide o parágrafo da página 63, linha 11). Em 30 uma realização desse processo, o histórico de tempo de cada pixel nos dados é seguido e modelado. A partir desse histórico, um nível de limite de pixel L_i é definido em

termos de uma quantidade que pode ser chamaada de "nível de discriminação de execução", M_i , para o processo aleatório que descreve o histórico de cada pixel.

Supondo que para a diferença de imagem δ_i é 5 possível determinar um nível de limite L_i acima do qual se acredita (de acordo com algum teste estatístico) que o valor do pixel não pode ser devido ao ruído: um "valor de pixel desviante". Então é possível decidir que na diferença de imagem δ_j deve ser considerado um pixel que tem um valor Δ_j 10 desviante se tivesse

$$|\Delta_j| > \lambda L_i \quad 17.1$$

para algum fator de segurança λ . (É reconhecido que para uma distribuição enviesada dos valores de pixels em δ_j é possível escolher limites diferentes para valores positivos e 15 negativos de Δ ; no entanto, por causa de simplicidade notacional, se supõe que estes sejam os mesmos).

Uma vez que as mudanças Δ_j nos valores do pixel são um processo aleatório não-estacionário, o valor de L_i deve refletir o envelope superior dos valores de $|\Delta_j|$. Os envelopes superiores são notoriamente difíceis de estimar 20 para tais processos e assim é preciso recorrer a algumas suposições simplificadas. Isto é especialmente verdadeiro, uma vez que isto tem que ser feito para cada pixel e há uma restrição de tempo computacional.

25 Nível de discriminação. Considera-se os valores precedentes de m de Δ_j , utilizando o cálculo desses valores, para cada pixel, um nível de discriminação M_j com base em uma fórmula tal como qualquer uma das seguintes:

$$M_j = \max\{|\Delta_{j-1}|, |\Delta_{j-2}|, |\Delta_{j-3}|, \dots, |\Delta_{j-m}|\}$$

$$30 M_j = \text{média}\{|\Delta_{j-1}|, |\Delta_{j-2}|, |\Delta_{j-3}|, \dots, |\Delta_{j-m}|\} + \kappa \quad 18.1$$

$$M_j = \beta |\Delta_{j-1}| + (1-\beta) M_{j-1}$$

A primeira destas é uma tentativa direta de atingir o envelope ao olhar as alturas de sinal em uma janela de intervalo de tempo m que se move. A segunda utiliza simplesmente a média do módulo das últimas alturas de sinal m junto com uma margem de segurança κ . A última destas é uma média pesada de tempo das alturas de sinal precedentes, sendo que a quantidade β reflete a pesagem de tempo relativa. Este é o mecanismo preferido.

Nível de Limiar de Pixel. Dado o nível de discriminação tal como definido acima (parágrafo da página 60, linha 1), é possível computar o nível de limiar de pixel L_j para cada pixel tal como segue. Ajusta o limite para que esse pixel seja

$$L_j = \alpha L_{j-1} + (1-\alpha) M_j \quad 19.1$$

para algum "parâmetro de memória" α . Observe que α não é o mesmo que a quantidade β que participa do cálculo do nível de discriminação M_j (a terceira das equações 18.2). É feita então a comparação para decidir se devemos "marcar" o pixel ou não como sendo desviante e para restaurar o valor de L_j para o cálculo seguinte do quadro caso o pixel seja desviante ou não:

$$L_j = \begin{cases} \text{if } \Delta_j > \lambda L_i, \\ 19.2 \end{cases}$$

$\{M_j$ em outro caso

Em outras palavras, não é atualizado o limiar para o pixel se esse pixel for considerado desviante. Isto evita a polarização que pode ser introduzida ao permitir que o limiar seja determinado por circunstâncias anômalas. Se o critério de aceitação fosse com base em desvios 3σ , por exemplo, este procedimento seria simplesmente equivalente à rejeição 3σ no cálculo do limiar.

Compensação para fundos que se movem. O que este

procedimento faz é permitir que o limiar se sobreponha aos picos de ruído. Para uma densidade de probabilidade conhecida para a distribuição de ruído os níveis podem ser ajustados de modo que haja uma probabilidade conhecida de que um pixel 5 seja considerado falsamente como sendo desviante. Na ausência de uma densidade de probabilidade conhecida da distribuição das diferenças de pixel, a decisão pode ser feita não-parametricamente utilizando testes padrão de graus variados de sofisticação.

10 O efeito líquido de um fundo que se move é a dessensibilização da detecção do movimento nas áreas onde a cena está mudando de uma maneira limitada e repetitiva. Isto pode acontecer, por exemplo, onde as sombras das árvores moldadas pelo sol estavam se movendo devido ao movimento do 15 vento: o limiar seria impulsionado porque a variância local das diferenças de imagem é aumentada.

Este é um mecanismo importante para evitar cascatas de alarmes falsos nos sistemas de detecção de vídeo. O inconveniente disto é que um mecanismo de detecção 20 suplementar pode ser requerido sob estas circunstâncias, uma vez que a dessensibilização cria o perigo de perder eventos importantes. Em uma realização, isto é resolvido ao utilizar moldes que tenham memórias relativamente longas, uma vez que tais moldes borram e absorvem tais movimentos. A comparação 25 da imagem é feita contra um fundo que é relativamente livre de características de fundo definidas móvientes (vide os parágrafos da página 40, linha 4 e da página 51, linha 24).

30 Os parâmetros. Na realização descrita, há diversos parâmetros que devem ser ajustados para a detecção de mudanças significativas dentro de um fluxo de imagem. Alguns destes parâmetros são fixados no início, sendo que outros irão variar com as condições ambientais e são "aprendidos".

É possível identificar diversos parâmetros que têm

que ser ajustados ou determinados ao utilizar o procedimento previamente descrito:

m

Este é o retardamento nos quadros para fazer a comparação. Claramente, a 25 quadros por segundo *m* será maior do que para três quadros por segundo. É óbvio que se tivesse sub-amostras dos 25 quadros por segundo para três quadros por segundo, seria utilizado por fim o mesmo valor de *m*. Desse modo, *m* é diretamente proporcional à taxa do quadro. O valor da constante de proporcionalidade depende do quanto rápido é o movimento que está sendo procurado em termos da velocidade transversal do quadro.

λ

Esta é a sensibilidade da detecção em um dado pixel: a quantia de anomalia do valor observado da mudança de pixel está relacionada aos valores observados previamente. Observe que é utilizado um critério máximo, em vez de um desvio médio ou padrão, a fim de testar os valores de pixel. *λ* está relacionado à estatística de primeira ordem na amostra de valores não-desviantes.

α

O fator da memória diz quanto do histórico passado de limiares é levado em consideração ao atualizar o valor do limiar para o quadro seguinte. Isto está relacionado à taxa de captação de quadro, uma vez que reflete a extensão do tempo em que as condições ambientais provavelmente mudarão bastante a fim de tornar um valor anterior do limiar irrelevante.

Esses parâmetros são ajustados com valores padrão e podem ser auto-ajustados após olhar para dez ou mais quadros. Este é um "ciclo de ensino", embora o método de aprendizagem não precise ser mais sofisticado (pode-se imaginar a tomada

das estatísticas do ruído durante um período de tempo e fazer um cálculo - isto funciona, mas na prática mal vale o esforço).

Análise de Pixel Desviante. A realização descrita 5 gera, dentro de uma imagem, um conjunto de pixels desviantes: pixels para os quais a mudança no valor dos dados excedeu algum limiar automaticamente atribuído. Até este ponto, a posição dos pixels na cena era irrelevante: foi meramente comparado o valor das mudanças em um determinado pixel com o 10 histórico precedente nesse ponto. Isto tinha a vantagem de poder manipular as distribuições de ruído não-uniformes espacialmente.

A questão agora é decidir se eles estão propensos a representar uma mudança genuína na imagem, ou eles são 15 simplesmente uma consequência das flutuações estatísticas no ruído da imagem e na condição ambientais. A fim de auxiliar nisto, é observada a coerência na distribuição espacial dos pixels desviantes.

Correlações espaciais de pixels desviantes. Se em 20 uma imagem forem encontrados, por exemplo, dez pixels desviantes, seria causada uma maior impressão se eles fossem aglomerados uns aos outros ao invés de serem distribuídos aleatoriamente por toda a imagem. Certamente, seria possível computar a probabilidade que se teria dez pixels desviantes 25 distribuídos aleatoriamente se fossem conhecidos os detalhes da distribuição do ruído.

Pontuação de bloco. Aqui apresentamos uma realização de um método simples para avaliar o grau de aglomeração dos pixels desviantes ao atribuir uma pontuação a

Pixel desviante	Pontuação 2
Cada vizinho horizontal ou verticalmente unido	Pontuação 2

Cada vizinho diagonalmente unido	Pontuação 1
----------------------------------	-------------

cada pixel desviante dependendo de quantos de seus vizinhos são eles próprios desviantes.

Uma série de padrões 3x3, com as pontuações atribuídas ao pixel central, é mostrada nos painéis de 5 "Pontuações de Pixels" da Figura 16.

A pontuação sobe rapidamente à medida que o número de vizinhos aumenta, embora pareça, à primeira vista, haver algumas anomalias ligeiras onde um padrão parece pontuar menos do que algum outro padrão que se pode considerar menos 10 significativo. Uma cruz horizontal-vertical de 5 pixels pontua 10, sendo que uma diagonal de 6 pixels pontua somente 9 (padrões 1 e 3 na última fileira).

A situação é resolvida quando se olha a pontuação total padrão, isto é, a pontuação total para todos os blocos 15 desviantes em uma dada região. O painel de "Pontuações de Padrões Especiais" da Figura 16 ilustra as pontuações de pixels desviantes totais em alguns blocos de 3x3, onde se supôs que o bloco de 3x3 está isolado e não tem nenhum pixel desviante adjacente. Há um reforço mútuo não-linear das 20 pontuações de bloco e desse modo a pontuação da telha é impulsionada se o padrão do bloco dentro da região de 3x3 for embalado firmemente.

Em uma realização, os blocos são pesados para favorecer as estruturas horizontais, verticais ou diagonais 25 de pontuação na imagem. Este é o primeiro estágio de classificação do padrão. Claramente, esse processo pode ser executado hierarquicamente: a única limitação quanto a isso é que ao fazer é duplicado o requisito para recursos computacionais.

30 Como um comentário final, deve-se observar que a Imagem sinótica dos pixels desviantes não precisa armazenar

as pontuações de pixels: estas podem ser sempre recalculadas sempre que necessário, contanto que as posições dos pixels desviantes sejam conhecidas. Desse modo, a Imagem Sinótica que indica os pixels desviantes é um bitmap simples de plano de um bit: igual a 1 somente se o pixel correspondente for desviante, ou então 0.

É isto o que torna a busca por Dados sinóticos para mudanças de retrato tão rápida.

Vetores de Movimento

10 O cálculo de vetores de movimento é uma parte essencial de muitos algoritmos de compressão e de algoritmos de reconhecimento de objeto. No entanto, não é necessário utilizar os vetores de movimento para a compressão, a menos que níveis extremos de compressão sejam requeridos.

15 São utilizados vetores de movimento para identificar e acompanhar objetos em cena. O método utilizado é novo pelo fato de que não é baseado em bloco nem baseado em correlação. O método se beneficia da utilização da técnica de substituição do kernel de ondeletas (parágrafos da página 48, linha 20 a página 49, linha 5) que, até uma extensão suficiente, elimina as variações sistêmicas na iluminação de fundo. (Os problemas com a iluminação de fundo são bem conhecidos por serem um problema com cálculos de fluxo óptico.)

25 A presente descrição aplica-se aos componentes $\{^jSS\}$ da transformação de ondeletas de kernel substituídos. Para cada nível de ondeletas é produzido o logaritmo dos valores de pixel em cada componente $\{^jSS\}$. A fim de evitar valores nulos e negativos (os últimos podendo ocorrer em 30 consequência da transformação de ondeletas) é adicionado um deslocamento constante dependente de nível para os valores de pixel de modo que todos os valores sejam estritamente positivos.

$${}^j\rho = \ln({}^j\kappa + {}^jg), \quad {}^jg \in \{{}^jSS\}$$

20.1

Todas as imagens utilizadas no cálculo apresentam os mesmos deslocamentos. Os valores de pixel logarítmicos são mantidos como números de ponto flutuante, mas para a questão 5 do cálculo da velocidade, podem ser novamente escalados para números inteiros assinados de 4 ou 5 bits.

A fim de avaliar os derivados de tempo de ${}^j\rho$ que se fazem necessários $\{{}^jSS\}$ em três períodos de tempo: a época atual e a época dos quadros precedentes e seguintes. Serão 10 denotados os valores dos dados nestes períodos com subscritos $-1, 0$ e $+1$. Desse modo,

$$\dot{\rho}_0 = (\rho_1 - \rho_{-1})$$

20.2

e

$$\ddot{\rho}_0 = \frac{1}{2}(\rho_1 - 2\rho_0 + \rho_{-1})$$

20.3

15 Para cada um desses campos são computados campos novos e altamente suavizados

$$\phi(x) = \sum_{vizinhos-i} w_i \dot{\rho}_0(x_i)$$

20.4

e

$$\phi(x) = \sum_{vizinhos-i} w_i \ddot{\rho}_0(x_i)$$

20.5

20 Os fatores de peso w_i são os mesmos para ambas as equações. Os pesos são escolhidos de modo que esses campos potenciais sejam soluções aproximadas da equação de Laplace com fontes que são os derivados da primeira e da segunda vez de ρ , a densidade logarítmica.

25 O campo da velocidade é calculado ao utilizar gradientes espaciais desses potenciais em todas as escalas da transformação de ondeletas.

Observe que em baixas taxas de quadros, o primeiro campo de derivado, ϕ , pode produzir um resultado nulo ainda

que haja uma intrusão. Isto se dá porque os campos da imagem em um ou outro lado podem ser os mesmos se a intrusão ocorrer somente no único quadro atual. No entanto, este seria fortemente selecionado no segundo campo do derivado, ϕ .

5 Por outro lado, um alvo que se move uniformemente lento poderia fornecer um segundo campo nulo de derivado, ϕ , mas este seria fortemente selecionado no primeiro campo de derivado, ϕ .

10 Observe que ambos os campos estão propensos a ser iguais a zero ou próximos de zero onde a análise de pixel desviante não mostra nenhuma mudança. Deve haver uma mudança a fim de medir uma velocidade!

COMPRESSÃO E ARMAZENAMENTO

15 Dados codificados por ondeletas. Neste estágio, o fluxo de dados é codificada como um fluxo de dados de ondeletas, ocupando mais memória do que os dados originais. A vantagem da representação de ondeletas é que esta pode ser consideravelmente comprimida. No entanto, o trajeto para a compressão substancial que retém alta qualidade não é de modo nenhum direto: uma série de técnicas tem que ser combinada.

20 Estrutura de dados. A Figura 17 resume os elementos do processo de compressão de dados. O fluxo dos dados de imagem original consiste em um conjunto de imagens $\{F_i\}$. Estas são construídas em uma seqüência de execução de moldes $\{T_i\}$ contra a qual várias comparações serão feitas. A partir desses dois fluxos, imagens e moldes, um outro fluxo é criado - um fluxo da diferença de retratos $\{D_i\}$.

30 As diferenças são diferenças entre quadros vizinhos ou entre quadros e um molde selecionado. Por "vizinhos" não se insiste que o vizinho seja o quadro predecessor: a comparação pode ser feita com uma retardaçāo de tempo que depende da taxa do quadro e de outros parâmetros do fluxo de

imagens.

Para uma discussão sobre a variedade de moldes possíveis vide os parágrafos da página 31, linha 4 et seq. e da página 51, linha 23 et seq. Vide também os parágrafos da 5 página 31, linha 4 e da página 51, linha 4, a respeito de alternativas para a utilização do "quadro atual". A discussão pode continuar com referência aos quadros e aos moldes sem perda de generalidade, reconhecendo que existem essas outras realizações possíveis do princípio.

10 É feita referência ao parceiro no processo de diferenciação como uma Imagem de referência $\{R_j\}$. Em outras palavras, R_j poderia ser T_i ou um dos F_i .

15 O objeto da compressão é o fluxo de dados que consiste nos dados $\{D_i\}$ e $\{R_j\}$. Ambos esses fluxos são transformadas por ondeletas utilizando uma ondeleta apropriada ou, tal como neste caso, um conjunto de ondeletas. As ondeletas podem ser de ponto flutuante ou inteiro, ou uma mistura de ambos. Simbolicamente, pode-se escrever:

$$F_k = R_i + D_k \quad 21.1$$

20 É uma questão importante a respeito de quanto do D_k deve ser utilizado com um dado R_j . A princípio, seria necessária somente uma imagem de referência, R_0 . No entanto, uma seqüência bastante longa seria desvantajosa, uma vez que (a) D_k se tornaria maior, sendo que quadros futuros deveriam 25 diferir mais da referência e (b) a descompressão de um D_k posterior deveria envolver a manipulação de uma seqüência muito longa de dados.

Por sua própria natureza, o $\{D_i\}$ individual irá comprimir muito mais do que os quadros da referência $\{R_j\}$. 30 Essa situação pode ser ajudada ao diferenciar os $\{R_j\}$ entre si e então ao representar a seqüência $\{R_j\}$ como uma nova seqüência $\{R_j, \{\delta_k\}\}$ de modo que

$$R_k = R_i + \delta_k$$

21.2

Por causa da similaridade prévia dos elementos da seqüência $\{R_j\}$, δ_k pode ser representado em menos bits do que R_k . A compressão de $\{R_j\}$ é um fator central na determinação 5 da qualidade das imagens restauradas. A compressão da seqüência de $\{\delta_k\}$ deve ser feita quase sem perda, uma vez que as perdas são equivalentes à diminuição da qualidade de $R_k = R_j + \delta_k$ restaurado. O fluxo de dados a ser comprimida pode ser representada como

$$10 \quad \{ \{R_i, D_i, D_{i+1}, \dots, D_{i+m-1}\}, \{\delta_k, D_k, D_{k+1}, \dots, D_{k+m-1}\}, \dots \}, k=m+i$$

A Figura 17 mostra esquematicamente como a diferenciação é organizada.

