



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

(21) PI 0719555-9 A2



* B R P I 0 7 1 9 5 5 5 A 2 *

(22) Data de Depósito: 30/11/2007
(43) Data da Publicação: 10/12/2013
(RPI 2240)

(51) Int.Cl.:
G06T 7/20

(54) Título: ESTIMATIVA DE UMA LOCALIZAÇÃO DE UM OBJETO EM UMA IMAGEM (57) Resumo:

(30) Prioridade Unionista: 01/12/2006 US 60/872.145,
01/12/2006 US 60/872.146, 19/01/2007 US 60/885.780

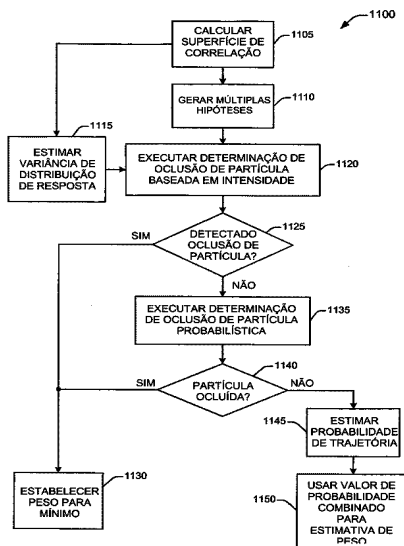
(73) Titular(es): Thomson Licensing

(72) Inventor(es): Joan Llach, Yu Huang

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

(86) Pedido Internacional: PCT US2007024713 de
30/11/2007

(87) Publicação Internacional: WO 2008/070012de
12/06/2008



“ESTIMATIVA DE UMA LOCALIZAÇÃO DE UM OBJETO EM UMA IMAGEM”
REFERÊNCIA CRUZADA A PEDIDOS RELACIONADOS

Este pedido reivindica o benefício de cada um dos três seguintes pedidos: (1) pedido provisório US 60/872.145 intitulado “Cluttered Backgrounds and Object Tracking” e depositado em 1 de dezembro de 2006 (protocolo representante PU060244), (2) pedido provisório US 60/872.146 intitulado “Modeling for Object Tracking” e depositado em 1 de dezembro de 2006 (protocolo representante PU060245), e (3) pedido provisório US 60/885.780 intitulado “Object Tracking” e depositado em 19 de janeiro de 2007 (protocolo representante PU070030). Todos estes três pedidos de prioridade estão incorporados em suas totalidades por meio desta referência para todos os propósitos.

CAMPO DA INVENÇÃO

Pelo menos uma implementação nesta revelação diz respeito a estimativa de estado dinâmico.

ANTECEDENTES DA INVENÇÃO

Um sistema dinâmico se refere a um sistema no qual um estado do sistema muda ao longo do tempo. O estado pode ser um conjunto de variáveis escolhidas arbitrariamente que caracterizam o sistema, mas o estado frequentemente inclui variáveis de interesse. Por exemplo, um sistema dinâmico pode ser construído para caracterizar um vídeo, e o estado pode ser escolhido para ser uma posição de um objeto em um quadro do vídeo. Por exemplo, o vídeo pode representar uma partida de tênis e o estado pode ser escolhido para ser a posição da bola. O sistema é dinâmico porque a posição da bola muda ao longo do tempo. Estimar o estado do sistema, isto é, a posição da bola, em um novo quadro do vídeo é de interesse.

SUMÁRIO

De acordo com um aspecto geral, uma trajetória é determinada. A trajetória é para um objeto em uma imagem particular em uma sequência de imagens digitais, e a trajetória é baseada em ou mais localizações prévias do objeto em uma ou mais imagens prévias na sequência. Um peso é determinado, para uma partícula em uma estrutura de trabalho baseada em partícula para rastrear o objeto, com base na distância da trajetória para a partícula. Uma localização estimada para o objeto é determinada usando a estrutura de trabalho baseada em partícula, a localização estimada sendo baseada no peso da partícula determinado.

Os detalhes de uma ou mais implementações estão expostos nos desenhos anexos e na descrição a seguir. Mesmo se descritas em uma maneira particular, deve ficar claro que implementações podem ser configuradas ou incorporadas de várias maneiras. Por exemplo, uma implementação pode ser executada como um método, ou incorporada como um aparelho configurado para executar um conjunto de operações, ou incorporada como um

aparelho armazenando instruções para executar um conjunto de operações, ou incorporada em um sinal. Outros aspectos e recursos se tornarão aparentes a partir da descrição detalhada a seguir considerada em conjunto com os desenhos anexos e as reivindicações.

DESCRIÇÃO RESUMIDA DOS DESENHOS

5 A figura 1 inclui um diagrama de blocos de uma implementação de um estimador de estado.

A figura 2 inclui um diagrama de blocos de uma implementação de um aparelho para implementar o estimador de estado da figura 1.

10 A figura 3 inclui um diagrama de blocos de uma implementação de um sistema para codificar dados com base em um estado estimado pelo estimador de estado da figura 1.

A figura 4 inclui um diagrama de blocos de uma implementação de um sistema para processar dados com base em um estado estimado pelo estimador de estado da figura 1.

A figura 5 inclui um diagrama que representa de forma pictorial várias funções executadas por uma implementação do estimador de estado da figura 1.

15 A figura 6 inclui um fluxograma de uma implementação de um método para determinar uma localização de um objeto em uma imagem em uma sequência de imagens digitais.

A figura 7 inclui um fluxograma de uma implementação de um processo para implementar um filtro de partículas.

20 A figura 8 inclui um fluxograma de um processo alternativo para implementar um filtro de partículas.

A figura 9 inclui um fluxograma de uma implementação de um processo para implementar um modelo dinâmico no processo da figura 8.

25 A figura 10 inclui um fluxograma de uma implementação de um processo para implementar um modelo dinâmico incluindo avaliar uma estimativa de movimento em um filtro de partículas.

A figura 11 inclui um fluxograma de uma implementação de um processo para implementar um modelo de medição em um filtro de partículas.

30 A figura 12 inclui um diagrama que representa de forma pictorial um exemplo de uma trajetória projetada com localizações de objeto ocluídas.

A figura 13 inclui um fluxograma de uma implementação de um processo para determinar se é para atualizar um modelo depois de estimar um estado usando um filtro de partículas.

35 A figura 14 inclui um fluxograma de uma implementação de um processo para determinar se é para atualizar um modelo e refinar posição de objeto depois de estimar um estado usando um filtro de partículas.

A figura 15 inclui um diagrama que representa de forma pictorial uma implementa-

ção de um método de refinar posição estimada de um objeto em relação a uma trajetória projetada.

A figura 16 inclui um fluxograma de uma implementação de um processo para estimar localização de um objeto.

5 A figura 17 inclui um fluxograma de uma implementação de um processo para selecionar estimativas de localização.

A figura 18 inclui um fluxograma de uma implementação de um processo para determinar uma posição de uma partícula em um filtro de partículas.

10 A figura 19 inclui um fluxograma de uma implementação de um processo para determinar se é para atualizar um modelo.

A figura 20 inclui um fluxograma de uma implementação de um processo para detectar oclusão de uma partícula em um filtro de partículas.

A figura 21 inclui um fluxograma de uma implementação de um processo para estimar um estado com base em partículas produzidas por um filtro de partículas.