O estágio final consiste em tomar a transformação de ondeletas de tudo que é requerido para fazer o fluxo de 15 dados comprimida:

$$R_k \rightarrow W_k \quad 22.3a$$

$$D_k \rightarrow V_k$$

e, se forem reorganizados os quadros de referência:

$$\delta_k \rightarrow \omega_k \quad 22.3b$$

20 O fluxo de transformação de ondeletas é então

$$\{ \{W_i, V_i, V_{i+1}, \dots, V_{i+m-1}\}, \{\omega_k, V_k, V_{k+1}, \dots, V_{k+m-1}\}, \dots \}, k=m+i$$

para algum comprimento de ciclo, m . Observe que nenhuma compressão ocorreu ainda.

25 Cada bloco de dados no fluxo de dados de ondeletas consiste em uma série de disposições de coeficientes de ondeletas:

$$v_j = \{^1Q_j, ^2Q_j, \dots, ^kQ_j\} \quad 23.4$$

onde

$$^nQ_j = \{^nSS, ^nDS, ^nSD, ^nDD\} \quad 23.5$$

30 é a disposição da transformação de ondeletas no nível N e, do mesmo modo, para as transformações W_i e ω_k . das imagens de

referência e suas diferenças. A menor dessas disposições, que aparece como o nível K de ondeletas, contém uma versão pequena da imagem: a também denominada "kernel de ondeletas". Na presente notação, o kernel de ondeletas é

5 Kernel de ondeletas de dados = KSS 23.6

10 Compressão. As transformações de cada um dos tipos diferentes de quadro, os quadros de referência R_i , a diferença de quadros D_i ou as referências diferenciadas δ_i , requerem seu próprio tratamento especial a fim de maximizar a eficácia da compressão ao mesmo tempo em que se mantém a qualidade elevada da imagem.

15 Aqui são recordados apenas os princípios genéricos: que o processo consiste na determinação de um limiar abaixo do qual os coeficientes serão ajustados a zero de alguma maneira apropriada, um método para quantizar os coeficientes restantes e, finalmente, uma maneira de representar ou de codificar eficientemente esses coeficientes.

20 Codificação adaptável. Também é recordado que regiões diferentes dos planos de ondeletas podem ter um limiar e quantização diferentes: cada região dos dados que mantém valores particulares de limite e quantização é definida por uma máscara. A máscara reflete o índice dos dados e é codificada com os dados.

25 Supondo que uma parte da imagem é identificada como sendo de interesse especial, talvez em virtude de seu movimento ou simplesmente porque há um detalhe delicado presente. É possível, para essas áreas de interesse especial, escolher um limiar inferior e um grau mais delicado de quantização (mais níveis). Uma tabela diferente de códigos de 30 coeficientes é produzida para essas áreas de interesse especial. Pode-se ainda utilizar os códigos mais curtos para os valores mais populosos; o truque é manter duas tabelas. Junto com as duas tabelas, também é necessário manter dois

valores do limiar e dois valores do fator de escala da quantização.

Thresholding. O thresholding é uma das ferramentas principais do controle da quantidade de compressão. Em algum nível, a limitação remove o que poderia ser considerado como ruído, mas à medida que o nível do limiar se eleva e mais coeficientes são zerados, as características da imagem ficam comprometidas. Uma vez que os componentes j_{SD} , j_{DS} e j_{DD} da matriz de transformação de ondeletas medem os aspectos da curvatura dos dados de imagem, são as partes de baixa curvatura da escala de pixels da imagem que sofrem primeiramente. Certamente, as imagens de ondeletas comprimidas têm uma aparência "vítreia" quando o thresholding foi intenso demais.

A eliminação dos componentes j_{SD} , j_{DS} e j_{DD} da matriz de transformação de ondeletas resulta em $j^{-1}SS$ de uma imagem, que é simplesmente uma ampliação suave do componente j_{SS} e fazer isto em mais de um nível produz imagens descaracterizadas.

A regra do polegar é que os níveis mais elevados (disposições menores) das ondeletas devem ser preservados com cuidado, sendo que os níveis inferiores (disposições maiores) podem ser dizimados sem muitos danos percebidos à imagem se a limitação for efetuada cuidadosamente.

Quantização. A quantização dos coeficientes de ondeletas também contribui para o nível de compressão ao reduzir o número de coeficientes e tornar possível codificá-los eficientemente. Idealmente, a quantização deve depender do histograma dos coeficientes, mas na prática esta coloca uma demanda muito elevada nos recursos computacionais. O método mais simples e geralmente eficiente de quantização consiste em escalar novamente os coeficientes e dividir o resultado em planos de bits. Isto é eficazmente uma

quantização de intervalo logarítmica. Se o histograma dos coeficientes for distribuído exponencialmente, este deve ser um método ideal.

Os efeitos da quantização inadequada se fazem 5 sentir particularmente na restauração de áreas lisas da imagem com gradientes pequenos da intensidade: a reconstrução mostra o contorno que pode ser bastante ofensivo. Felizmente, a reconstrução inteligente, utilizando por exemplo, a difusão dos erros, pode eliminar o aparecimento do problema sem 10 danificar outras partes da imagem (vide os parágrafos da página 47, linha 3 e da página 73, linha 6).

O fator de escala do plano de ondeletas deve ser mantido como uma parte do cabeçalho de dados comprimido.

Codificação. Uma vez que a transformação de 15 ondeletas tenha sido submetida a thresholding e quantizada, o número de valores do coeficiente distintos é bastante pequeno (depende do número de valores quantizados) e códigos do tipo Huffman podem ser atribuídos.

A tabela de código deve ser preservada com cada 20 plano de ondeletas. É geralmente possível utilizar a mesma tabela para um grande número de quadros da mesmo fluxo de vídeo: uma técnica de compressão de cabeçalho adequada irá manipular isto eficientemente, reduzindo desse modo as despesas gerais do armazenamento de diversas tabelas por 25 quadro. A unidade de armazenamento são os grupos de ondeletas comprimidos (vide abaixo) e é possível que grupos inteiros utilizem a mesma tabela.

Empréstimo de Bits. A utilização de um conjunto 30 muito pequeno de números para representar os valores dos dados tem muitos inconvenientes e pode ser seriamente nociva à qualidade da imagem reconstruída. A situação pode ser ajudada consideravelmente por qualquer técnica de uma variedade de técnicas conhecidas. Em uma realização desse

processo, os erros da quantização de um ponto de dados são levados a se difundir através dos pontos de dados vizinhos, conservando desse modo tanto quanto possível o índice das informações totais da área local. A redistribuição uniforme 5 do restante ajuda a suprimir o contorno nas áreas de iluminação uniforme. Além disso, a redistribuição criteriosa desse resto onde há características irá ajudar a suprimir os danos para o detalhe da imagem e irá produzir desse modo resultados visuais consideravelmente melhores. Isto reduz o 10 contorno e outros tais artefatos. Isto é denominado "emprestimo de bits".

Validação e criptografia. Deseja-se saber, quando se vê para uma imagem, se é de fato a mesma imagem que foi capturada, comprimida e armazenada. Este é o processo de 15 validação da imagem.

Também é possível restringir o acesso aos dados de imagem e desse modo criptografar os coeficientes de reconstrução, convertendo os mesmos em valores corretos se o usuário fornecer uma chave válida de descriptografia.

20 Ambos esses problemas podem ser resolvidos ao mesmo tempo ao criptografar a tabela de coeficientes quantizados de ondeletas. Se o acesso não for restrito, uma chave geral é utilizada com base nos próprios dados do fluxo. Se os dados forem autênticos, os dados serão corretamente descomprimidos. 25 Uma segunda chave é utilizada se o acesso aos dados for restrito.

Acondicionamento. Os dados de imagem comprimidos vêm em "pacotes" que consistem em um quadro ou molde de referência comprimido ou seguido por um conjunto de quadros 30 que são derivados dessa referência. Isto é denominado "Grupo do Quadro". Isto é análogo a um "Grupo de Retratos" em outros esquemas de compressão, exceto pelo fato que aqui o quadro de referência pode ser uma construção inteiramente artificial, e

desse modo é preferível utilizar um nome ligeiramente diferente. Este é o menor pacote que pode ser utilmente armazenado.

5 O grupo de transformações de ondeletas a partir dessas imagens que compreendem um grupo de quadros pode ser do mesmo modo chamado de um grupo de ondeletas.

É útil o acondicionamento de diversos tais Grupos de Quadros em um pacote maior que será denominado, na falta de um termo melhor, um "Bocado de Dados" e o pacote dos dados 10 comprimidos que derivam deste um "bocado de dados comprimido".

Os grupos de quadros podem estar tipicamente na ordem de um megabyte ou menos, sendo que o tamanho conveniente do bocado pode ser de diversas dezenas de 15 megabytes. A utilização de elementos de armazenamento maiores torna o acesso aos dados a partir dos drives de disco mais eficiente. Também é vantajoso ao gravar para uma mídia removível tal como DVD+RW.

DADOS SINÓTICOS

20 Compressão e criptografia. Os dados sinóticos consistem em um conjunto de imagens dos dados, sendo que cada um deles resume algum aspecto específico da imagem original da qual foi derivado. Uma vez que os aspectos que são resumidos são geralmente apenas uma parte pequena das 25 informações contidas dentro da imagem, os dados sinóticos serão comprimidos até um tamanho que seja substancialmente menor do que a imagem original. Por exemplo, se parte dos dados sinóticos indicar as áreas da imagem onde o movimento em primeiro plano foi detectado, os dados em cada pixel podem 30 ser representados por um único bit (detectado ou não). Haverá geralmente muitos zeros a partir das áreas onde nada está acontecendo no primeiro plano.

Os dados sinóticos são comprimidos sem perda.

Acondicionamento. O tamanho dos dados de imagem sinóticos é bem menor do que os dados originais, mesmo que os dados originais estejam limpos e comprimidos.

Para fins de conveniência de acesso, os dados sinóticos são acondicionados exatamente da mesma maneira que os dados de ondeletas comprimidos. Todas as imagens sinóticas que se relacionam às imagens em um grupo de quadro são acondicionadas em um grupo de imagem sinótico e esses grupos são então acondicionados em bocados que correspondem precisamente aos Bocados de dados de ondeletas comprimidos.

BANCO DE DADOS

Linha do Tempo. Uma vez que os dados originais vêm em um fluxo, é apropriado endereçar os dados de todas as formas em termos de um identificador de quadro ou um identificador de tempo ou ambos em que o quadro foi capturado.

Os dados comprimidos são armazenados nos Bocados que contêm muitos grupos de quadro. O banco de dados mantém uma lista de todos os bocados disponíveis junto com uma lista dos conteúdos (os grupos de quadros) de cada bocado e uma lista dos índices de cada grupo de quadros.

A lista mais simples do banco de dados para um item de dados armazenado consiste em um identificador construído com um número de identificação e tempos de início e fim do item de dados armazenado, quer seja ele um bocado, um grupo de quadro ou simplesmente um quadro. A manutenção das informações sobre o tamanho em bytes do elemento de dados também é útil para uma recuperação eficiente.

A Figura 18 mostra como há uma correspondência de um-para-um entre os dados sinóticos da imagem e os dados de ondeletas comprimidos. A linha do tempo pode ser utilizada para acessar as imagens sinóticas para análise ou os dados de ondeletas comprimidas para visualização.

Observe que não é necessário manter os dados sinóticos e os dados de ondeletas comprimidos no mesmo lugar.

Divisão lógica do tempo. Uma vez que uma aplicação principal deste procedimento é a gravação de imagens digitais com potencialidade de análise pós-gravação, faz sentido 5 armazenar os dados em uma base do calendário.

Imagens sinóticas. As imagens sinóticas são geralmente imagens de resolução variável de um plano de um bit. Não faz nenhum sentido exibir as mesmas, mas elas são 10 muito eficientes para a busca.

Dados de imagem comprimidos. Os dados de imagem comprimidos são os dados finais que o usuário verá em resposta a uma pergunta.

Esses precisam ser armazenados no mesmo repositório 15 que os dados sinóticos, mas eles têm que ser denominados pelo banco de dados e pelos dados sinóticos.

ARMAZENAMENTO DE DADOS

Bancos de dados. Finalmente, os dados têm que ser armazenados em algum tipo de mídia de armazenamento, quer 20 seja um disco rígido ou um DVD ou qualquer outro dispositivo.

No nível mais simples, os dados podem ser armazenados como uma parte do sistema de arquivamento do próprio computador. Nesse caso, é útil armazenar os dados no formato lógico de calendário. Cada dia uma pasta é criada 25 para esse dia e os dados são armazenados em uma base de hora em hora em uma pasta com base em hora. (A utilização do padrão de tempo UTC evita os caprichos associados com as mudanças nos relógios devido ao horário de verão).

Em um nível mais elevado, o próprio banco de dados 30 pode ter seu próprio sistema de armazenamento e acesso aos elementos de dados armazenados em termos de suas próprias convenções de armazenamento.

O mecanismo de armazenamento é independente do

sistema de perguntas utilizado: a interface do banco de dados deve prover acesso aos dados que foram pedidos, não importando o mecanismo de armazenamento e se o mesmo foi armazenado.

5 Mídia. As mídias de armazenamento do computador são bastante diversas. A classificação mais simples aqui é entre as mídias removível e não-removível. Os exemplos de mídia não-removível podem ser discos rígidos, embora alguns discos rígidos sejam removíveis.

10 A diferença prática é que as mídias removíveis devem manter seus próprios bancos de dados: isto as torna não somente removíveis, mas também móveis. O controle das mídias removíveis dessa maneira não é sempre simples; depende do banco de dados que é utilizado e se o mesmo tem essa 15 facilidade. As mídias removíveis também devem manter cópias da auditoria que descreve como, quando e de onde esses dados foram tirados.

RECUPERAÇÃO DE DADOS

20 A Figura 19 mostra as etapas na recuperação de dados e o Ciclo de análise. Em resposta a uma pergunta do usuário, os dados sinóticos são procurados por correspondências à pergunta. Com paradas bem-sucedidas, os 25 eventos são construídos e adicionados a uma lista de eventos que é retornada ao usuário. Os dados principais da imagem não são tocados até que o usuário deseje visualizar os eventos na lista. A Figura 18 descreve como os dados armazenados principais são associados com os dados sinóticos.

30 Com base no que é apresentado, o usuário pode refinar as buscas até que uma lista aceitável de eventos seja encontrada. A lista de eventos selecionada pode ser convertida em um formato de armazenamento diferente, pode ser anotada, acondicionada e exportada para utilização futura.

PERGUNTAS

Critérios de busca. Este tipo de sistema de armazenamento de dados, em uma realização particular, permite pelo menos dois tipos de busca de dados:

5 Busca por hora e data: O usuário solicita os dados capturados em um dado instante de um fluxo de vídeo escolhida. Se, nos dados sinóticos, houvesse um evento que tivesse ocorrido perto da hora especificada, isto seria informado ao usuário.

10 Busca por evento ou objeto: O usuário especifica uma área da cena um fluxo de vídeo escolhida e a um intervalo de tempo de busca onde um evento particular pode ter acontecido. Os dados sinóticos para esse intervalo de tempo são procurados e todos os eventos encontrados são informados ao usuário. A busca é bastante rápida (diversas semanas de dados podem ser procuradas em um minuto) e desse modo o usuário pode buscar eficientemente extensões de tempo enormes.

20 Deve-se recordar que a busca por evento dentro dos dados sinóticos não está predicada em nenhum critério de seleção pré-gravado.

25 Busca de fluxos múltiplos. Listas de dados sinóticos de fluxos múltiplos podem ser construídas e combinadas de acordo com a lógica ajustada pelo usuário. O mecanismo para permitir essa lógica é até a interface do usuário; a busca produz simplesmente uma lista de todas as batidas em todos os fluxos solicitados e então combina os mesmos de acordo com os critérios lógicos ajustados pelo usuário.

30 O usuário pode, por exemplo, querer visualizar o que estava acontecendo em outros fluxos de vídeo em resposta a uma parada em um de seus fluxos da busca. O usuário pode desejar visualizar somente aqueles fluxos que pontuaram paradas ao mesmo tempo ou dentro de algum determinado

intervalo de tempo. O usuário pode desejar visualizar as paradas em um fluxo que era contingente às paradas que estão sendo visualizadas em outros fluxos.

Eventos - o resultado da pergunta bem-sucedida. O resultado de uma pergunta bem-sucedida deve ser a apresentação de um clipe de filme que o usuário pode examinar e avaliar. O clipe de filme deve mostrar um número suficiente de quadros de vídeo para permitir que o usuário faça essa avaliação. Se a pergunta envolveu múltiplos fluxos de vídeo, a exibição deve envolver uma regravação de vídeo sincronizada desses fluxos.

A técnica utilizada aqui é a construção de uma lista de paradas bem-sucedidas nos Dados Sinóticos e acondicionar as mesmas com outros quadros em filmes pequenos ou "Eventos". O usuário vê somente eventos, não quadros individuais, a menos que isso seja solicitado.

BUSCA POR DADOS SINÓTICOS

Paradas. A busca por Dados Sinóticos atinge a busca de uma seqüência de imagens por características particulares. A vantagem aqui é que os dados são geralmente de um único plano de bits e apenas é preciso buscar uma área nomeada pelo usuário pelos bits que são ativados. Este é um processo extremamente rápido que pode ser adicionalmente acelerado se o mapa de Dados Sinóticos for apropriadamente codificado.

As paradas podem vir de fluxos múltiplos de vídeo, combinando os resultados de buscas de fluxos múltiplos com o conjunto de lógica pela pergunta.

As paradas podem ser modificadas de acordo com os valores de uma variedade de outros atributos que estão disponíveis direta ou indiretamente a partir dos Dados Sinóticos tais como a pontuação total do bloco ou a direção do movimento ou o tamanho.

Exibição. Após encontrar as paradas dentro das

séries de Dados Sinóticos, as paradas de Dados Sinóticos têm que ser construídas em um Evento que possa ser exibido. Há então duas opções para a exibição e avaliação. (1): Mostrar os Excertos de Filmes se os mesmos foram armazenados. (2):

5 Obter os dados completos.

10 Velocidade. A busca por Dados Sinóticos pode ser muito rápida porque a análise já foi feita. Além disso, o tamanho do conjunto de dados sinóticos é geralmente muitas ordens de magnitude menor do que os dados originais. A parte mais lenta da busca é, de fato, o acesso aos dados a partir 15 da mídia de armazenamento.

20 Isto é especialmente verdadeiro se a mídia de armazenamento for um DVD (velocidade de acesso de aproximadamente 10 megabytes por segundo) em cujo caso é normalmente útil o armazenamento do banco de dados sinóticos inteiro na memória. A execução de tarefas múltiplas inteligente da interface do usuário pode fazer isto de uma maneira fácil: a primeira busca será o momento de ler os dados sendo que as buscas seguintes serão quase instantâneas.

25 As buscas em uma rede são extremamente eficientes uma vez que os dados sinóticos são mantidos em um disco rígido com acesso local rápido e apenas os resultados têm que ser transmitidos ao cliente.

RECUPERAÇÃO DE DADOS ASSOCIADOS

30 Definição e construção de eventos. Um evento é uma coleção de quadros de dados consecutivos de uma ou mais fontes de dados. Pelo menos um dos quadros que constituem esta coleção, o quadro chave, irá satisfazer algum critério especificado que foi formulado como uma pergunta do usuário endereçada aos dados sinóticos. A pergunta pode concernir atributos tais como o tempo, a posição, a cor em alguma região, a velocidade do movimento, e assim por diante. É feita referência a um resultado bem-sucedido à pergunta como

uma "parada".

Considerar uma realização do processo em que, se houver uma única "parada", o usuário irá querer visualizar alguns segundos de vídeo antes da "parada" e alguns segundos em seguida a fim de apreciar a ação. Se duas ou mais paradas ocorrem dentro de alguns segundos entre si, elas também podem ser combinadas para formar um clipe mais longo do evento. Desse modo, nesta realização, as paradas sucessivas são combinadas no mesmo clipe se o intervalo entre as paradas for menor do que a soma dos tempos de pré e pós-parada especificados pelo usuário.

É possível que um único quadro chave de um fluxo de dados represente um evento que cobre fluxos múltiplos: dessa maneira, todos os fluxo de dados associadas com o(s) quadro(s) chave podem ter uma referência cruzada. Um evento pode compreender uma pluralidade de quadros de dados antes e depois do quadro chave sendo que eles mesmos não satisfazem ao critério de quadro chave (tal como nas seqüências de imagem de pré e pós-alarme).

A Figura 20 mostra como os dados são adquiridos, processados, armazenados e recuperados. Em resposta a uma pergunta chave os quadros são encontrados e os eventos são construídos, transpondo esses quadros chave.

Construção do clipe do evento. Cada quadro de dados sinóticos é associado com o quadro original a partir do qual foi derivado nos dados de vídeo originais (Ondeleta comprimida).

Os quadros referidos a um evento, tal como definido pelas paradas nos dados sinóticos, são recuperados a partir do fluxo de dados comprimidos de ondeletas. Eles são validados e descriptografados, caso necessário) e descomprimidos. Em seguida, eles são convertidos em um formato de dados interno que é apropriado para a

visualização.

O formato de dados pode ser um formato de computador (tal como DIB ou JPG) se deve ser visualizado no computador do usuário ou pode ser convertido novamente em um formato de vídeo de CCTV análogo por um chip ou cartão gráfico codificador para a visualização em um monitor de TV.

Análise do evento. Uma vez adquiridos os quadros de vídeo originais para a parada de Dados sinóticos, eles podem ser analisados para a visualização se satisfazem outros critérios que não foram incluídos nos dados sinóticos. Desse modo, os dados sinóticos podem, por causa dos limites de recursos computacionais no momento do processamento, não ter classificado os objetos em pessoas, animais ou veículos. Essa classificação pode ser feita a partir de combinações de quaisquer dados sinóticos que estejam disponíveis para esses fluxos e a partir da imagem armazenada.

Adição de dados de áudio. Quando um evento é gravado ou exportado, pode ser necessário o acesso a todos os canais de áudio que possam acompanhar a seqüência.

O canal de áudio é, do ponto da vista desta discussão, meramente uma outro fluxo de dados e desse modo é acessado e apresentado exatamente da mesma maneira que qualquer outro fluxo.

FLUXO DE TRABALHO

Acesso e validação dos dados. Se os dados forem criptografados, então a interface do usuário deve pedir autorização para descriptografar os dados antes de sua apresentação. Todos os dados gravados no mesmo computador terão o mesmo código de acesso do usuário. Diferentes fluxos podem ter códigos de acesso de fluxo suplementares se tiverem níveis de segurança diferentes.

A validação de dados é feita ao mesmo tempo em que a descriptografia, uma vez que o código de validação de dados

é um resultado quase exclusivo de uma fórmula de verificação de dados construída nos dados de imagem. (É indicado "quase exclusivo" uma vez que o código tem um número finito de bits. Portanto, é concebível, embora astronomicamente improvável, 5 que duas imagens tenham o mesmo código.)

Perguntas repetidas ou refinadas. A interface do usuário tem a opção de repetir um inquérito ou de refinar um inquérito ou ainda combinar o resultado de um inquérito com o resultado de outro em um fluxo de dados inteiramente 10 diferente.

O procedimento de busca dentro dos dados sinóticos é tão rápido que custa pouco para simplesmente executar novamente um inquérito com parâmetros diferentes ou lógicas diferentes. Esta é meramente uma questão de eficiência 15 programática.

Exportação de dados - auditorias. Uma vez que o usuário tem um conjunto de eventos que satisfazem a pergunta, há a necessidade de armazenar esses eventos descobertos de uma maneira tal que eles possam ser utilizados por outros 20 programas ou ser utilizados para finalidades de exibição e de informações.

Um exame sobre como os resultados foram atingidos é publicado junto com a exportação, de modo que o procedimento possa ser novamente executado, caso necessário. (A 25 possibilidade de repetir o resultado de uma busca é às vezes requerida em casos legais).

DADOS EXPORTADOS

Os dados do evento podem ser exportados para qualquer formato de uma série de formatos padrão. A maioria 30 destes consiste em formatos que são compatíveis com o software Microsoft WindowsTM, alguns com Linux. Muitos são baseados nos padrões de MPEG (que não é suportado pelas versões atuais do Windows Media Player!).

Embora a presente invenção tenha sido descrita de acordo com as realizações mostradas, um técnico no assunto irá reconhecer facilmente que poderia haver variações nas realizações e essas variações estariam dentro do caráter e do 5 âmbito da presente invenção. Conseqüentemente, muitas modificações podem ser feitas por um técnico no assunto sem que se desvie do caráter e do âmbito das reivindicações em anexo.

NOTAÇÃO

10 Notação Simbólica

A seguir, para maior clareza, serão utilizados símbolos para denotar dados e imagens de vários tipos.

Dados, Imagens e Operadores

15 Os processos que agem nestas imagens ou combinações dos mesmos serão denotados como operadores. Desse modo, se F denotar um quadro da imagem e N denotar um operador que filtre o ruído, NF denotará o resultado desse processo e $F-NF$ denotará o resíduo a ser identificado como o componente do ruído de F .

20 Os operadores que agem seqüencialmente são tomados para agir da direita para a esquerda. Desse modo, se N_1 e N_2 forem dois operadores que podem agir em um quadro de imagem F , N_2N_1F é o resultado da primeira aplicação de N_1 para F e então N_2 .

25 Os operadores não precisam ser lineares e os operadores não precisam comutar. Em outras palavras, se N_1 e N_2 forem dois operadores que podem agir em um quadro de imagem F , N_1N_2F e N_2N_1F não serão necessariamente a mesma coisa.

30 A dependência de espaço-tempo genérica de um quadro F pode ser denotada pelo símbolo $F(x,t)$, onde x são os dados de imagem bidimensionais do quadro no tempo t .

Também será utilizado um pseudocódigo para mostrar

como estas várias imagens são geradas e relacionadas. Maiores detalhes podem ser encontrados no Apêndice.

Notação

	Interpretação
F_n	n^{th} . quadro bruto na seqüência. O índice n funcionará a partir de valores negativos no passado a valores positivos no futuro. n=0 refere-se ao quadro que está sendo atualmente processado.
\bar{F}_n	n^{th} . quadro processado na seqüência. A sobre-barra pode ser substituída por qualquer uma de uma diversidade de decorações apropriadas: pontos, chapéus, tils, etc., para indicar os resultados de processos diferentes.
F_n^S, F_n^M, F_n^D	Componentes estáticos, Móveis e Dinâmicos do quadro após a análise de resolução do tempo.
$\Delta_n = F_n - F_{n-1}$	Diferença entre dois quadros brutos sucessivos
$\sigma(D)$	Variação do histograma dos dados de imagem D
Mediana (D)	Mediana do histograma da imagem D. O operador da Mediana pode ser igualmente Suave, Filtro, Máscara, Transformação ou qualquer palavra descritiva que descreva a operação. O nome do operador pode ser alternativamente representado por um símbolo em negrito como em S, F, M, T.
$M_n = M(F_n)$	Versão mascarada do n^{th} . quadro
$G(\underline{x}, t)$	Dados gerais dependentes da posição \underline{x} e do tempo, t .
$G^C(\underline{x}, t)$	Um componente específico de G etiquetado com o conteúdo C que descreve a natureza do componente. Os exemplos específicos de componentes incluem o ruído (C='N'), os dados de fundo (C='B') e os dados dinâmicos (C='D').

A notação pode ficar bastante pesada: considera-se

- 5 o caso onde dados gerais são descritos por uma matriz de valores cujo tamanho se deseja indicar especificamente. É preciso dar um passo na simplificação usual ao manter somente os subscritos e sobrescritos necessários, deixando de fora aqueles que podem ser deduzidos a partir do contexto.

GLOSSÁRIOEMPRÉSTIMO DE BITS

O procedimento por meio do qual algumas partes de uma imagem onde um nível mais elevado de fidelidade é 5 necessário podem ser comprimidas para uma melhor qualidade do que outras partes da imagem. De fato, bits são emprestados de uma parte da imagem para melhor representar outras partes. Isto é conseguido durante a codificação de coeficientes de ondeletas antes do armazenamento. Uma tabela especial de 10 códigos do coeficiente é produzida para estas áreas de interesse especial. Pode-se ainda utilizar os códigos mais curtos para os valores mais populosos; o truque consiste em manter duas tabelas. Junto com as duas tabelas também é necessário manter dois valores do limiar e dois valores do 15 fator de escala da quantização.

CDF(2,2)

Um elemento simples de uma ampla classe de ondeletas biortogonais de Cohen, Daubechies e de Feaveau, também conhecida como a ondeleta 5-3 biortogonal uma vez que 20 utiliza cinco pontos para seu filtro de passagem elevada e três pontos para seu filtro de passagem baixa.

IMAGEM ATUAL

A imagem em uma seqüência que seja o foco atual de interesse. Embora esta seja geralmente a imagem mais recente 25 capturada no fluxo, pode ser a penúltima, a antepenúltima ou a última menos n se o processamento da imagem atual depende de uma série de imagens subseqüentes (como pode acontecer se forem estimados derivados de tempo das imagens).

DCD

30 Detecção da Mudança de Dados: a forma geral de VMD. Vide também VMD, Detecção de Movimento de Vídeo.

PIXEL DESVIANTE

Um pixel em uma imagem que é considerado, com base

na análise da série de tempo de seu histórico passado, como tendo um valor que é excepcional com relação ao que esse histórico deve indicar. Os pixels desviantes são definidos em termos do comportamento do tempo em cada ponto e a sua 5 importância é avaliada em termos de sua proximidade relativa em relação aos outros ao pontuar padrões espaciais de pixels desviantes.

DivX

Um formato de arquivo de vídeo muito popular devido 10 à sua capacidade de comprimir segmentos de vídeo longos em tamanhos pequenos ao mesmo tempo em que mantém uma qualidade visual relativamente elevada. DivX utiliza a compressão com perda de MPEG-4 Parte 2: o codec de DivX é inteiramente compatível com o MPEG-4-Advanced Simple Profile. O formato de 15 DivX está agora sujeito às limitações de patente e não é mais uma Fonte Aberta. DivX é inferior ao novo H.264/MPEG-4 AVC, também conhecido como MPEG-4 Parte 10, mas é bem menos intensivo ao cpu.

No domínio público, foi substituído pelo formato de 20 Fonte Aberta conhecido como Xvid.

PRIMEIRO PLANO DINÂMICO

As características na cena que entram ou saem de cena ou executam movimentos substanciais, durante o período da aquisição de dados, compreendem o primeiro plano dinâmico. 25 (Ao contrário dos componentes de fundo estático e estacionário).

Vide também Fundo Estático, Fundo Estacionário.

EVENTO

Um evento é uma coleção de quadros de dados 30 consecutivos de uma ou mais fontes de dados. Pelo menos um dos quadros que constituem essa coleção, o quadro chave, irá satisfazer algum critério especificado (tal como o tempo, a posição, a cor em alguma região, a velocidade do movimento,

etc.). É possível que um único quadro chave de um fluxo de dados represente um evento que cobre fluxos múltiplos: dessa maneira, todos os fluxo de dados associadas com o(s) quadro(s) chave podem ter uma referência cruzada. Um evento 5 pode compreender uma pluralidade de quadros de dados antes e depois do quadro chave sendo que eles mesmos não satisfazem o critério de quadro chave (tal como nas seqüências de imagem de pré e pós-alarme).

Vide também Detecção de Evento de Vídeo

10 IGU

Interface Gráfica do Usuário. Este é um programa de computador, que funciona em um computador, assistente de dados pessoal, telefone móvel etc., que apresenta ao usuário uma vista "em janela" ou "gráfica" de programas e dados 15 disponíveis. Os programas de controles do usuário e os dados dos acessos através de um dispositivo indicador, tais como um mouse e um teclado. A IGU define as facilidades e a funcionalidade com que um usuário pode executar programas e armazenar dados.

20 MÁSCARA DA IMAGEM

Regiões em uma imagem que devem ser protegidas de determinadas operações nos dados de imagem. Desse modo, uma máscara pode ser construída para cobrir as bordas das características em uma imagem de modo que uma operação de 25 suavização não crie características opacas.

MOLDE DA IMAGEM

Uma imagem construída a partir da imagem atual e possivelmente um número de suas predecessoras. A finalidade de tal imagem é enfatizar aspectos específicos de uma imagem 30 e de seu histórico. Um exemplo de um molde pode ser a imagem que consiste unicamente nas bordas da imagem atual. Outra pode ser uma imagem que é alguma média do tempo específico das imagens precedentes. Ao comparar a imagem atual com um

molde especialmente projetado é possível isolar mudanças em aspectos específicos da imagem.

MÁSCARA

Uma máscara da imagem é um mapa da região em uma imagem, sendo que todos os pontos do mesmo dividem alguma propriedade particular. O mapa é em si uma imagem, embora um tanto simplificada, uma vez que descreve geralmente se um ponto na imagem tem essa propriedade particular ou não. Um mapa com dois valores (Sim ou Não) é representado como um plano de um bit. As máscaras são utilizadas para resumir as informações específicas sobre uma ou mais imagem tais como onde há uma cor vermelha dominante, onde há um movimento em uma direção particular, e assim por diante. A máscara é, portanto, um mapa juntamente com uma lista de atributos e seus valores que definem o conteúdo de informações do mapa.

As informações de uma ou mais máscaras vão para os Dados Sinóticos em construção para o fluxo de dados.

As máscaras também podem ser utilizadas para proteger partes particulares de uma imagem dos processos que poderiam ser destruídos se não fossem mascarados.

Vide também Dados Sinóticos

MPEG

O Grupo de Experts em Retratos em Movimento: uma organização que existe desde 1988. Eles são responsáveis pelo desenvolvimento de padrões para a representação codificada de sinais digitais de áudio e de vídeo. Os padrões resultam em formatos do arquivo de dados do tipo MPEG-1, MPEG-2, MPEG-4 e MP3. A documentação dos padrões não está livremente disponível e a utilização do padrão está sujeita a acordos de licença. MPEG não é realmente um padrão de fonte aberta.

RUÍDO

O componente do ruído é essa parte dos dados de imagem que não representa exatamente nenhuma parte da cena.

Geralmente, surge dos efeitos instrumentais e serve para diminuir uma apreciação clara dos dados de imagem. Geralmente, pensa-se no componente do ruído como sendo não-correlato com ou ortogonal aos dados de imagem (por exemplo: 5 "neve" de vídeo superposta), mas este não é necessariamente o caso, uma vez que o ruído pode depender diretamente da natureza local da imagem.

DECOMPOSIÇÃO PIRAMIDAL

A redução da escala e a decomposição sucessivas de 10 dados n-dimensionais em versões re-escaladas de resolução inferior das próprias versões seguem os preceitos da decomposição de Resolução Múltipla de Mallat. Os erros na reconstrução de um conjunto de dados mais elevado a partir da resolução inferior de seu predecessor também são armazenados. 15 Um exemplo disto é a transformação de ondeletas, mas nem todas as decomposições piramidais são com base em ondeletas: sendo a transformação do número médio piramidal não-linear um exemplo importante.