15 A figura 22 inclui um fluxograma de uma implementação de um processo para mudar uma posição estimada de um objeto.

A figura 23 inclui um fluxograma de uma implementação de um processo para determinar uma localização de objeto.

DESCRIÇÃO DETALHADA

20 Uma ou mais modalidades fornecem um método de estimativa de estado dinâmico. Uma ou mais modalidades fornecem um método de estimar estados dinâmicos. Um exemplo de uma aplicação na qual estimativa de estado dinâmico é usada é prever o movimento de um recurso no vídeo entre quadros. Um exemplo de vídeo é vídeo compactado, o qual pode ser compactado, a título de exemplo, no formato MPEG-2. Em vídeo compactado, somente um subconjunto dos quadros tipicamente contém informação completa tal como para a imagem associada com os quadros. Tais quadros contendo informação completa são chamados de quadros I no formato MPEG-2. A maioria dos quadros somente fornece informação indicando diferenças entre o quadro e um ou mais quadros próximos, tais como os quadros I próximos. No formato MPEG-2, tais quadros são denominados quadros P e quadros B. É um desafio incluir informação suficiente para prever o progresso de um recurso em vídeo enquanto mantendo ainda compactação de dados.

25

30

Um exemplo de um recurso em vídeo é uma bola em um evento esportivo. Exemplos incluem bolas de tênis, bolas de futebol e de basquetebol. Um exemplo de uma aplicação na qual o método é usado é prever a localização de uma bola entre quadros em um vídeo de múltiplos quadros. Uma bola pode ser um objeto relativamente pequeno, tal como ocupando menos que cerca de 30 pixels. Um exemplo adicional de um recurso é um jogador ou um árbitro em um evento esportivo.

35

Um desafio para rastrear movimento de um objeto entre quadros em vídeo é oclusão do objeto em um ou mais quadros. Oclusão pode ser na forma do objeto ficar escondido atrás de um recurso no primeiro plano. Isto é referido como “oclusão real”. Por exemplo, em uma partida de tênis uma bola de tênis pode passar atrás de um jogador. Tal oclusão pode ser referida de várias maneiras, tal como, por exemplo, o objeto sendo escondido, bloqueado ou encoberto. Em um outro exemplo, oclusão pode ser na forma de um plano de fundo que torna a determinação da posição do objeto difícil ou impossível. Isto é referido como “oclusão virtual”. Por exemplo, uma bola de tênis pode passar na frente de um plano de fundo desordenado, tal como uma multidão que inclui inúmeros objetos aproximadamente do mesmo tamanho e cor da bola de tênis, de maneira que seleção da bola a partir dos outros objetos é difícil ou impossível. Em um outro exemplo, uma bola pode passar na frente de um campo da mesma cor que a bola, de maneira que a localização da bola é impossível ou difícil de determinar. Oclusão, incluindo desordem, torna difícil formar uma estimativa de probabilidade precisa de partículas em um filtro de partículas. Oclusão, incluindo desordem, frequentemente resulta em ambiguidade no rastreamento de objeto.

Estes problemas são frequentemente maiores para pequenos objetos, ou para objetos se deslocando rapidamente. Isto é porque, por exemplo, as localizações de um pequeno objeto em imagens sucessivas (por exemplo, quadros) em um vídeo frequentemente não se sobrepõem uma à outra. Quando as localizações não se sobrepõem, o objeto propriamente dito não se sobrepõe, significando que o objeto se deslocou pelo menos a sua própria largura no intervalo de tempo entre as duas imagens sucessivas. A falta de sobreposição frequentemente torna mais difícil descobrir o objeto na próxima imagem, ou ter uma alta confiança de que o objeto foi encontrado.

Ambiguidade em rastreamento de objeto não é limitada a pequenos objetos. Por exemplo, um plano de fundo desordenado pode incluir recursos similares a um objeto. Nesse caso, independente de tamanho de objeto, ambiguidade em rastreamento pode resultar.

Determinação de se um objeto está ocluído também pode ser desafiadora. Por exemplo, um método conhecido de determinar oclusão de objeto é uma razão típico/atípico. Com pequenos objetos e/ou um plano de fundo desordenado, a razão típico/atípico pode ser difícil de determinar.

Uma implementação aborda estes desafios pela formação de uma superfície métrica em um quadro baseado em partículas. Uma outra implementação aborda estes desafios ao empregar e avaliar estimativas de movimento em um quadro baseado em partículas. Uma outra implementação aborda estes desafios pelo emprego de múltiplas hipóteses em estimativa de probabilidade.

Em um quadro baseado em partículas, uma simulação de Monte Carlo é tipicamente conduzida sobre inúmeras partículas. As partículas podem representar, por exemplo,

possíveis localizações diferentes de um objeto em um quadro. Uma partícula particular pode ser selecionada com base na probabilidade determinada de acordo com uma simulação de Monte Carlo. Um filtro de partículas é um quadro baseado em partículas exemplar. Em um filtro de partículas, inúmeras partículas são geradas, representando possíveis estados, os quais podem corresponder a possíveis localizações de um objeto em uma imagem. Uma probabilidade, também referida como um peso, é associada com cada partícula no filtro de partículas. Em um filtro de partículas, partículas tendo uma baixa probabilidade ou baixo peso são tipicamente eliminadas em uma ou mais etapas de reamostragem. Um estado representando um resultado de um filtro de partículas pode ser uma média ponderada de partículas, por exemplo.

Referindo-se à figura 1, em uma implementação um sistema 100 inclui um estimador de estado 110 que pode ser implementado, por exemplo, em um computador. O estimador de estado 110 inclui um módulo de algoritmo de partículas 120, um módulo de modo local 130, e um módulo adaptador de número 140. O módulo de algoritmo de partículas 120 executa um algoritmo baseado em partículas, tal como, por exemplo, um filtro de partículas (PF), para estimar estados de um sistema dinâmico. O módulo de modo local 130 aplica um mecanismo de procura de modo local, tal como, por exemplo, pela execução de uma análise de mudança de meio, às partículas de um PF. O módulo adaptador de número 140 modifica o número de partículas usado no algoritmo baseado em partículas, tal como, por exemplo, pela aplicação de um processo de amostragem de distância de Kullback-Leibler (KLD) às partículas de um PF. Em uma implementação, o filtro de partículas pode amostrar adaptativamente dependendo do tamanho do espaço de estado onde as partículas são encontradas. Por exemplo, se as partículas forem todas encontradas em uma pequena parte do espaço de estado, um menor número de partículas pode ser amostrado. Se o espaço de estado for grande, ou se a incerteza de estado for alta, um maior número de partículas pode ser amostrado. Os módulos 120-140 podem ser, por exemplo, implementados separadamente ou integrados em um único algoritmo.

O estimador de estado 110 acessa como entrada tanto um estado inicial 150 quanto uma entrada de dados 160, e fornece como saída um estado estimado 170. O estado inicial 150 pode ser determinado, por exemplo, por meio de um detector de estado inicial ou por meio de um processo manual. Exemplos mais específicos são fornecidos ao considerar um sistema para o qual o estado é a localização de um objeto em uma imagem em uma sequência de imagens digitais, tal como um quadro de um vídeo. Em um sistema como este, a localização inicial de objeto pode ser determinada, por exemplo, por meio de um processo automatizado de detecção de objeto usando detecção de borda e comparação de modelo, ou manualmente por meio de um usuário visualizando o vídeo. A entrada de dados 160 pode ser, por exemplo, uma sequência de imagens de vídeo. O estado estimado 170 pode ser,

por exemplo, uma estimativa da posição de uma bola em uma imagem de vídeo particular.