MOVIMENTO ALEATÓRIO DA CÂMERA

20 O movimento aleatório e limitado da câmera faz com que a seqüência de imagens percebida se agite, tendo por resultado a falsa detecção do movimento. O movimento aleatório da câmera pode ser superposto ao movimento sistêmico da câmera, que no caso se vê como desvios 25 aleatórios a partir das mudanças de outra maneira suaves no aspecto da imagem.

Vide também Movimento Sistêmico da Câmera

IMAGEM DE REFERÊNCIA

Uma imagem, possivelmente artificial, contra a qual 30 irá se decidir há algum evento significativo ocorrendo na cena atual. As imagens artificiais podem ser construídas a partir de outras imagens que foram tiradas no passado (uma média seria um exemplo de tais). Também é possível (e

certamente desejável se puder ser feito) incluir as imagens subseqüentes a essa que está sendo atualmente analisada.

Vide também Molde.

DESLOCAMENTO DE CENA

5 Um momento em um fluxo de vídeo quando a vista do componente de fundo estático da cena muda tanto que as cenas imediatamente antes e depois do deslocamento não se correlacionam espacialmente.

Vide também Marcador de Cena

10 MARCADOR DE CENA

Nota onde há uma mudança significativa da cena. Tal mudança se dá geralmente pelo movimento sistêmico de uma câmera que começa uma seqüência com visualização diferente ou a uma mudança na câmera que fornece a seqüência. Pode, no 15 entanto, marcar um lugar onde, por exemplo, as luzes devem ser jogadas.

Vide também Deslocamento de Cena

PENEIRAR

O verbo "peneirar" é sinônimo de "investigar". As 20 definições no dicionário são "examinar a fim de testar a adequabilidade", "verificar e classificar com cuidado" e "distinguir e separar". Uma Peneira (substantivo) é um dispositivo que permite que se peneire. Neste documento, o substantivo é utilizado no sentido do conceito matemático 25 exemplificado pela Peneira de Erastóstenes, que é um algoritmo para distinguir e separar todos os números primos até um dado número, N. Desse modo, apresentamos um processo por meio do qual é possível distinguir e separar os atributos em fluxo de dados.

30 Vide também Investigar, Peneira Espacial, Peneira Temporal

INVESTIGAR

Segundo o dicionário, isto significa "investigar

especialmente para classificar o que é útil ou de valor <investigou a evidência> -- utilizado freqüentemente com através <investigar através dos sinais escolhidos pelo telescópio Arecibo". Também significa "Empregar uma 5 peneira", "Distinguir como se estivesse separando com uma peneira" e "Fazer um exame cuidadoso". Vide Peneira. (A utilização do dicionário não deve ser confundida com o acrônimo SIFT, que foi adotado para "Scale-invariant feature transform": um algoritmo para a visualização pelo computador 10 que extrai as características distintivas das imagens, para ser utilizado nos algoritmos para tarefas como combinar vistas diferentes de um objeto ou cena (por exemplo, para a visualização estereofônica) e Reconhecimento do Objeto).

Vide também Peneirar

15 INSTANTÂNEO

Um "Instantâneo" é uma imagem única tirada de um evento que fornece uma pequena vista de thumbnail de um quadro da ação. Tais quadros podem fazer parte de um trecho de filme, ou podem ser quadros especialmente construídos que 20 são mantidos na Sinopse.

PENEIRA ESPACIAL

Um algoritmo ou dispositivo que extrai e preserva as características que podem estar presentes em um sinal ou série de sinais que variam espacialmente. A transformação 25 Hough é uma peneira espacial.

Vide também Peneirar, Peneira Temporal.

FUNDO ESTÁTICO

Consiste nos elementos da cena que são fixos e que mudam somente em virtude das mudanças na resposta, na 30 iluminação ou na oclusão da câmera ao mover objetos. Um fundo estático pode existir mesmo quando uma câmera está girando em movimento panorâmico, inclinando ou fazendo zoom. A repetição de uma cena em momentos diferentes irá mostrar os mesmos

elementos de fundo estático. Os edifícios e as estradas são exemplos de elementos que constituem o fundo estático.

Vide também Fundo Estacionário, Primeiro Plano Dinâmico

5 FUNDO ESTACIONÁRIO

Consiste nos elementos da cena que são fixos no sentido que a repetição de uma cena em momentos diferentes irá mostrar os mesmos elementos em formas ligeiramente deslocadas. Os ramos e as folhas que se movem em uma árvore 10 são exemplos de componentes de fundo estacionário. O movimento é localizado e limitado e a sua variação de tempo pode ser episódica. Os reflexos em uma janela devem entrar nesta categoria.

Vide também Fundo Estático, Primeiro Plano Dinâmico

15 DADOS SINÓTICOS

Os dados sinóticos consistem em um conjunto de imagens dos dados, sendo que cada uma resume algum aspecto específico da imagem original a partir da qual foi derivado.

MOVIMENTO SISTÊMICO DA CÂMERA

20 As câmeras podem ter a facilidade de girar em movimento panorâmico, inclinar e fazer zoom sob o controle de um operador ou de um programa. Sob tais circunstâncias, é observado que um deslocamento sistêmico na cena pode ser modelado através de uma série de transformações afins. Se o 25 movimento for muito rápido, cenas consecutivas podem ter pouca relação ou relação nenhuma entre si.

Vide também Movimento Aleatório da Câmera.

MOLDE

Uma imagem, possivelmente artificial, contra a qual 30 será decidido se há algum evento significativo ocorrendo na cena atual. As imagens artificiais podem ser construídas a partir de outras imagens que foram tiradas no passado (uma média seria um exemplo de tais). Também é possível (e

certamente desejável se puder ser feito) incluir as imagens subsequentes a essa que está sendo atualmente analisada. Vide também Imagem de Referência.

PENEIRA TEMPORAL

5 Um algoritmo ou dispositivo que extrai e preserva as características que podem estar presentes em um sinal ou série de sinais. A banda de passagem de um filtro é uma peneira de freqüência que seleciona no índice da freqüência de sinal.

10 Vide também Peneira, Peneira Espacial

THUMBNAIL

15 Um retrato imóvel pequeno que mostra a cena onde a atividade foi detectada. Essas imagens pequenas podem ser armazenadas como um fluxo de dados paralela ou como parte dos Dados Sinóticos. Elas podem ser indicadas no lugar da imagem completa quando uma busca rápida por clipes de filmes é requerida.

EXCERTOS DE FILMES

20 Versões pequenas e com sub-amostras dos quadros que constituem um evento. Esses pequenos quadros podem ser armazenados como um fluxo de dados paralela ou como parte dos dados sinóticos. Eles podem ser novamente gravados no lugar dos dados completos quando uma busca rápida por clipes de filmes é requerida. Um trecho de filme não é uma coleção de 25 Thumbnails: isso seria caro demais para se armazenar.

DETECÇÃO DE EVENTO DE VÍDEO

30 Um evento de vídeo é uma coleção de quadros de vídeo consecutivos de uma ou mais fontes de dados de vídeo. Pelo menos um dos quadros que constituem essa coleção, o quadro chave, é especial de alguma maneira e define o evento. A coleção de quadros consecutivos é uma coleção que transpõe todos os quadros que contêm os quadros chave: haverá um critério para o quão grande uma abertura entre os quadros

chave delineia eventos diferentes. A coleção pode ainda incluir uma série de quadros que precedem o primeiro quadro chave e que seguem o último quadro chave: esta é a essência da gravação de pré e pós-evento. Isto contrasta com a

5 Detecção de Movimento de Vídeo, que se refere à detecção do movimento em alguma região de um quadro único de vídeo. O quadro de vídeo onde o movimento foi detectado pela Detecção de Movimento de Vídeo é freqüentemente um quadro chave que define um evento de vídeo.

10 Vide também Evento, VMD, Detecção de Movimento de Vídeo

QUADRO DE VÍDEO

Um quadro tal como utilizado na presente invenção é definido como a menor unidade temporal de uma seqüência de 15 vídeo a ser representada como uma única imagem.

SEQÜÊNCIA DE VÍDEO

Uma seqüência de vídeo tal como utilizada na presente invenção é definida como uma seqüência temporalmente requisitada das imagens digitais individuais que podem ser 20 geradas diretamente a partir de uma fonte digital, tal como uma câmera ou aplicação de arte gráfica em um computador, ou pode ser produzida pela conversão digital (digitalização) da porção visual dos sinais análogos, tais como aquelas produzidas pela transmissão da televisão ou mídia de gravação 25 ou pode ser produzida pela conversão digital (digitalização) da película do filme.

DETECÇÃO DE MOVIMENTO DE VÍDEO

Detecção de Movimento de Vídeo: um dos objetivos primários consiste em encontrar as mudanças na cena que não 30 se dão simplesmente devido às variações nas condições ambientais. Os movimentos são de diversos tipos. São distinguidas as mudanças gerais (tais como as árvores que se movem ao vento) das mudanças devido a intrusões (tais como

veículos). O movimento anterior é reconhecido pelo fato de que tal movimento está limitado dentro da cena e é manifestamente refluxo.

VMD

5 Vide Detecção de Movimento de Vídeo

COEFICIENTES DE ONDELETAS

A representação de uma imagem por meio da transformação de ondeletas produz uma disposição de números que podem ser utilizados para reconstruir precisamente a imagem. A transformação é efetuada ao processar grupos de pixels da imagem com um conjunto de números referidos como coeficientes de ondeletas. Há muitos tipos de ondeletas, sendo que cada um deles é representado por seu próprio conjunto de coeficientes particular. Do ponto de vista da compressão da imagem, esses conjuntos do coeficiente que permitem a compressão máxima são vantajosos. No entanto, os dados produzidos por esses coeficientes serão censurados e aproximados a fim de ganhar um nível de compressão maior. Desse modo, conjuntos de coeficientes que fornecem uma reconstrução resistente e exata face à esta censura e aproximação, também devem ser preferidos. Muitos debates focalizam quais conjuntos de coeficientes de ondeletas particulares fazem um melhor trabalho com relação a estas duas abordagens.

25 COMPRESSÃO DE ONDELETAS

Dois fatores tornam possível conseguir a compressão significativa de dados de ondeletas. A estrutura hierárquica da representação de ondeletas da imagem predispõe em direção a isto, sendo grande o número de coeficientes avaliados como quase zero que são hierarquicamente relacionados. O processo de thresholding coeficiente intensifica o número de valores zero nessa hierarquia e o processo de quantização assegura que os valores diferentes de zero sejam representados

eficientemente. É, portanto, possível representar os dados de uma maneira bem mais eficiente, consumindo bem menos espaço de armazenamento.

CRIPTOGRAFIA DE ONDELETAS

5 Quando os coeficientes de ondeletas foram quantizados, há relativamente poucos valores representados pelos códigos que são armazenados em uma tabela de consulta (vide a Quantização de ondeletas). O número de código pode ser consultado para a reconstrução. No entanto, antes do 10 armazenamento, é possível criptografar a tabela que fornece os valores de código, em consequência do que os programas sem acesso ao método de criptografia não poderão reconstruir a imagem.

KERNEL DE ONDELETAS

15 A transformação de ondeletas de uma imagem consiste em uma hierarquia das imagens com um tamanho sempre reduzido. O fator de escala entre os níveis de hierarquia é geralmente, mas não necessariamente, um fator linear de 2: um bloco de 2x2 de quatro pixels transforma-se em um pixel. É feita referência ao menor nível que é utilizado como o "Kernel de 20 ondeletas" uma vez que todas as imagens mais amplas (maiores) são construídas a partir desta através de uma transformação de ondeletas.

QUANTIZAÇÃO DE ONDELETAS

25 A transformação de ondeletas dos dados consiste em um conjunto de números que podem ser utilizados para reconstruir os dados originais. A fim de atingir níveis de compressão substanciais, é útil simplificar esses números, representando os valores reais por alguns valores 30 representativos. A maneira com a qual os valores representantes são selecionados tem que ser tal que o resultado não faça uma mudança perceptível aos dados reconstruídos. Esse processo é denominado quantização, uma

vez que muda o que é essencialmente um conjunto de valores contínuo (os coeficientes de ondeletas originais) em um conjunto de valores distintos apropriado. Poucos valores distintos podem ser codificados, substituindo cada valor por 5 um código específico que pode ser consultado durante o processo de reconstrução. Desse modo, o valor 29.6135 pode ser representado pela letra 'W' e cada 'W' é substituído por 29.6135 na reconstrução. A codificação abre a possibilidade de encriptação dos dados.

10 THRESHOLDING DE ONDELETAS

A transformação de ondeletas dos dados consiste em um conjunto de números que podem ser utilizados para reconstruir os dados originais. A fim de atingir níveis de compressão substanciais, é útil descartar aqueles números que 15 são tão pequenos que a sua perda não fará uma mudança perceptível aos dados reconstruídos. O thresholding é de sentido único sendo que uma decisão é feita com relação a se um número pode ser rejeitado com segurança ou não. Há muitas maneiras de decidir quais poderiam ser os valores mais 20 favoráveis do limite e o que fazer com os dados uma vez que o thresholding tenha sido executado. Tal método é denominado como "ERIS" (para "Estimador de Risco Imparcial de Stein").

TRANSFORMAÇÃO DE ONDELETAS

Uma transformação de dados seqüenciais ou de imagem 25 em que os dados transformados têm a metade do comprimento da escala linear dos dados originais. O conjunto de dados reduzido é mantido com um outro conjunto de dados que contém as informações necessárias para reconstrução dos dados originais a partir da versão reduzida. A possibilidade de 30 reconstruir os dados originais a partir dos dados encolhidos é uma característica chave das ondeletas.

XviD

XviD é um codec de vídeo de MPEG-4 de fonte aberta

e livre. XviD foi criado por um grupo de programadores voluntários depois que a fonte aberta DivX foi fechada em julho de 2001. Nos releases 1.0.x, uma licença de GNU GPL v2 é utilizada sem nenhuma restrição geográfica explícita; no 5 entanto, a utilização legal de XviD pode ainda ser restrita por leis locais. Observe que os arquivos codificados XviD podem ser gravados em um CD ou DVD e podem ser executados em um tocador de DVD compatível com DivX.

REFERÊNCIA REMISSIVA A PEDIDO DE PATENTE CORRELATO

10 O presente pedido de patente reivindica o benefício do Pedido de Patente Provisório U.S. N°. 60/712.810, depositado em 1º de setembro de 2005, cuja totalidade é aqui incorporada a título de referência neste pedido.

P.03.2023

REIVINDICAÇÕES

1. MÉTODO PARA INTERROGAR OU BUSCAR UM CORPO DE DADOS DIGITADOS SEQÜENCIAIS, caracterizado pelo fato de utilizar as seguintes etapas:

5 (a) decompor os dados utilizando uma decomposição piramidal;

(b) aplicar um processo de classificação para separar as informações sobre atributos dos dados (dados sinóticos);

10 (c) gerar uma máscara para cada um dos atributos dos dados, delineando onde nos dados o atributo está localizado;

(d) gerar dados sinóticos, que tomam a forma de uma ou mais máscaras;

15 (e) armazenar os dados e os dados sinóticos com um índice;

(f) configurar os critérios de interrogação ou de busca;

(g) recuperar os dados sinóticos;

20 (h) aplicar os critérios de interrogação ou busca aos dados sinóticos recuperados.

2. MÉTODO, de acordo com a reivindicação 1, caracterizado pelo fato de que o índice é utilizado para recuperar os dados principais correspondentes.

25 3. MÉTODO, de acordo com a reivindicação 1, caracterizado pelo fato de que a decomposição é feita utilizando wavelets.

30 4. MÉTODO, de acordo com a reivindicação 2, caracterizado pelo fato de que a decomposição é feita utilizando uma hierarquia adaptável de wavelets.

5. MÉTODO, de acordo com a reivindicação 1, caracterizado pelo fato de que o processo de classificação é utilizado para extrair atributos de ruído.

6. MÉTODO, de acordo com a reivindicação 1, caracterizado pelo fato de que o processo de classificação é utilizado para extrair as informações sobre um fundo estático.

5 7. MÉTODO, de acordo com a reivindicação 1, caracterizado pelo fato de que o processo de classificação é utilizado para extrair as informações sobre um fundo estacionário.

10 8. MÉTODO, de acordo com a reivindicação 1, caracterizado pelo fato de que o processo de classificação é utilizado para extrair as informações sobre movimentos dinâmicos.

15 9. MÉTODO, de acordo com a reivindicação 1, caracterizado pelo fato de que o processo de classificação é utilizado para extrair as informações sobre objetos.

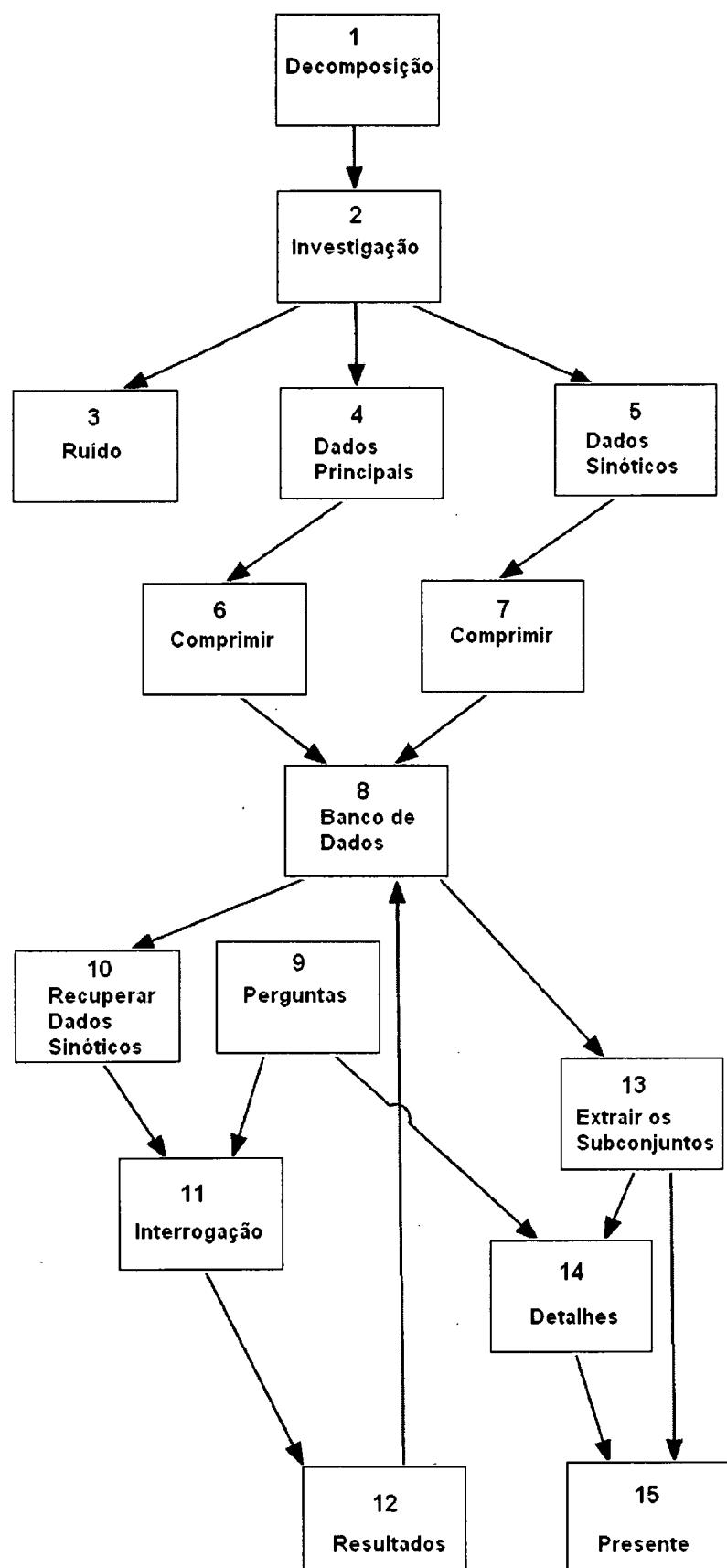


FIG. 1

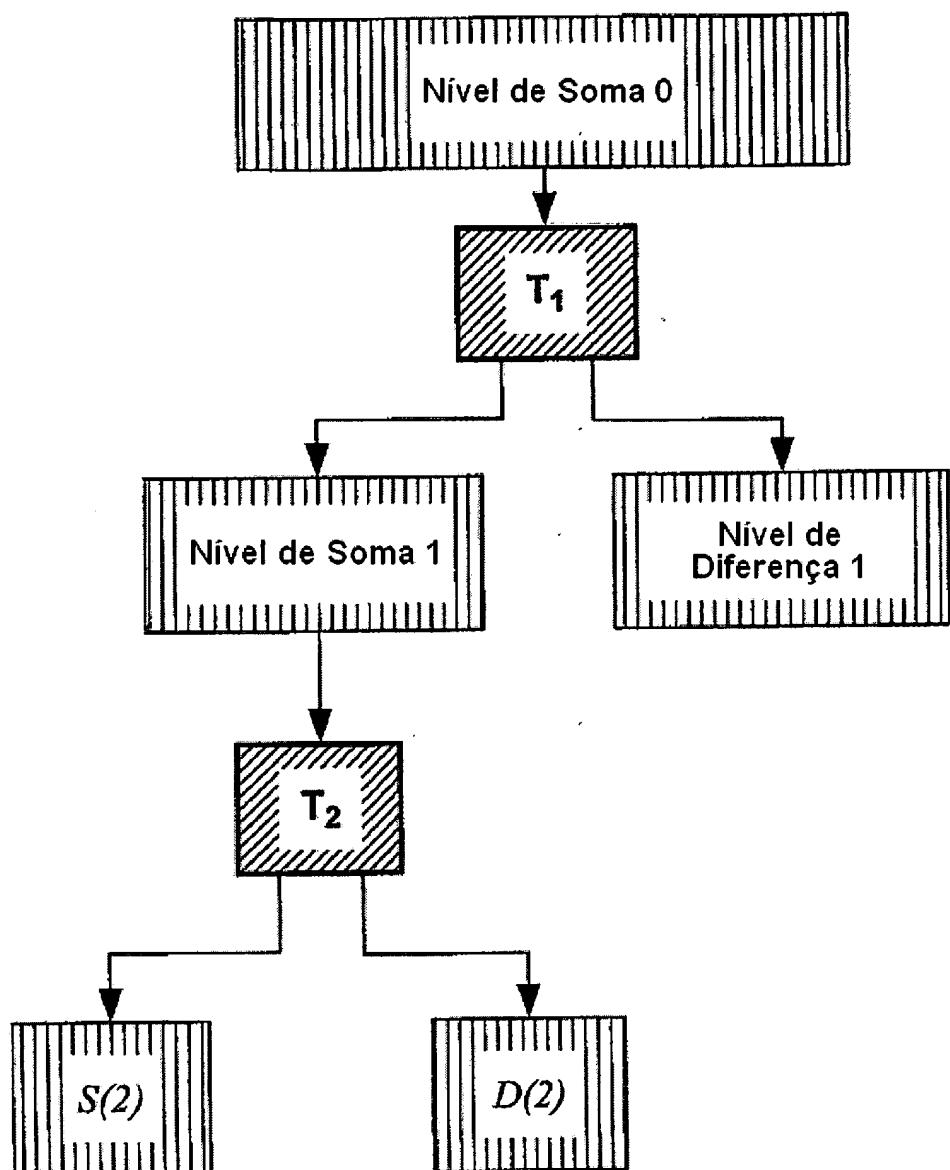


FIG. 2

3/19

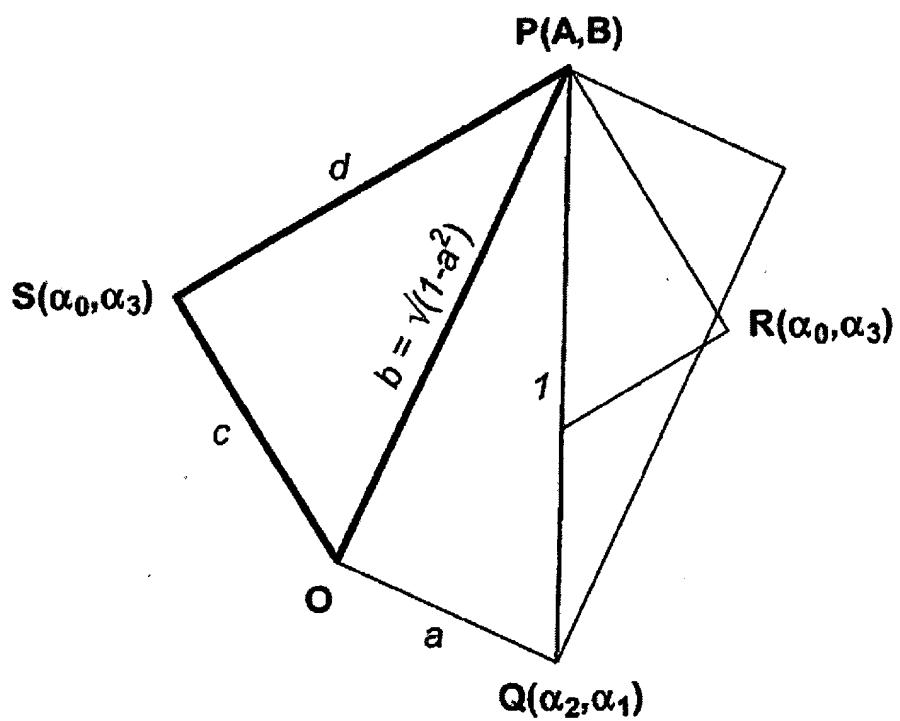
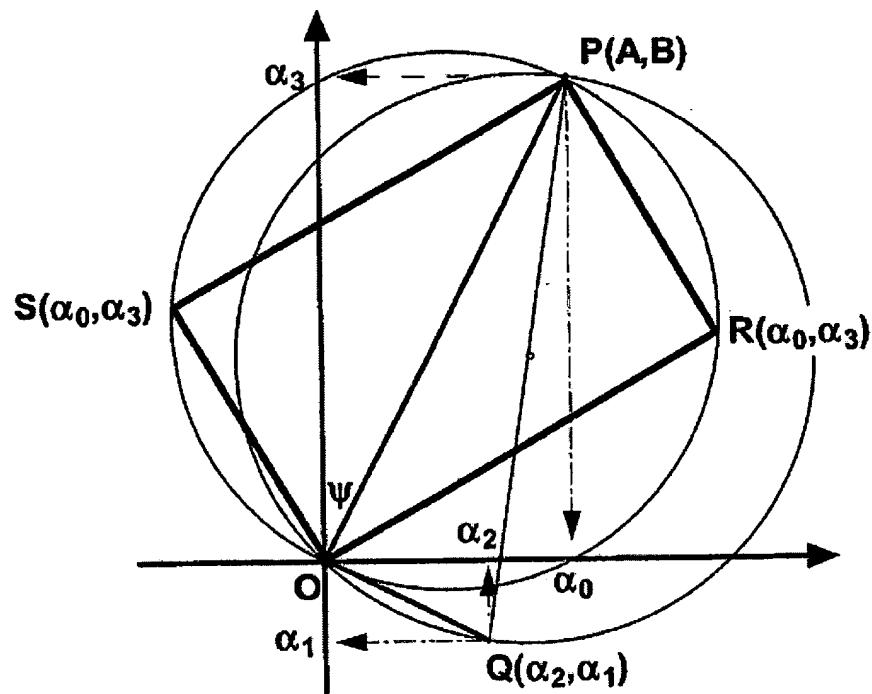