Na figura 2 está mostrado um aparelho exemplar 190 para implementar o estimador de estado 110 da figura 1. O aparelho 190 inclui um dispositivo de processamento 180 que recebe o estado inicial 150 e a entrada de dados 160, e fornece como saída um estado estimado 170. O dispositivo de processamento 180 acessa um dispositivo de armazenamento 185, o qual pode executar dados de armazenamento se relacionando com uma imagem particular em uma sequência de imagens digitais.

O estado estimado 170 pode ser usado para uma variedade de propósitos. Para fornecer contexto adicional, diversas aplicações estão descritas usando as figuras 3 e 4.

Referindo-se à figura 3, em uma implementação um sistema 200 inclui um codificador 210 acoplado a um dispositivo de transmissão/armazenamento 220. O codificador 210 e o dispositivo de transmissão/armazenamento 220 podem ser implementados, por exemplo, em um computador ou em um codificador de comunicações. O codificador 210 acessa o estado estimado 170 fornecido pelo estimador de estado 110 do sistema 100 na figura 1, e acessa a entrada de dados 160 usada pelo estimador de estado 110. O codificador 210 codifica a entrada de dados 160 de acordo com um ou mais de uma variedade de algoritmos de codificação, e fornece uma saída de dados codificados 230 para o dispositivo de transmissão/armazenamento 220.

Adicionalmente, o codificador 210 usa o estado estimado 170 para codificar diferencialmente diferentes partes da entrada de dados 160. Por exemplo, se o estado representar a posição de um objeto em um vídeo, o codificador 210 pode codificar uma parte do vídeo correspondendo à posição estimada usando um primeiro algoritmo de codificação, e pode codificar uma outra parte do vídeo não correspondendo à posição estimada usando um segundo algoritmo de codificação. O primeiro algoritmo, por exemplo, pode fornecer mais redundância de codificação do que o segundo algoritmo de codificação, de maneira que a posição estimada do objeto (e esperançosamente o objeto propriamente dito) será esperada ser reproduzida com mais detalhes e resolução do que outras partes do vídeo.

Assim, por exemplo, uma transmissão de uma maneira geral de baixa resolução pode fornecer maior resolução para o objeto que está sendo rastreado, permitindo, por exemplo, a um usuário ver uma bola de golfe em uma partida de golfe com maior facilidade. Uma implementação como esta permite a um usuário observar a partida de golfe em um dispositivo móvel através de um vínculo de baixa largura de banda (baixa taxa de dados). O dispositivo móvel pode ser, por exemplo, um telefone celular ou um assistente digital pessoal. A taxa de dados é mantida baixa pela codificação do vídeo da partida de golfe em uma baixa taxa de dados, mas usando bits adicionais, quando comparada às outras partes das imagens, para codificar a bola de golfe.

O dispositivo de transmissão/armazenamento 220 pode incluir um ou mais de um

dispositivo de armazenamento ou um dispositivo de transmissão. Desta maneira, o dispositivo de transmissão/armazenamento 220 acessa os dados codificados 230 e ou transmite os dados 230 ou armazena os dados 230.

Referindo-se à figura 4, em uma implementação um sistema 300 inclui um dispositivo de processamento 310 acoplado a um dispositivo de armazenamento local 315 e acoplado a um mostrador 320. O dispositivo de processamento 310 acessa o estado estimado 170 fornecido pelo estimador de estado 110 do sistema 100 na figura 1, e acessa a entrada de dados 160 usada pelo estimador de estado 110. O dispositivo de processamento 310 usa o estado estimado 170 para aprimorar a entrada de dados 160 e fornece uma saída de dados aprimorados 330. O dispositivo de processamento 310 pode produzir dados, incluindo o estado estimado, a entrada de dados, e elementos dos mesmos para serem armazenados no dispositivo de armazenamento local 315, e pode recuperar tais dados do dispositivo de armazenamento local 315. O mostrador 320 acessa a saída de dados aprimorados 330 e exibe os dados aprimorados no mostrador 320.

Referindo-se à figura 5, um diagrama 400 inclui uma função de distribuição de probabilidade 410 para um estado de um sistema dinâmico. O diagrama 400 representa de forma pictorial várias funções executadas por uma implementação do estimador de estado 110. O diagrama 400 representa uma ou mais funções em cada um dos níveis A, B, C e D.

O nível A representa a geração das quatro partículas A1, A2, A3 e A4 por meio de um PF. Por conveniência, linhas tracejadas verticais separadas indicam a posição da função de distribuição de probabilidade 410 acima de cada uma das quatro partículas A1, A2, A3 e A4.

O nível B representa a mudança das quatro partículas A1-A4 para as partículas correspondentes B1-B4 por meio de um algoritmo de busca de modo local com base em uma análise de mudança de meio. Por conveniência, linhas cheias verticais indicam a posição da função de distribuição de probabilidade 410 acima de cada uma das quatro partículas B1, B2, B3 e B4. A mudança de cada uma das partículas A1-A4 está mostrada graficamente pelas setas correspondentes MS1-MS4, as quais indicam o movimento de partícula das posições indicadas pelas partículas A1-A4 para as posições indicadas pelas partículas B1-B4, respectivamente.

O nível C representa as partículas ponderadas C2-C4, as quais têm as mesmas posições que as partículas B2-B4, respectivamente. As partículas C2-C4 têm tamanhos variando indicando uma ponderação que foi determinada para as partículas B2-B4 no PF. O nível C também reflete uma redução no número de partículas, de acordo com um processo de amostragem, tal como um processo de amostragem KLD, no qual a partícula B1 foi descartada.

O nível D representa três novas partículas geradas durante um processo de rea-

mostragem. O número de partículas geradas no nível D é igual ao número de partículas no nível C, tal como indicado por uma seta R (R se mantém para reamostragem).

Referindo-se agora à figura 6, está ilustrado um fluxo de processo de alto nível 600 de um método para determinar uma localização de um objeto em uma imagem em uma sequência de imagens digitais. Uma trajetória do objeto pode ser estimada com base na informação de localização de quadros anteriores 605. Estimativa de trajetória é conhecida para os versados na técnica. Um filtro de partículas pode ser operado 610. Várias implementações de filtro de partículas são descritas a seguir. A localização do objeto predito por uma saída do filtro de partículas pode ser verificada para oclusão 615. Implementações de métodos de verificar oclusão são explicadas a seguir. Se oclusão for encontrada 620, então uma posição pode ser determinada usando projeção e interpolação de trajetória 625. Implementações de determinação de posição são explicadas a seguir com relação à figura 16, por exemplo. Se oclusão não for encontrada, então a saída de filtro de partículas é usada para determinar posição de partícula 630. Se oclusão não for encontrada, então o modelo é verificado para deslocamento 635. Deslocamento se refere a uma mudança no modelo, tal como pode ocorrer, por exemplo, se o objeto está se movendo para mais longe ou mais perto, ou mudando de cor. Se for encontrado deslocamento acima de um limiar 635, então um modelo de objeto não é atualizado 640. Isto pode ser útil, por exemplo, porque grandes valores de deslocamento podem indicar uma oclusão parcial. Atualizar o modelo com base em uma oclusão parcial pode produzir um modelo inferior para ser usado. De outro modo, se deslocamento não estiver acima do limiar, então um modelo pode ser atualizado 645. Quando ocorrem pequenas mudanças (pequenos valores de deslocamento), existe tipicamente mais confiabilidade ou confiança de que as mudanças são mudanças verdadeiras para o objeto e não mudanças causadas, por exemplo, por oclusão.