FIG. 3

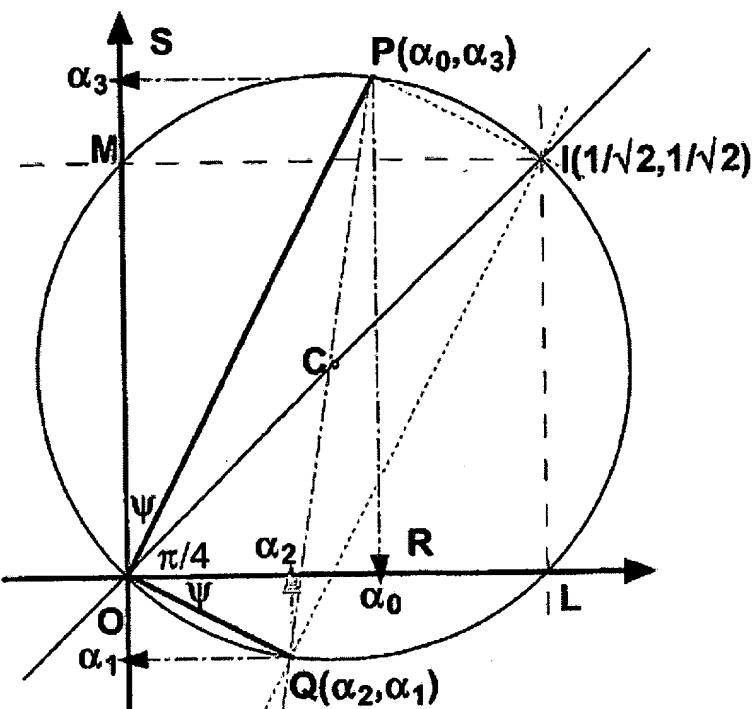


FIG. 4

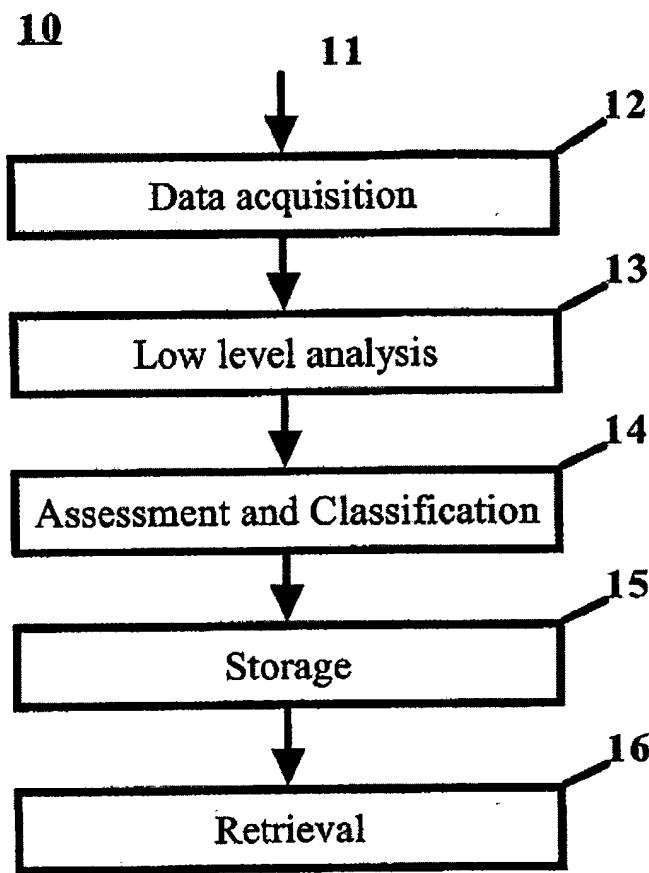


FIG. 5

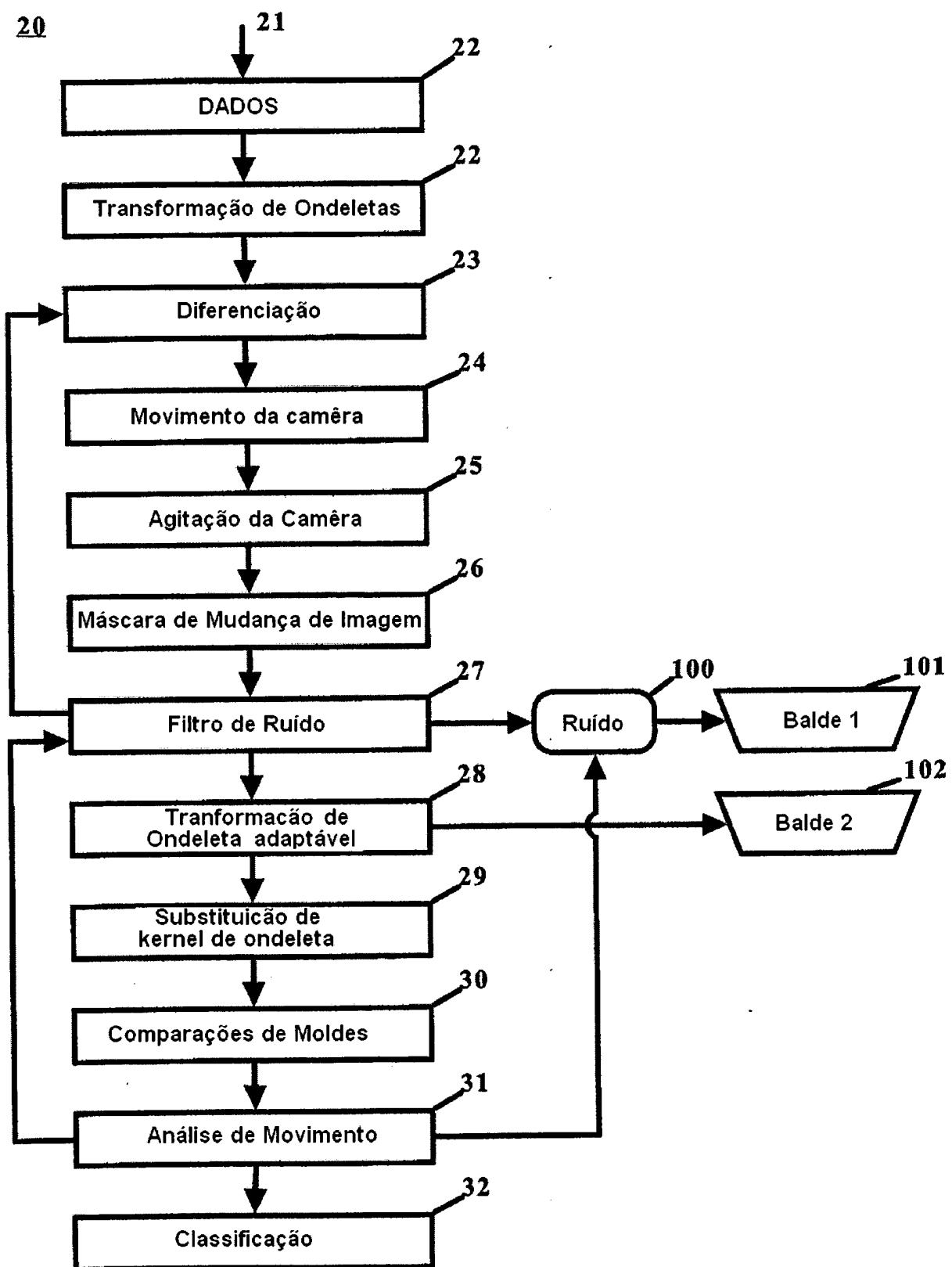


FIG. 6

6/19

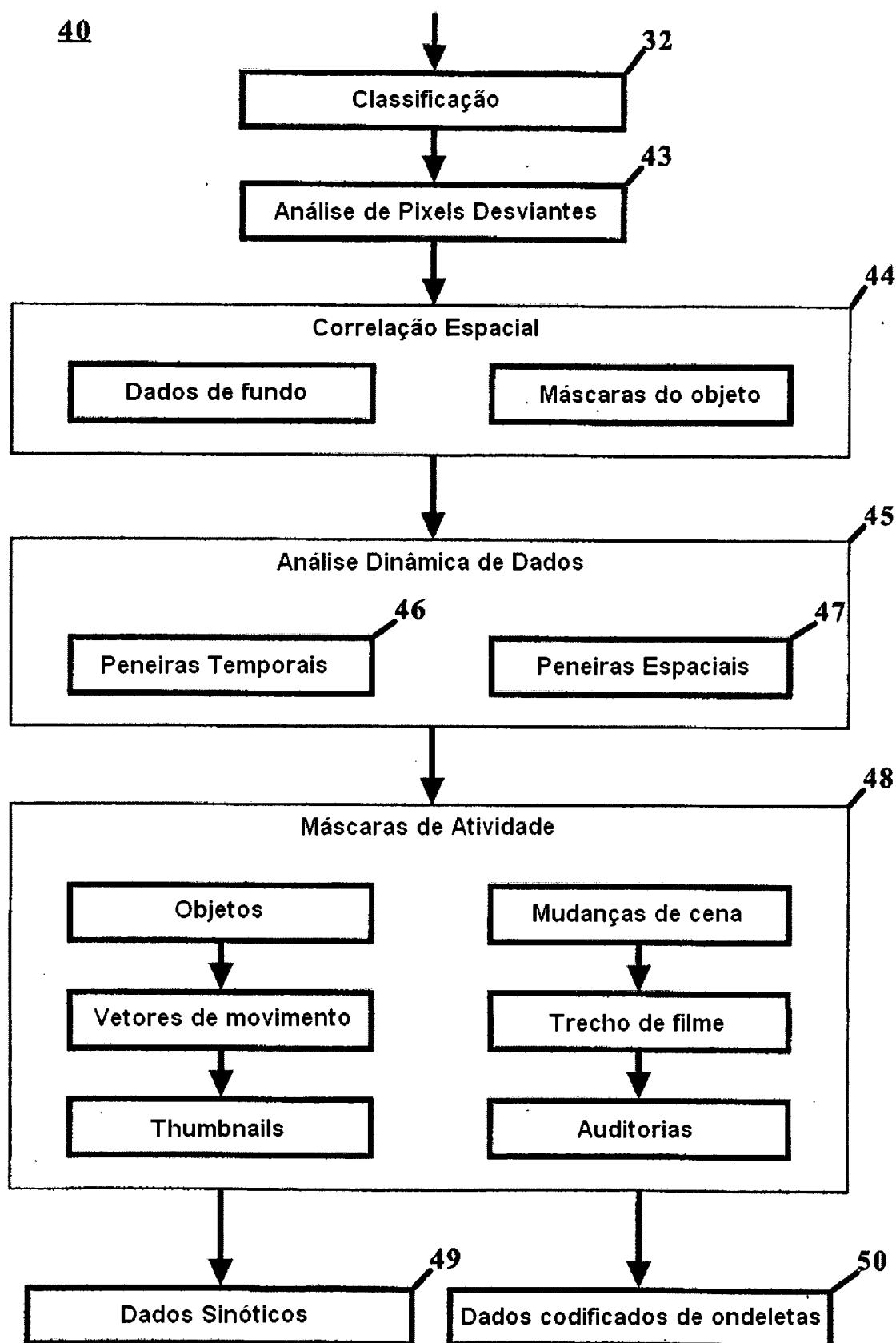


FIG. 7

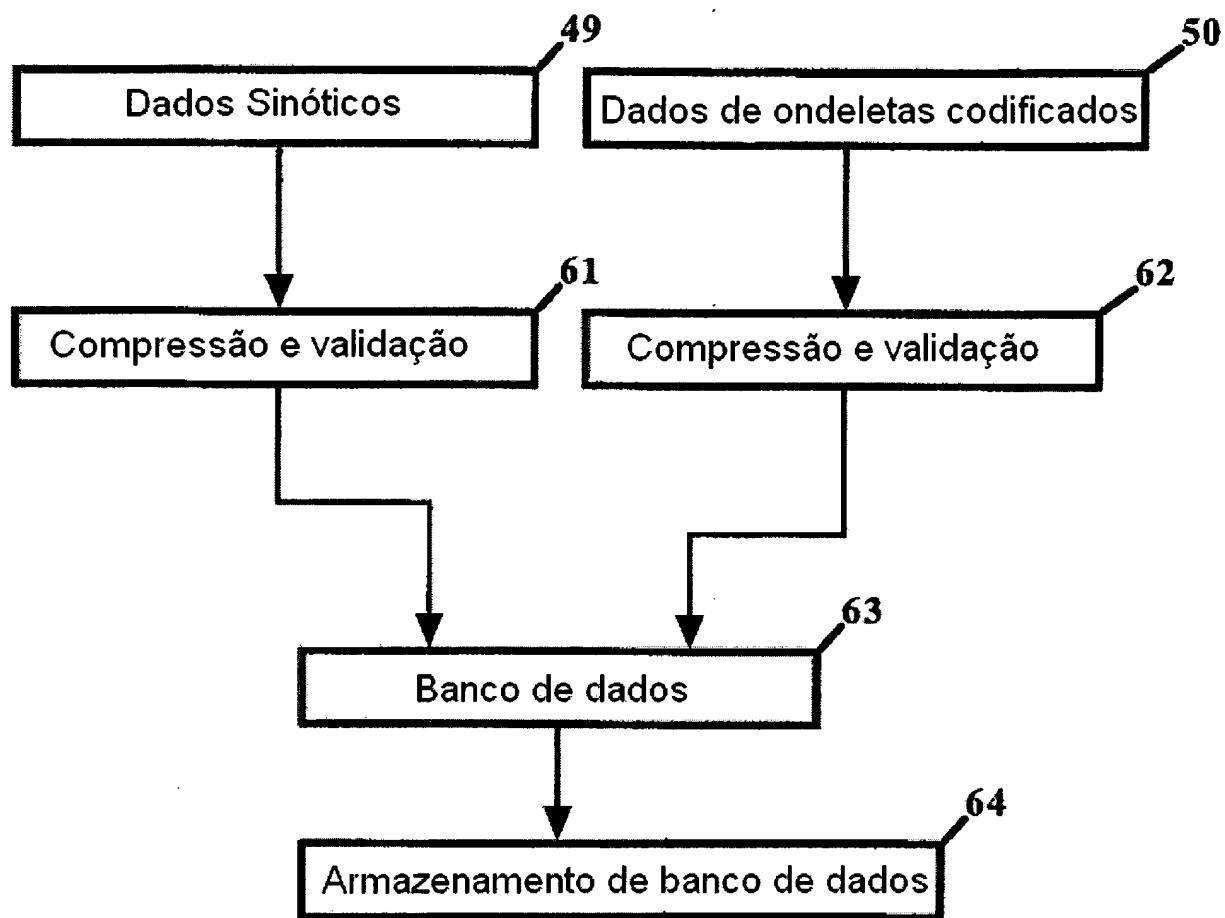
60

FIG. 8

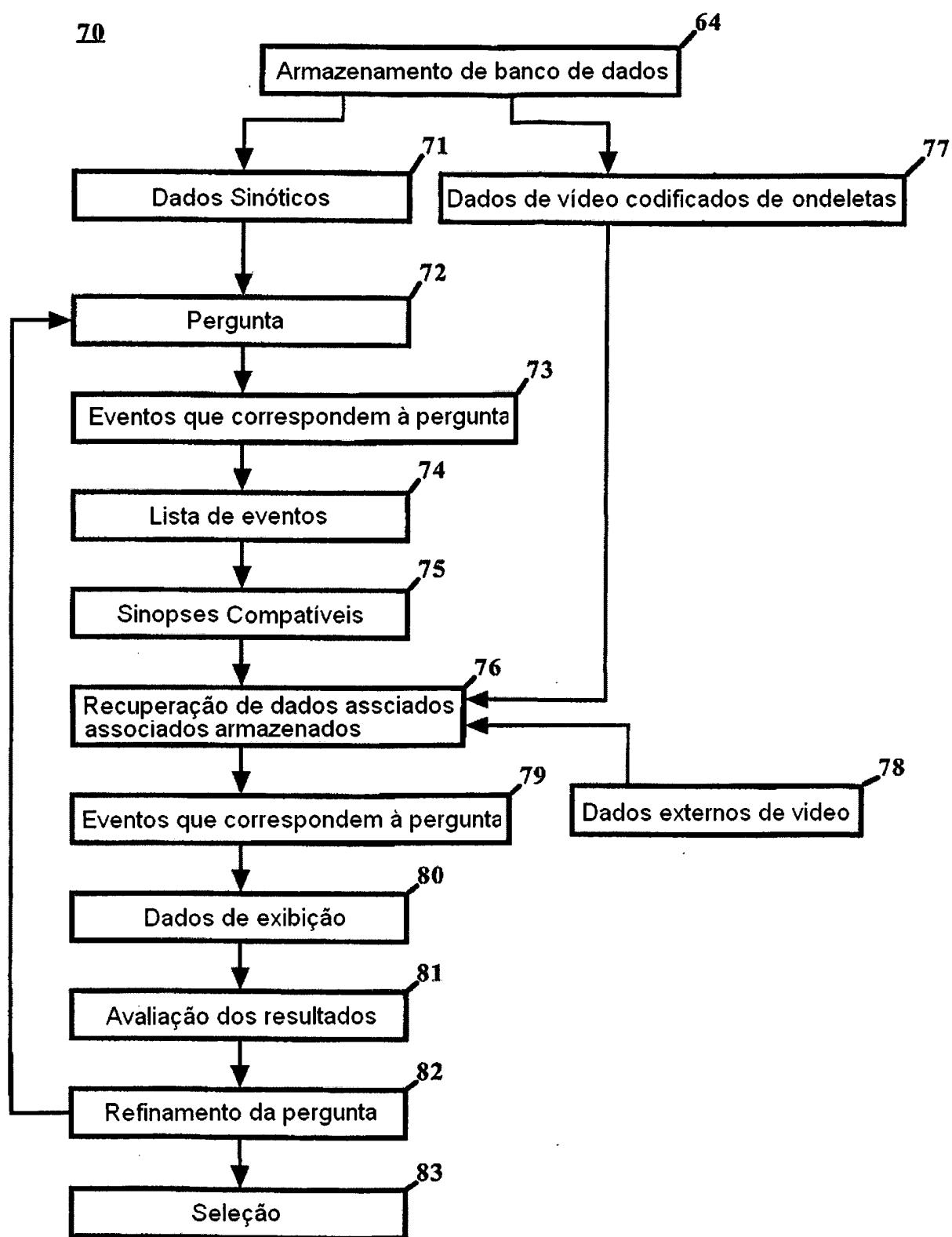


FIG. 9

90

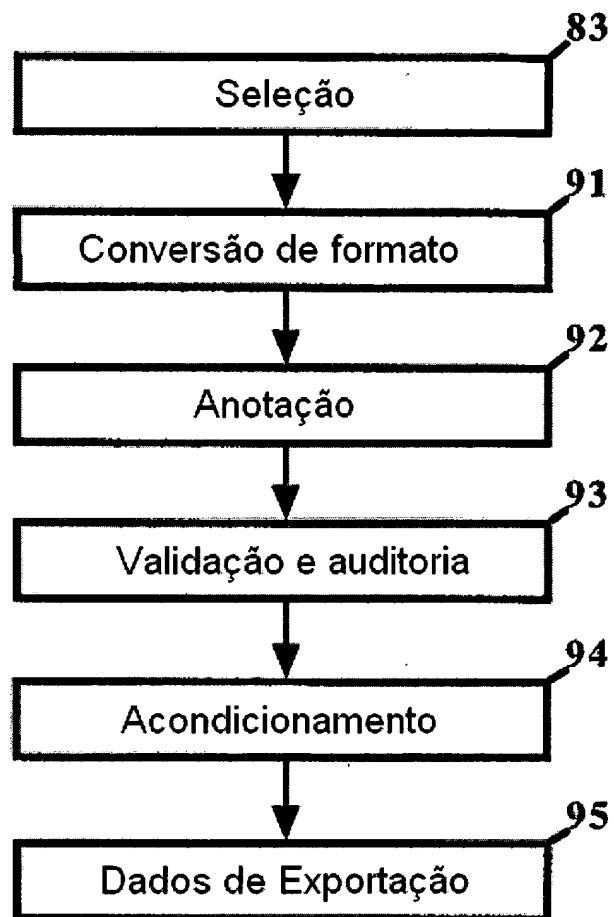


FIG. 10

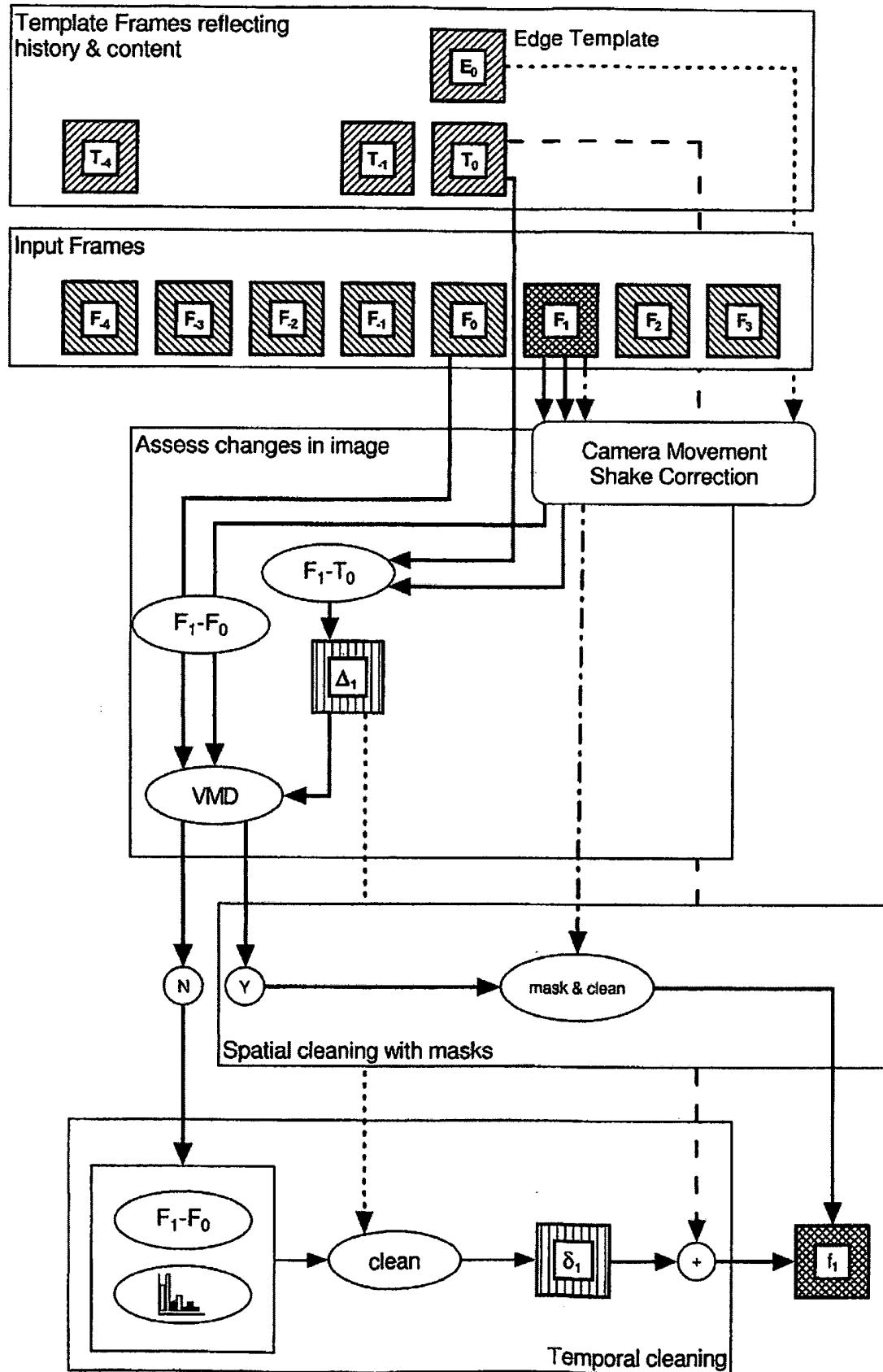


FIG. 11

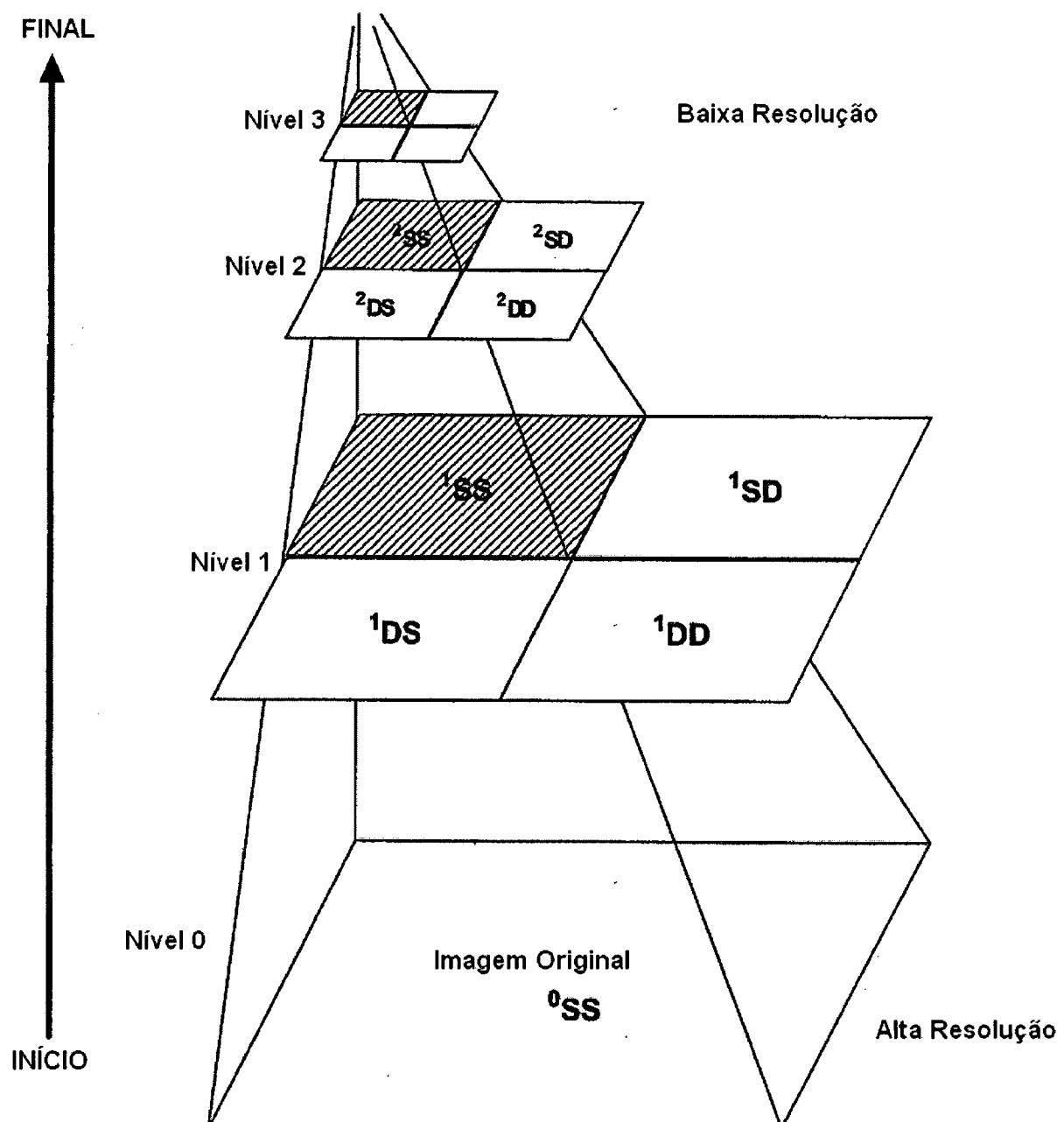