Referindo-se agora à figura 7, será discutido um processo 500 de implementar um filtro de partículas. O processo 500 inclui acessar um conjunto inicial de partículas e fatores de ponderação cumulativos de um estado anterior 510. Fatores de ponderação cumulativos podem ser gerados de um conjunto de pesos de partícula e tipicamente permitem processamento mais rápido. Deve-se notar que na primeira vez através do processo 500 o estado anterior será o estado inicial e o conjunto inicial de partículas e pesos (fatores de ponderação cumulativos) necessitará ser gerado. O estado inicial pode ser fornecido, por exemplo, como o estado inicial 150 (da figura 1).

Referindo-se de novo à figura 7, uma variável de controle de laço "it" é inicializada 515 e um laço 520 é executado repetidamente antes de determinar o estado atual. O laço 520 usa a variável de controle de laço "it", e executa "reiteração" número de vezes. Dentro do laço 520, cada partícula no conjunto inicial de partículas é tratada separadamente em um laço 525. Em uma implementação, o PF é aplicado ao vídeo de uma partida de tênis para

rastrear uma bola de tênis, e o laço 520 é executado um número de vezes predeterminado (o valor da variável de repetição de laço “reiteração”) para cada novo quadro. Cada repetição do laço 520 é esperada para melhorar a posição das partículas, de maneira que quando a posição da bola de tênis é estimada para cada quadro, a estimativa é suposta estar baseada em boas partículas.

O laço 525 inclui selecionar uma partícula com base em um fator de ponderação cumulativo 530. Este é um método para selecionar a localização de partícula remanescente com o maior peso, tal como é conhecido. Deve-se notar que muitas partículas podem estar na mesma localização, em cujo caso tipicamente somente é necessário executar o laço 525 uma vez para cada localização. O laço 525 inclui então atualizar a partícula ao predizer uma nova posição no espaço de estado para a partícula selecionada 535. A predição usa o modelo dinâmico do PF. Esta etapa será explicada com mais detalhes a seguir.

O modelo dinâmico caracteriza a mudança do estado de objeto entre quadros. Por exemplo, um modelo de movimento, ou estimativa de movimento, que reflete a cinemática do objeto pode ser empregado. Em uma implementação, um modelo de velocidade constante fixada com variância de ruído fixada pode ser encaixado às posições de objeto em quadros passados.

O laço 525 inclui então determinar o peso da partícula atualizada usando o modelo de medição do PF 540. Determinar o peso envolve, como é conhecido, analisar os dados observados/medidos (por exemplo, os dados de vídeo no quadro atual). Continuando a implementação de partida de tênis, dados do quadro atual, na localização indicada pela partícula, são comparados com dados da última localização da bola de tênis. A comparação pode envolver, por exemplo, analisar histogramas de cor ou executar detecção de borda. O peso determinado para a partícula é baseado em um resultado da comparação. A operação 540 também inclui determinar o fator de ponderação cumulativo para a posição de partícula.

O laço 525 inclui então determinar se mais partículas são para ser processadas 542. Se mais partículas forem para ser processadas, o laço 525 é repetido e o processo 500 salta para a operação 530. Depois de executar o laço 525 para cada partícula no conjunto de partículas inicial (ou “antigo”), um conjunto completo de partículas atualizadas é gerado.

O laço 520 inclui então gerar um “novo” conjunto de partículas e novos fatores de ponderação cumulativos usando um algoritmo de reamostragem 545. O algoritmo de reamostragem é baseado nos pesos das partículas, focalizando assim em partículas com maiores pesos. O algoritmo de reamostragem produz um conjunto de partículas em que cada uma tem o mesmo peso individual, mas certas localizações tipicamente têm muitas partículas posicionadas nessas localizações. Assim, as localizações de partículas tipicamente têm diferentes fatores de ponderação cumulativos.

Reamostragem tipicamente também ajuda a reduzir o problema de degeneração

que é comum em PFs. Existem diversos modos para reamostrar, tais como reamostragem polinomial, residual, estratificada e sistemática. Uma implementação usa reamostragem residual porque reamostragem residual não é sensível à ordem de partícula.

5 O laço 520 continua ao incrementar a variável de controle de laço "it" 550 e comparar "it" com a variável de repetição "reiteração" 555. Se uma outra repetição através do laço 520 for necessária, então o novo conjunto de partículas e seus fatores de ponderação cumulativos são tornados disponíveis 560.

Depois de executar o laço 520 um "reiterado" número de vezes, o conjunto de partículas é esperado ser um "bom" conjunto de partículas, e o estado atual é determinado 565.
10 O novo estado é determinado, como é conhecido, pelo cálculo da média das partículas no novo conjunto de partículas.

Referindo-se agora à figura 8, será explicada uma outra implementação de um fluxo de processo incluindo um filtro de partículas. O fluxo de processo global é similar ao fluxo de processo descrito anteriormente com referência à figura 7, e elementos comuns à figura 7 e
15 à figura 8 não serão descritos aqui detalhadamente. O processo 800 inclui acessar um conjunto inicial de partículas e fatores de ponderação cumulativos de um estado anterior 805. Uma variável de controle de laço "it" é inicializada 810 e um laço é executado repetidamente antes de determinar o estado atual. No laço, uma partícula é selecionada de acordo com um fator de ponderação cumulativo. O processo atualiza então a partícula ao predizer uma nova
20 posição no espaço de estado para a partícula selecionada 820. A predição usa o modelo dinâmico do PF.

O modo local da partícula é então procurado usando uma superfície de correlação, tal como uma superfície de correlação baseada em SSD 825. Um local mínimo da SSD é identificado, e então a posição da partícula é mudada para o local mínimo identificado da
25 SSD. Outras implementações, usando uma superfície apropriada, identificam um local máximo da superfície e mudam a posição da partícula para o local máximo identificado. O peso da partícula deslocada é então determinado 830 a partir do modelo de medição. A título de exemplo, uma superfície de correlação e múltiplas hipóteses podem ser empregadas na computação do peso, tal como descrito a seguir. Se existirem mais partículas a processar
30 835, então o laço retorna para escolher uma partícula. Se todas as partículas tiverem sido processadas, então as partículas são reamostradas com base nos novos pesos, e um novo grupo de partículas é gerado 840. A variável de controle de laço "it" é incrementada 845. Se "it" for menor do que o limiar de repetição 850, então o processo muda para o grupo de partículas antigo 870, e repete o processo.

35 Se a repetição final tiver sido completada, uma etapa adicional é conduzida antes de obter o estado atual. Um indicador de oclusão para o objeto no quadro anterior é verificado 855. Se o indicador de oclusão mostrar oclusão no quadro anterior, então um subconjunto

to de partículas é considerado para seleção do estado atual 860. O subconjunto de partículas é selecionado pelas partículas tendo o peso mais alto. Em uma modalidade, o subconjunto de partículas é a partícula tendo o peso mais alto. Se mais de uma partícula tiver o mesmo peso mais alto, então todas as partículas tendo o peso mais alto são incluídas no subconjunto. O estado da partícula pode ser suposto um estado de detecção. A seleção de um subconjunto de partículas é feita porque oclusão afeta negativamente a confiabilidade de partículas tendo menores pesos. Se o indicador de oclusão mostrar que não existe oclusão no quadro anterior, então uma média do novo grupo de partículas pode ser usada para determinar o estado atual 865. Neste caso, o estado é um estado de rastreamento. Será percebido que a média pode ser ponderada de acordo com pesos de partícula. Também deve ser percebido que outras medidas estatísticas além de uma média (por exemplo, um meio) podem ser empregadas para determinar o estado atual.

Referindo-se à figura 9, é explicada uma implementação 900 do modelo dinâmico (820 da figura 8). No modelo dinâmico, informação de movimento de quadros anteriores pode ser empregada. Pelo uso de informação de movimento de quadros anteriores, as partículas provavelmente estarão mais próximas da posição real do objeto, aumentando assim eficiência, precisão ou ambas. No modelo dinâmico, como uma alternativa, um passo aleatório pode ser empregado na geração de partículas.

O modelo dinâmico pode empregar um modelo de espaço de estado para rastreamento de objeto pequeno. Um modelo de espaço de estado para rastreamento de objeto pequeno, para uma imagem, em uma sequência de imagens digitais, no tempo t , pode ser formulado como:

$$\begin{aligned} X_{t+1} &= f(X_t, \mu_t), \\ Z_t &= g(X_t, \xi_t), \end{aligned}$$

onde X_t representa o vetor de estado de objeto, Z_t é o vetor de observação, f e g são duas funções avaliadas por vetor (o modelo dinâmico e o modelo de observação, respectivamente), e μ_t e ξ_t representam o ruído de processo ou dinâmico, e ruído de observação respectivamente. Em estimativa de movimento, o vetor de estado de objeto é definido como $X = (x, y)$, onde (x, y) são as coordenadas do centro de uma janela de objeto. O movimento estimado é preferivelmente obtido de dados de quadros anteriores, e pode ser estimado a partir da equação de fluxo ótico. O movimento estimado para um objeto em uma imagem no tempo t pode ser V_t . O modelo dinâmico pode ser representado por:

$$X_{t+1} = X_t + V_t + \mu_t$$

A variância de ruído de predição μ_t pode ser estimada dos dados de movimento, tal como de uma medida de erro de estimativa de movimento. Um movimento residual da equação de fluxo ótico pode ser empregado. Alternativamente, a variância de ruído de predição pode ser um critério baseado em intensidade, tal como um residual de compensação de

movimento; entretanto, uma variância baseada em dados de movimento pode ser preferível a uma variância baseada em dados de intensidade.

Para cada partícula, um indicador de oclusão armazenado é lido, indicado pelo bloco 905. O indicador de oclusão indica se o objeto foi determinado para verificar oclusão no quadro anterior. Se a leitura do indicador 910 indicar que o objeto foi ocluído, então nenhuma estimativa de movimento é empregada no modelo dinâmico 915. Será percebido que oclusão reduz a precisão de estimativa de movimento. Um valor de variância de ruído de predição para a partícula pode ser estabelecido para um máximo 920. Ao contrário, se a leitura do indicador de oclusão mostrar que não existe oclusão no quadro anterior, então o processo usa estimativa de movimento 925 ao gerar partículas. Um método de variância de ruído de predição pode ser estimado 930, tal como a partir dos dados de movimento.

Referindo-se agora à figura 10, está ilustrada uma implementação de um fluxo de processo 1000 executado com relação a cada partícula em um modelo dinâmico dentro de um filtro de partículas, antes da amostragem. Inicialmente, é verificado um indicador de oclusão na memória 1005. O indicador de oclusão pode indicar oclusão do objeto no quadro anterior. Se oclusão do objeto no quadro anterior for encontrada 1010, então estimativa de movimento não é usada para o modelo dinâmico 1030, e a variância de ruído de predição para a partícula é estabelecida para um máximo 1035. Se o indicador de oclusão armazenado não indicar oclusão do objeto no quadro anterior, então estimativa de movimento é executada 1015.

Estimativa de movimento pode ser baseada no uso de posições do objeto em quadros passados na equação de fluxo ótico. A equação de fluxo ótico é conhecida para os versados na técnica. Depois da estimativa de movimento, detecção de falha 1020 é executada na localização de partícula resultante da estimativa de movimento. Várias métricas podem ser usadas para detecção de falha. Em uma implementação, uma média de uma diferença de intensidade absoluta entre a imagem de objeto tal como refletida no modelo e uma mancha de imagem centralizada em volta da localização de partícula derivada da estimativa de movimento pode ser calculada. Se a média exceder um limiar selecionado, então a estimativa de movimento é suposta ter falhado 1025, e nenhum uso é feito dos resultados de estimativa de movimento 1030 para a partícula. A variância de ruído de predição para a partícula pode ser estabelecida para o seu máximo 1035. Se a estimativa de movimento for suposta como não tendo falhado, então o resultado de estimativa de movimento é salvo 1040 como a predição para essa partícula. Variância de ruído de predição pode então ser estimada 1045. Por exemplo, a equação de fluxo ótico pode ser usada para fornecer um valor de movimento residual o qual pode ser usado como a variância de ruído de predição.

Referindo-se agora à figura 11, será discutida uma implementação de computar peso de partícula usando o modelo de medição. O método 1100 é executado com relação a

cada partícula. O método 1100 começa com o cálculo de uma superfície métrica, a qual pode ser uma superfície de correlação, tal como indicado pelo bloco 1105. Uma superfície métrica pode ser empregada para medir a diferença entre um modelo, ou modelo alvo, e a atual partícula candidata. Em uma implementação, uma superfície métrica pode ser gerada como se segue.

Uma métrica para a diferença entre o modelo e a partícula candidata pode ser uma superfície métrica, tal como uma superfície de correlação. Em uma implementação, uma superfície de soma dos quadrados das diferenças (SSD) é usada, a qual tem a seguinte fórmula:

$$Z_t = \arg \min_{X_t \in Viz} \sum_{\chi \in W} [T(\chi) - I(\chi + X_t)]^2$$

Aqui, W representa a janela de objeto, Viz é uma pequena vizinhança em volta do centro de objeto X_t . T é o modelo de objeto e I é a imagem no quadro atual. Em um pequeno objeto com um plano de fundo desordenado, esta superfície pode não representar uma estimativa precisa de uma probabilidade. Uma superfície de correlação exemplar adicional pode ser:

$$r(X_t) = \sum_{\chi \in W} [T(\chi) - I(\chi + X_t)]^2, X_t \in Viz$$

O tamanho da superfície de correlação pode ser variado. Dependendo da qualidade da estimativa de movimento, a qual pode ser determinada como o inverso da variância, o tamanho da superfície de correlação pode ser variado. Em geral, com maior qualidade de estimativa de movimento, a superfície de correlação pode ser feita menor.

Múltiplas hipóteses para o movimento da partícula podem ser geradas 1110 com base na superfície métrica. Hipóteses candidatas são associadas com um local mínimo ou máximo da superfície de correlação. Por exemplo, se J candidatos da superfície de correlação SSD forem identificados na área de suporte Viz , $J+1$ hipóteses podem ser definidas como:

$$H_0 = \{c_j = C : j = 1, \dots, j\},$$

$$H_j = \{c_j = T, c_i = C : i = 1, \dots, j, i \neq j\}, j = 1, \dots, j,$$

onde $c_j = T$ significa que o candidato de ordem j é associado com o casamento verdadeiro, $c_j = C$ de outro modo. A hipótese H_0 significa que nenhum dos candidatos está associado com o casamento verdadeiro. Nesta implementação, desordem é assumida estar distribuída uniformemente sobre a vizinhança Viz e de outro modo a medição orientada por casamento verdadeiro é uma distribuição gaussiana.