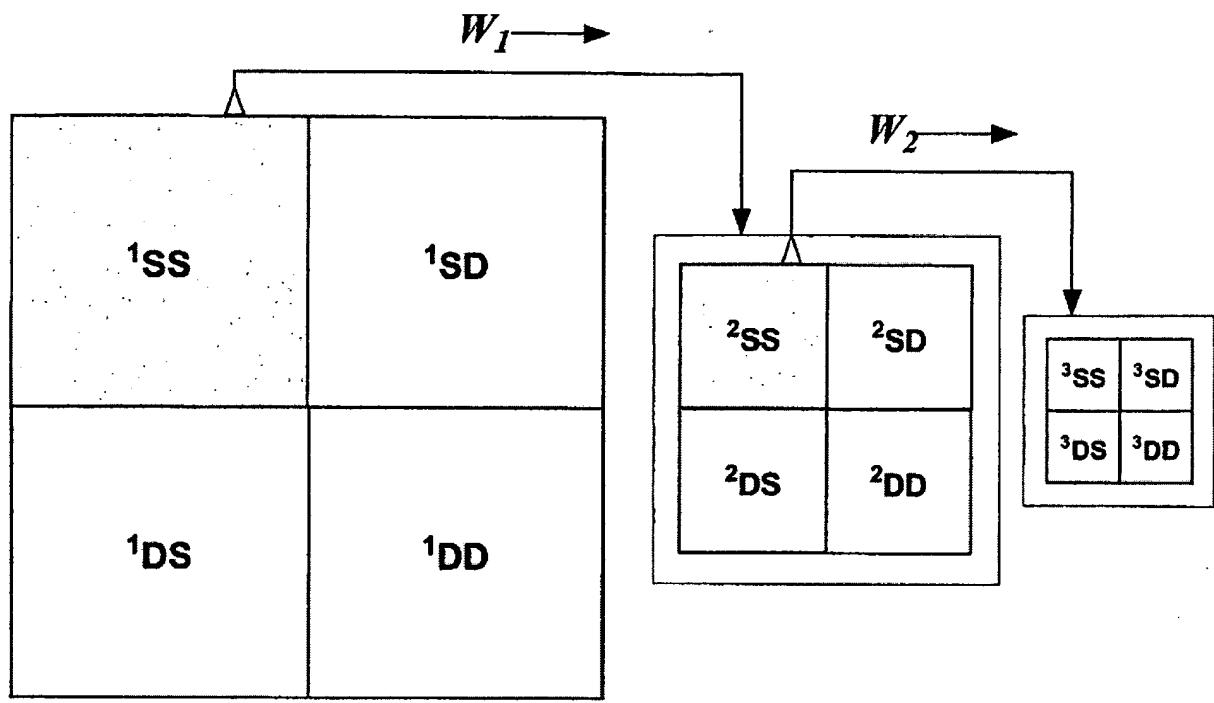
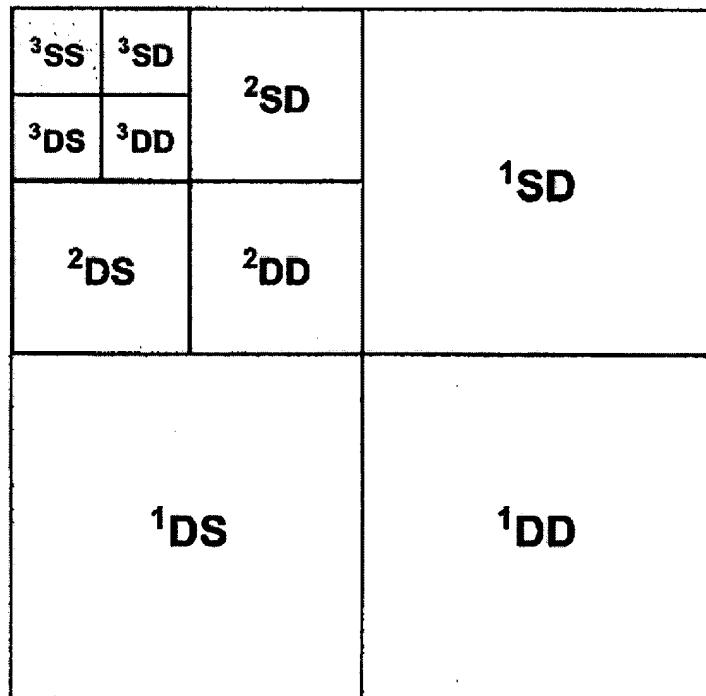


FIG. 12

12/19



Criando a hierarquia



Transformação de ondeletas para 3 níveis

FIG. 13

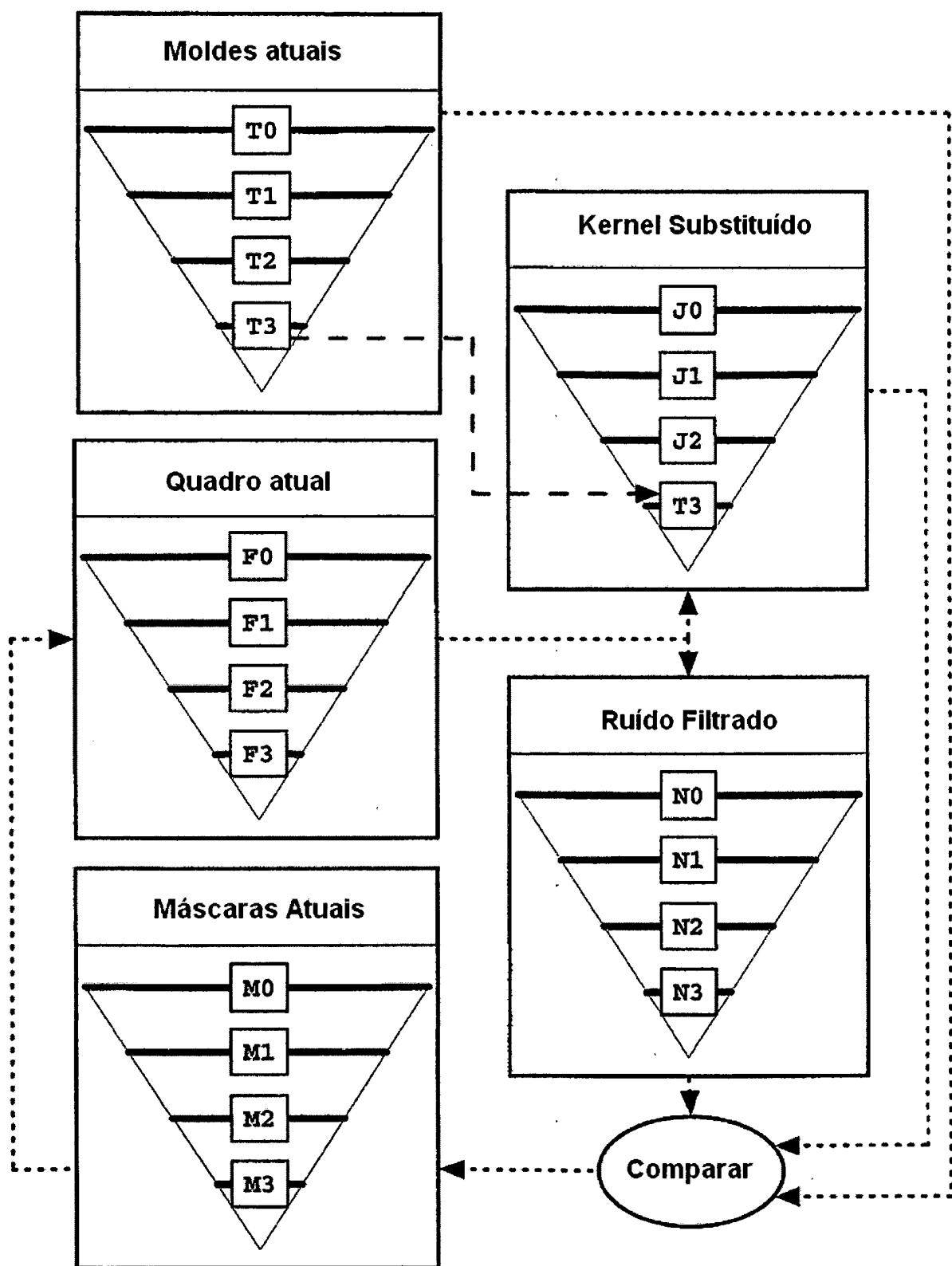
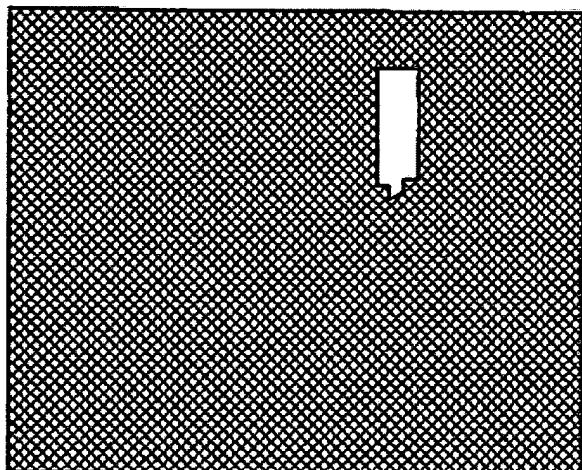


FIG. 14

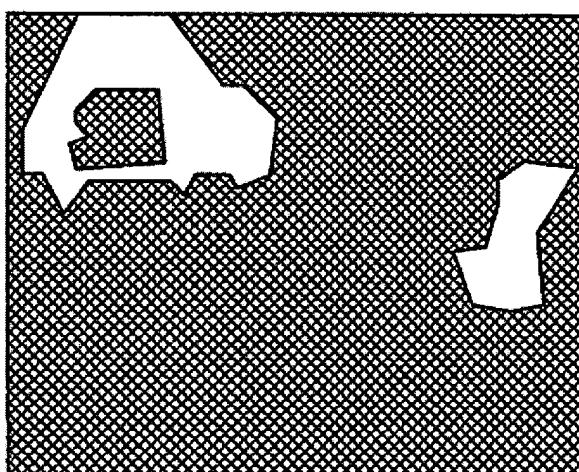
Imagen Original
Máscaras Superpostas



Máscara VMD



Fundo Móvel



Máscara de Iluminância

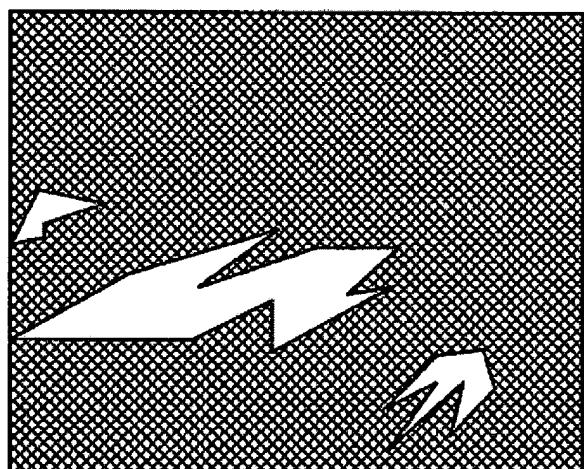
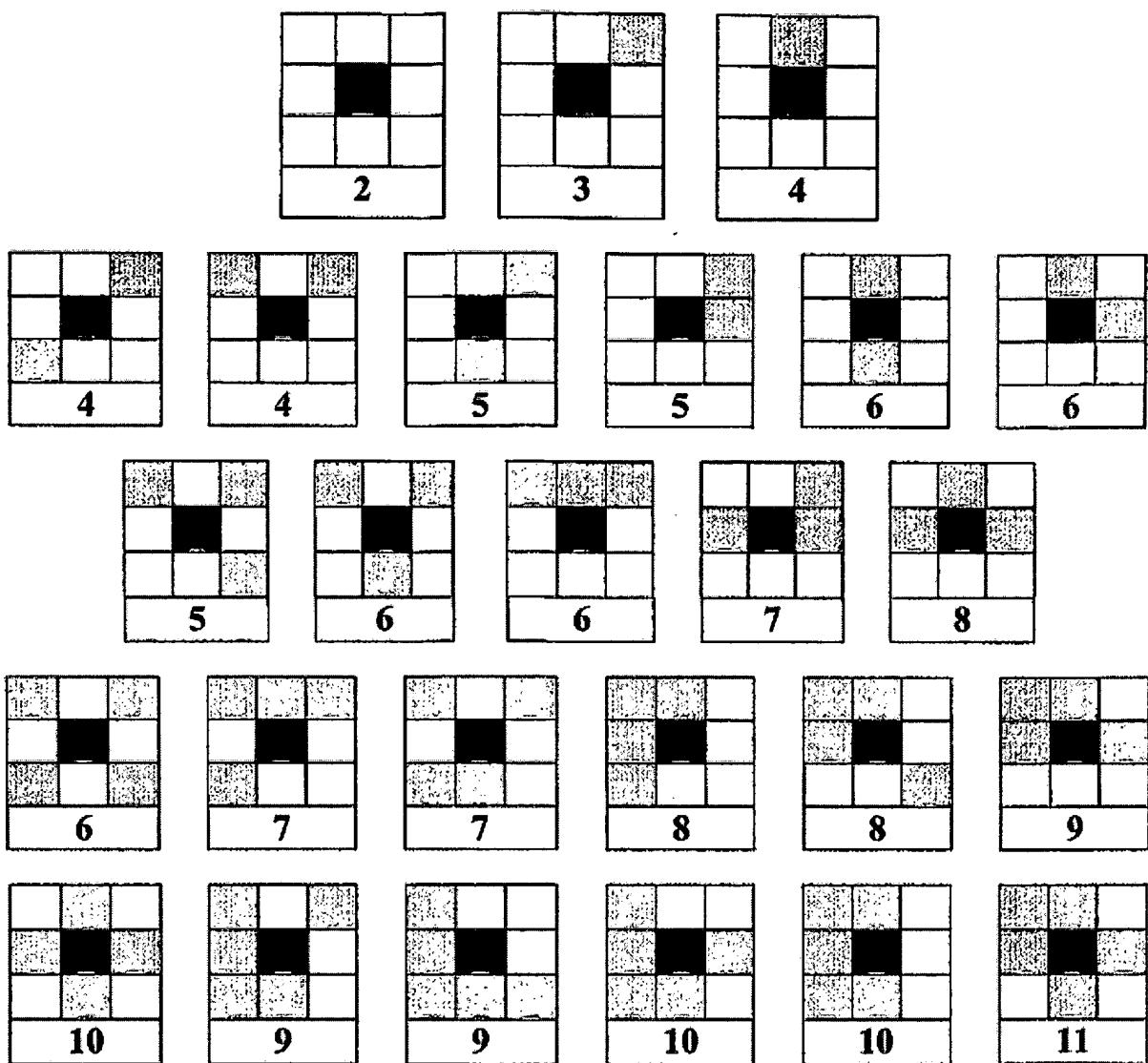