Com essas suposições, a probabilidade associada com cada partícula pode ser expressada por:

$$P(z_t | X_t) = q_0 U(\cdot) + C_N \sum_{j=1}^J q_j N(r_t, \sigma_t),$$

de tal maneira que $q_0 + \sum_{j=1 \dots J} q_j = 1$

onde C_N é um fator de normalização, q_0 é a probabilidade anterior da hipótese H_0 e q_j é a probabilidade para a hipótese H_j , $j=1, \dots, J$. Desta maneira, a medição de probabilidade usando o SSD é refinada considerando desordem pelo uso de múltiplas hipóteses.

5 Uma estimativa de variância de distribuição de resposta 1115 também é feita.

Uma determinação pode ser feita tal como para se a partícula está ocluída. Determinação de oclusão de partícula pode ser com base em uma avaliação baseada em intensidade 1120, tal como uma métrica SAD (soma das médias das diferenças), a qual pode ser usada para comparar um modelo de objeto e a partícula candidata. Tais avaliações são conhecidas para os versados na técnica. Com base na SAD, uma determinação pode ser feita tal como para partículas que muito provavelmente devem estar ocluídas. Avaliações de oclusão baseadas em intensidade são de computação relativamente barata, mas em um plano de fundo desordenado podem não ser altamente precisas. Por estabelecer um alto limiar, certas partículas podem ser determinadas para estar ocluídas usando uma avaliação baseada em intensidade 1125, e seus pesos estabelecidos para um mínimo 1130. Em tais casos, pode existir uma alta confiança de que oclusão ocorreu. Por exemplo, um limiar pode ser selecionado de tal maneira que o caso de oclusão real sem desordem é identificado, mas outros casos de oclusão não são identificados.

Se a avaliação baseada em intensidade não indicar oclusão, então uma determinação de oclusão de partícula probabilística pode ser feita 1135. A detecção de oclusão de partícula probabilística pode ser baseada nas múltiplas hipóteses geradas e na estimativa de variância de distribuição de resposta. Uma distribuição a pode ser gerada para aproximar a superfície SSD e oclusão é determinada (ou não) com base nessa distribuição usando um autovalor de uma matriz de covariância, tal como discutido a seguir.

25 Uma distribuição de resposta pode ser definida para aproximar uma distribuição de probabilidades na localização de casamento verdadeiro. Em outras palavras, uma probabilidade D de que a localização de partícula é uma localização de casamento verdadeiro pode ser:

$$D(X_t) = \exp(-\rho \cdot r(X_t)),$$

30 onde ρ é um fator de normalização. O fator de normalização pode ser escolhido para assegurar uma resposta máxima selecionada, tal como um máximo de 0,95. Uma matriz de covariância R_t associada com a medição Z_t é construída a partir da distribuição de resposta como

$$R_t = \left[\begin{array}{cc} \sum_{(x,y) \in Viz} D_t(x,y)(x-x_p)^2 & \sum_{(x,y) \in Viz} D_t(x,y)(x-x_p)(y-y_p) \\ \sum_{(x,y) \in Viz} D_t(x,y)(x-x_p)(y-y_p) & \sum_{(x,y) \in Viz} D_t(x,y)(y-y_p)^2 \end{array} \right] / N_R,$$

35 onde (x_p, y_p) é o centro de janela de cada candidato e $N_R = \sum_{(x,y) \in Viz} D_t(x,y)$ é o fator

de normalização de covariância. Os recíprocos dos autovalores de R_t podem ser usados como uma métrica de confiança associada com o candidato. Em uma implementação, o autovalor máximo de R_t pode ser comparado a um limiar; se o autovalor máximo exceder o limiar, oclusão é detectada. Em resposta a uma detecção de oclusão 1140, à partícula é

5 dado o menor peso disponível 1130, o qual de uma maneira geral será um peso diferente de zero. Se oclusão não for detectada, uma probabilidade pode ser calculada.

Em uma implementação, se oclusão for detectada, em vez de estabelecer o peso ou probabilidade para o menor valor, a probabilidade de partícula pode ser gerada com base em intensidade e movimento, mas sem consideração à trajetória. Por outro lado, se oclusão

10 não for detectada, probabilidade para a partícula pode ser gerada com base em intensidade, por exemplo.

Em uma implementação, pesos a ser designados para partículas podem ser baseados pelo menos em parte em consideração a pelo menos uma parte da imagem próxima da posição indicada pela partícula. Por exemplo, para uma dada partícula, uma mancha, tal

15 como um bloco 5x5 de pixels de um modelo de objeto, é comparada com a posição indicada pela partícula e com outras áreas. A comparação pode ser baseada em uma matriz de soma das diferenças absolutas (SAD) ou em um histograma, particularmente para objetos maiores. O modelo de objeto é assim comparado à imagem em volta da posição indicada pela partícula. Se as comparações fora de posição forem suficientemente diferentes, então o peso

20 designado para a partícula pode ser maior. Por outro lado, se a área indicada pela partícula for mais similar às outras áreas, então o peso da partícula pode ser reduzido correspondentemente. Uma superfície de correlação, tal como uma SSD, pode ser gerada que modela as áreas fora de posição, com base nas comparações.

Se o resultado da determinação for que a partícula não está ocluída, então uma estimativa pode ser feita da probabilidade de trajetória 1145. Com relação à estimativa do peso de partícula, uma determinação ponderada pode ser empregada 1150.

A determinação ponderada pode incluir uma ou mais de probabilidade de intensidade (por exemplo, comparação de modelo), probabilidade de movimento (por exemplo, uma extrapolação linear de localizações passadas de objeto) e probabilidade de trajetória. Estes

30 fatores podem ser empregados para determinar uma probabilidade ou peso de cada partícula no filtro de partículas. Em uma implementação, uma suposição pode ser feita de que movimento de câmera não afeta uniformidade de trajetória e, portanto, não afeta a probabilidade de trajetória. Em uma implementação, uma probabilidade de partícula pode ser definida como:

$$P(z_t|X_t) = P(Z_t^{int}|X_t)P(Z_t^{mov}|X_t)P(Z_t^{trj}|X_t),$$

onde $Z_t = \{Z_t^{int}, Z_t^{mov}, Z_t^{trj}\}$, em que uma medição de intensidade, a qual pode ser baseada em superfície SSD, é Z_t^{int} , uma probabilidade de movimento é dada por Z_t^{mov} e uma

probabilidade de trajetória é dada por Z_t^{trj} . Estes três valores podem ser assumidos para ser independentes. O cálculo da probabilidade de intensidade $P(Z_t^{int} | X_t)$ é conhecido para os versados na técnica.

A probabilidade de movimento pode ser calculada com base na diferença entre a mudança de posição da partícula (velocidade) e a mudança média na posição do objeto através de quadros recentes:

$$d_{mov}^2 = \left(\Delta x_t | - \overline{\Delta x} \right)^2 + \left(\Delta y_t | - \overline{\Delta y} \right)^2, t > 1$$

onde $(\Delta x_t, \Delta y_t)$ é a mudança de posição da partícula com relação a (X_{t-1}, Y_{t-1}) , e $(\overline{\Delta x}, \overline{\Delta y})$ é a velocidade média de objeto através de uma seleção de quadros recentes, isto é,

$$\overline{\Delta x} = \sum_{s=1}^{t-1} |x_s - x_{s-1}| / (t-1), \quad \overline{\Delta y} = \sum_{s=1}^{t-1} |y_s - y_{s-1}| / (t-1).$$

Consequentemente a probabilidade de movimento pode ser calculada com base em uma distância d_{mov} (por exemplo, a distância euclidiana) entre a posição predita pelo modelo dinâmico e a posição de partícula como

$$P(Z_t^{mov} | X_t) = \frac{1}{\sqrt{2\pi\sigma_{mov}^2}} \exp\left(-\frac{d_{mov}^2}{2\sigma_{mov}^2}\right).$$

Em uma implementação, uma probabilidade de uniformidade de trajetória pode ser estimada a partir da proximidade da partícula a uma trajetória que é calculada com base em uma sequência de posições do objeto em quadros recentes do vídeo. A função de trajetória pode ser representada como $y = f(x)$, cuja forma paramétrica pode ser:

$$y = \sum_{i=0}^m a_i x^i,$$

onde a_i representa os coeficientes polinomiais e m é a ordem da função polinomial (por exemplo, $m = 2$). No cálculo da função de trajetória a fórmula pode ser modificada. Uma primeira modificação pode envolver desconsiderar ou descontar posições de objeto, se a posição de objeto for determinada para corresponder a um estado ocluído no quadro passado particular. Segundo, um fator de ponderação, o qual pode ser chamado de fator de esquecimento, é calculado para ponderar a proximidade da partícula à trajetória. Quanto mais quadros nos quais o objeto é ocluído tanto menos segura é a trajetória estimada, e consequentemente tanto maior o fator de esquecimento.

O “fator de esquecimento” é simplesmente um valor de confiança. Um usuário pode designar um valor para o fator de esquecimento com base em uma variedade de considerações. Tais considerações podem incluir, por exemplo, se o objeto está ocluído em uma imagem anterior, o número de imagens anteriores nas quais o objeto está ocluído, o número de imagens anteriores consecutivas nas quais o objeto está ocluído, ou a confiabilidade de dados não ocluídos. Cada imagem pode ter um fator de esquecimento diferente.

Em uma implementação exemplar, a probabilidade de uniformidade de trajetória pode ser dada por:

$$P(Z_t^{trj} | X_t) = \frac{1}{\sqrt{2\pi\sigma_{trj}}} \exp\left(-\frac{[d_{trj} / (\lambda_f)^{-o_{cl}}]^2}{2\sigma_{trj}^2}\right)$$

onde o valor de proximidade é $d_{trj} = |y - f(x)|$, λ_f é a razão de esquecimento selecionada manualmente, $0 < \lambda_f < 1$ (por exemplo, $\lambda_f = 0,9$), e t_{ocl} é o número de quadros recentes nos quais o objeto está ocluído.

Em uma implementação, se uma determinação for feita de que o objeto está ocluído no quadro precedente, então uma probabilidade de partícula pode ser determinada com base em uma probabilidade de intensidade e em uma probabilidade de trajetória, mas não considerando uma probabilidade de movimento. Se uma determinação for feita de que o objeto não está ocluído no quadro precedente, então uma probabilidade de partícula pode ser determinada com base em uma probabilidade de intensidade e em uma probabilidade de movimento, mas não considerando uma probabilidade de trajetória. Isto pode ser vantajoso porque, quando a localização do objeto é conhecida no quadro anterior, de forma típica existe relativamente pouco benefício em fornecer restrições de trajetória. Além disso, incorporar restrições de trajetória pode violar a suposição de cadeia de Markov temporal, isto é, o uso de restrições de trajetória apresenta o estado seguinte dependendo do estado de quadros a não ser o quadro imediatamente precedente. Se o objeto estiver ocluído, ou uma determinação tiver sido feita de que estimativa de movimento estará abaixo de um limiar, então tipicamente não existe nenhum benefício em incluir probabilidade de movimento na determinação de probabilidade de partícula. Nesta implementação, a probabilidade de partícula pode ser expressada por:

$$P(Z_t | X_t) = P(Z_t^{int} | X_t) P(Z_t^{mov} | X_t)^{O_t-1} P(Z_t^{trj} | X_t)^{1-O_t-1}$$

onde $O_t = 0$ se o objeto estiver ocluído e 1 de outro modo.

Referindo-se agora à figura 12, está mostrada uma ilustração de um encaixe exemplar de uma trajetória de objeto para localizações de objeto em quadros de um vídeo. Os elementos 1205, 1206 e 1207 representam localizações de um pequeno objeto em três quadros de um vídeo. Os elementos 1205, 1206 e 1207 estão em uma zona 1208 e não estão ocluídos. Os elementos 1230 e 1231 representam localizações de um pequeno objeto em dois quadros do vídeo, depois dos quadros representados pelos elementos 1205, 1206 e 1207. Os elementos 1230 e 1231 estão na zona 1232 e foi determinado estar ocluídos, e assim existe um alto nível de incerteza a respeito das localizações determinadas. Assim, na figura 12, $t_{ocl} = 2$. Uma trajetória real 1210 está mostrada, a qual é projetada para uma trajetória predita 1220.

Referindo-se agora à figura 13, está ilustrado um fluxo de processo de uma imple-

mentação de um modelo. No começo do fluxo de processo da figura 13, um novo estado de um objeto foi estimado, tal como por meio de um filtro de partículas. O novo estado estimado corresponde, por exemplo, a uma localização estimada de um objeto em um novo quadro. O fluxo de processo 1300 da figura 13 pode ser empregado para determinar se é para reutilizar um modelo existente na estimativa do estado para o próximo quadro seguinte. Tal como indicado pela etapa 1305, detecção de oclusão é executada na nova localização estimada do objeto no quadro atual. Se oclusão for detectada 1310, então um indicador de oclusão é estabelecido na memória 1330. Esta indicação pode ser empregada no filtro de partículas para o quadro seguinte, por exemplo. Se oclusão não for detectada, então o fluxo de processo prossegue para detectar deslocamento 1315. Em uma implementação, deslocamento pode ser na forma de um movimento residual entre a imagem do objeto no novo quadro e o modelo inicial. Se o deslocamento exceder um limiar 1320, então o modelo não é atualizado 1335. Se deslocamento não exceder um limiar então o modelo pode ser atualizado 1325, com uma imagem de janela de objeto do quadro atual. Parâmetros de movimento de objeto também podem ser atualizados.

Referindo-se agora à figura 14, está ilustrado um fluxograma de uma implementação alternativa para o processo 1300 para atualizar modelos de objeto e refinar estimativas de posição. No processo 1400, depois da determinação do estado atual de objeto, é executada detecção de oclusão para a localização de objeto determinada e o quadro atual 1405. Se oclusão for detectada 1410, então a posição de objeto estimada pode ser modificada. Tal modificação pode ser útil porque, por exemplo, a oclusão pode reduzir a confiança de que a localização de objeto determinada é precisa. Assim, uma estimativa de posição refinada pode ser útil. Em um exemplo, a determinação de oclusão pode ser baseada na existência de desordem, e a localização de objeto determinada pode ser realmente a localização de algum da desordem.