FIG. 15

15/19



Pontuação de Pixels

Pontuação de Padrão Especial

FIG. 16

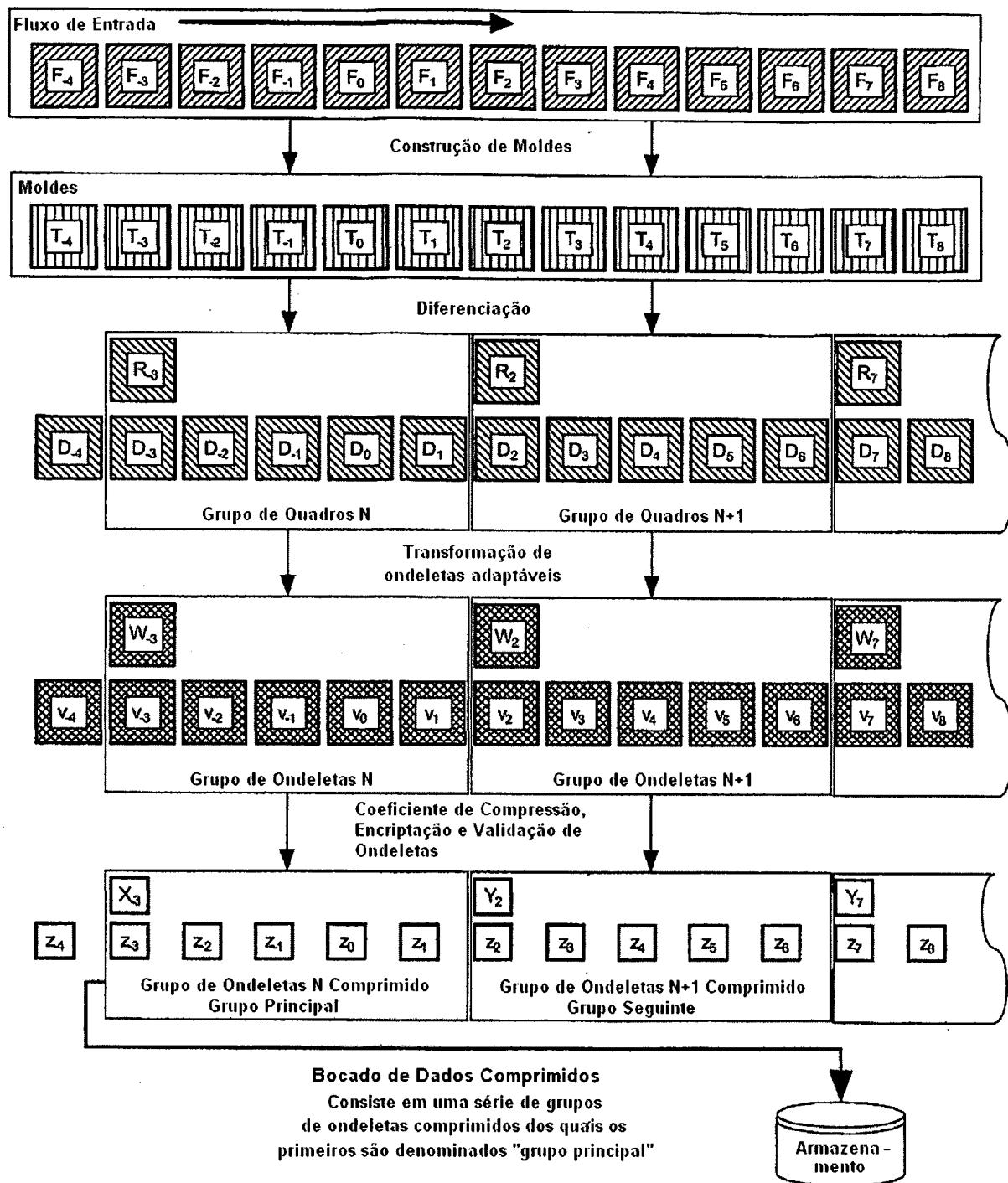


FIG. 17

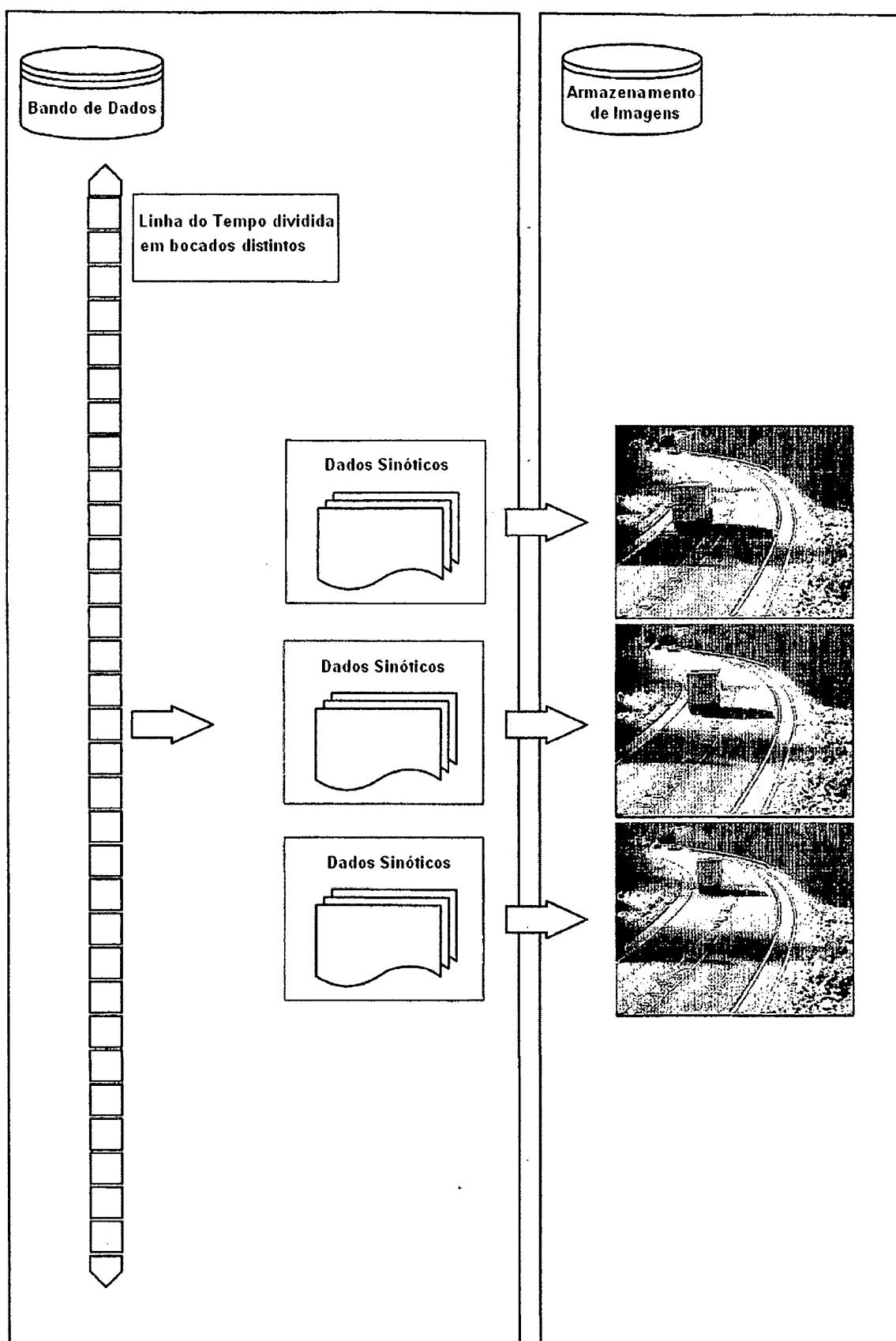


FIG. 18

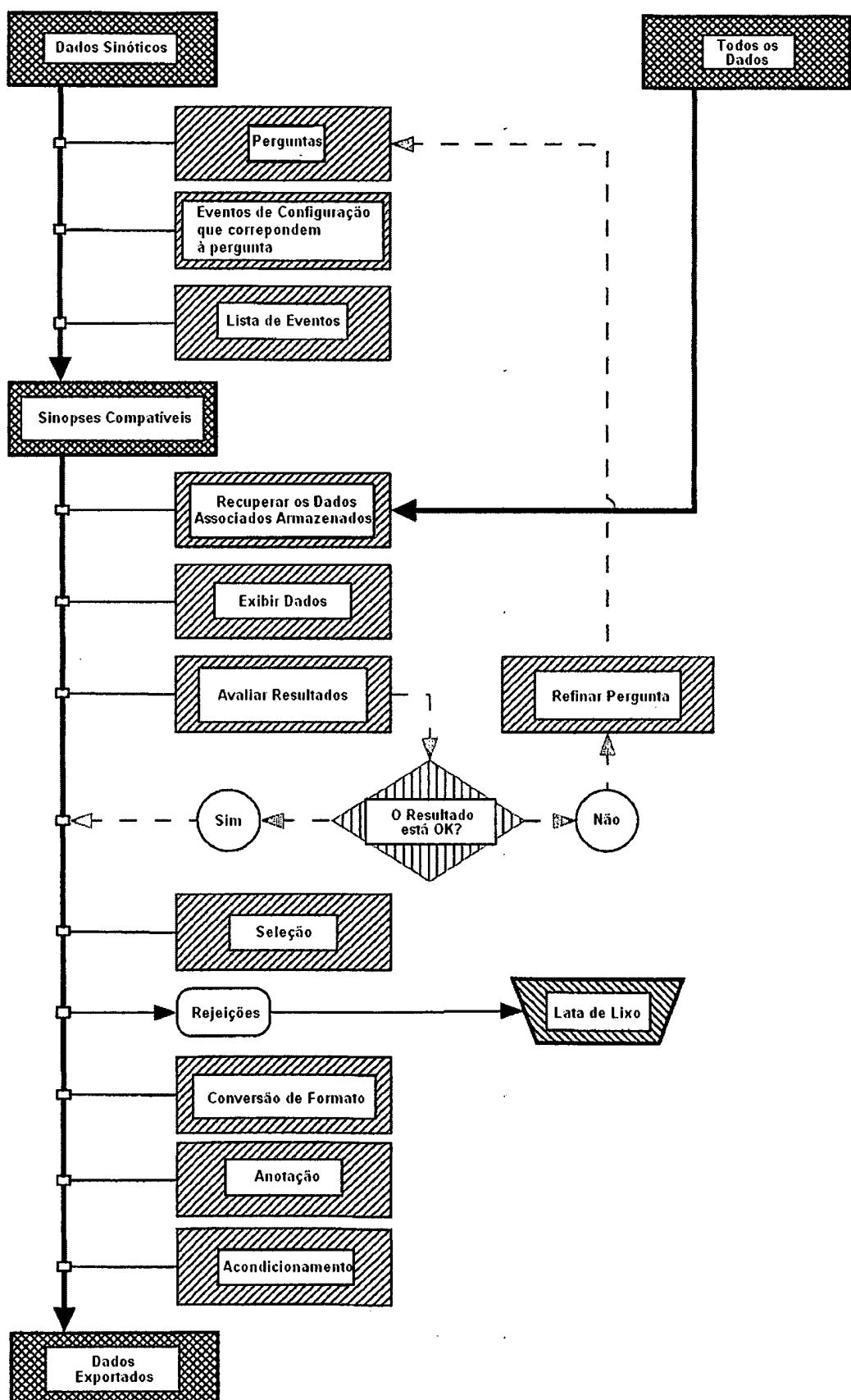


FIG. 19

19/19

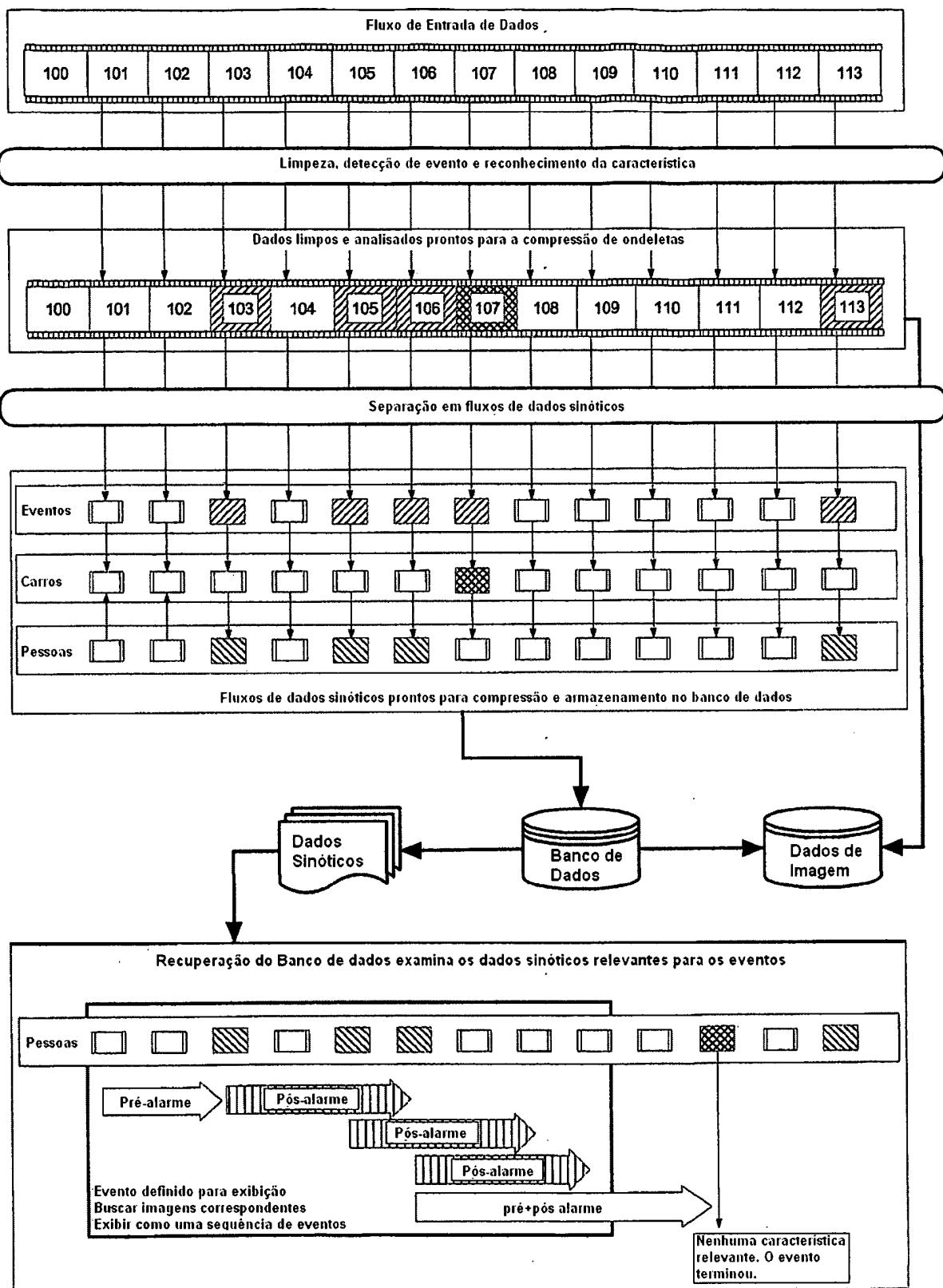


FIG. 20

RESUMO

MÉTODO PARA INTERROGAR OU BUSCAR UM CORPO DE DADOS DIGITADOS SEQÜENCIAIS, MÉTODO PARA AJUDAR NA COMPUTAÇÃO DE WAVELETS PARA APLICAÇÕES E MÉTODO PARA PROCESSAR UMA SEQÜÊNCIA DE DADOS DIGITADOS

Quando são feitas gravações de dados digitais utilizando alguma forma de computador ou calculadora, os dados são inseridos em uma variedade de maneiras e são armazenados em alguma forma de meio eletrônico. Durante esse processo, os cálculos e as transformações são executados nos dados para otimizar os mesmos para o armazenamento. A presente invenção envolve projetar dos cálculos de tal maneira que eles incluem o que é necessário para cada um de muitos processos diferentes, tais como a compressão de dados, a detecção da atividade e o reconhecimento do objeto. Enquanto os dados de entrada são sujeitados a esses cálculos e armazenados, as informações sobre cada um dos processos são extraídas ao mesmo tempo. Os cálculos para os diferentes processos podem ser executados tanto em série em um único processador, ou em paralelo a múltiplos processadores distribuídos. O processo de extração é denominado como "decomposição sinótica", e as informações extraídas são denominadas como "dados sinóticos". O termo "dados sinóticos" não inclui normalmente o corpo principal de dados originais. Os dados sinóticos são criados sem nenhuma tendência prévia a interrogações específicas que podem ser feitas, de modo que é desnecessário inserir critérios de busca antes de fazer a gravação. Nem depende da natureza dos algoritmos/cálculos utilizados para fazer a decomposição sinótica. Os dados resultantes, que compreendem os dados originais (processados) juntamente com os dados sinóticos (processados), são então

armazenados em um banco de dados relacional. Alternativamente, os dados sinóticos de uma forma simples podem ser armazenados como parte dos dados principais. Depois que a gravação é feita, os dados sinóticos podem ser analisados sem a necessidade de examinar o corpo principal dos dados. Essa análise pode ser feita muito rapidamente porque o volume dos cálculos necessários já foi executado no momento da gravação original. A análise dos dados sinóticos fornece marcadores que podem ser utilizados para acessar os dados relevantes da gravação dos dados principais caso requerido. O efeito global de fazer uma análise desta maneira é que uma grande quantidade de dados digitais gravados, que pode levar dias ou semanas para analisar por meios convencionais, pode ser analisada em segundos ou minutos. A presente invenção também se refere a um processo para a geração de famílias parameterizadas contínuas de wavelets. Muitas das wavelets podem ser expressas exatamente dentro de representações de 8 bit ou 16 bits. A presente invenção também se refere aos processos para a utilização de wavelets adaptáveis para extrair informações que são resistentes a variações em condições ambientais, e para executar a compressão de dados utilizando a quantização localmente adaptável e esquemas de limitação, e para executar a análise da pós-gravação.