A modificação pode ser implementada usando informação relacionada a uniformidade de trajetória. Uma posição de objeto pode ser projetada em uma determinada trajetória 1415 usando informação de dados de posição em quadros anteriores. Uma projeção de linha reta usando velocidade constante, por exemplo, pode ser empregada. A posição pode ser refinada 1420.

Referindo-se à figura 15, é fornecida uma ilustração de um processo de projetar uma localização de objeto em uma trajetória e refinar a localização. Uma trajetória 1505 está mostrada. A posição 1510 representa uma posição de objeto em um quadro anterior. O ponto de dados 1515 representa a posição X_j em um quadro anterior no tempo j . O ponto de dados 1520 representa uma posição X em um quadro anterior no tempo i . Os pontos de dados 1510, 1515 e 1520 representam posições de objetos não ocluídos, e assim são dados de qualidade relativamente alta. Os pontos de dados 1525, 1530, 1535, 1540 representam

posições do objeto em quadros anteriores, mas sujeitos a oclusão. Desta maneira, estes pontos de dados podem ser desconsiderados ou receber um menor peso nos cálculos de trajetória. A trajetória 1505 foi desenvolvida anteriormente baseada no encaixe destes pontos de dados, sujeitos a ponderação para oclusão de certos pontos de dados.

5 Um cálculo inicial da posição do objeto no quadro atual, isto é, no tempo atu, pode ser determinado usando uma linha reta e velocidade constante, usando a fórmula:

$$\hat{X}_{atu} = X_i + (X_i - X_j) * (atu - i) / (i - j).$$

Isto está representado por meio de uma projeção de linha reta 1550 (também referida como uma extrapolação linear) para obter uma localização estimada inicial de quadro atual 1545 (também referida como uma estimativa de localização linear). A localização estimada inicial de quadro atual pode então ser projetada na trajetória calculada como \tilde{X}_{atu} (também referida como um ponto de projeção), o qual é o ponto na trajetória mais próximo a \hat{X}_{atu} . A projeção pode usar a fórmula:

$$\hat{X}_{atu} = (1 - \lambda_f^{t_ocl}) \hat{X}_{atu} + \tilde{X}_{atu} * \lambda_f^{t_ocl},$$

15 onde λ_f é a razão de esquecimento, $0 < \lambda_f < 1$ (por exemplo, $\lambda_f = 0,9$), e t_ocl é o número de quadros em que o objeto foi ocluído desde a última vez que ele esteve visível. Em uma implementação, uma projeção pode ser um ponto na trajetória interpolada entre \hat{X}_{atu} e \tilde{X}_{atu} . Assim, a projeção será em uma linha entre \hat{X}_{atu} e \tilde{X}_{atu} . Em uma implementação como esta, a projeção pode ser representada por:

$$20 \quad X_{atu} = (1 - \lambda_f^{t_ocl}) \hat{X}_{atu} + \tilde{X}_{atu} * \lambda_f^{t_ocl}.$$

Na figura 15, o objeto foi ocluído nos dois últimos quadros, tal como representado pelas posições 1530, 1535, $t_ocl = 2$. A aplicação desta fórmula desloca de uma maneira geral a localização de objeto para uma posição interpolada entre a trajetória e a projeção de linha reta. À medida que t_ocl se torna maior a trajetória é menos certa, e a localização fica mais próxima da projeção de linha reta. No exemplo dado pela figura 15, a posição interpolada 1540 é determinada. A posição 1540 está ocluída, já que ela está dentro de uma zona ocluída 1545.

Referindo-se de novo à figura 14, será explicado o fluxo de processo quando o resultado da verificação para oclusão resulta em nenhuma oclusão encontrada. Deslocamento do modelo de objeto é determinado 1425. Deslocamento do modelo pode ser detectado pela aplicação de estimativa de movimento tanto ao modelo atual quanto ao modelo inicial. Os resultados são comparados. Se a diferença entre os dois modelos depois da aplicação de estimativa de movimento estiver acima de um limiar 1430, então ocorreu deslocamento. Nesse caso, então o modelo anterior não é atualizado 1445, e um novo modelo é obtido. Se a diferença não estiver acima de um limiar, então o modelo é atualizado 1435.

O fluxo de processo também inclui atualização do indicador de oclusão na memória 1440. O indicador de oclusão para o quadro anterior será então verificado no filtro de partículas durante a estimativa de posição de objeto para o próximo quadro.

Referindo-se agora à figura 16, um método 1600 inclui formar uma superfície métrica em um quadro baseado em partículas para rastrear um objeto 1605, a superfície métrica se relacionando com uma imagem particular em uma sequência de imagens digitais. Múltiplas hipóteses são formadas de uma localização do objeto na imagem particular com base na superfície métrica 1610. A localização do objeto é estimada com base nas probabilidades das múltiplas hipóteses 1615.

Referindo-se agora à figura 17, um método 1700 inclui avaliar uma estimativa de movimento para um objeto em uma imagem particular em uma sequência de imagens digitais 1705, a estimativa de movimento sendo baseada em uma imagem anterior na sequência. Pelo menos uma estimativa de localização é selecionada para o objeto com base em um resultado da avaliação 1710. A estimativa de localização é parte de um quadro baseado em partículas para rastrear o objeto.

Referindo-se agora à figura 18, um método 1800 inclui selecionar uma partícula em um quadro baseado em partículas usado para rastrear um objeto entre imagens em uma sequência de imagens digitais 1805, a partícula tendo uma localização. O método 1800 inclui acessar uma superfície que indica a extensão para a qual uma ou mais partículas casam com o objeto 1810. O método 1800 inclui adicionalmente determinar uma posição na superfície 1815, a posição sendo associada com a partícula selecionada, e indicar a extensão para a qual a partícula selecionada casa com o objeto. O método 1800 inclui associar um local mínimo ou máximo da superfície com a posição determinada 1820. O método 1800 também inclui deslocar a localização da partícula selecionada para corresponder ao local mínimo ou máximo determinado 1825.

Referindo-se agora à figura 19, um método 1900 inclui formar um modelo de objeto 1905 para um objeto em uma sequência de imagens digitais. O método 1900 também inclui formar uma estimativa de uma localização do objeto 1910 em uma imagem particular na sequência, a estimativa sendo formada usando um quadro baseado em partículas. O modelo de objeto é comparado a uma parte da imagem particular na localização estimada 1915. É determinado se é para atualizar o modelo de objeto dependendo do resultado da comparação 1920.

Referindo-se agora à figura 20, um método 2000 inclui executar uma avaliação com base em intensidade para detectar oclusão 2005 em um quadro baseado em partículas para rastrear um objeto entre imagens em uma sequência de imagens digitais. Em uma implementação, a avaliação baseada em intensidade pode ser com base em associação de dados. Se oclusão não for detectada, 2010, então uma avaliação probabilística é executada

para detectar oclusão 2015. Em uma implementação, a avaliação probabilística pode incluir o método descrito anteriormente baseado em uma superfície de correlação. Um indicador do resultado do processo de detectar oclusão é armazenado opcionalmente 2020.

Referindo-se agora à figura 21, um método 2100 inclui selecionar um subconjunto de partículas disponíveis 2105 para rastrear um objeto entre imagens em uma sequência de imagens digitais. Em uma implementação, tal como mostrado na figura 21, a(s) partícula(s) tendo uma probabilidade mais alta é(são) selecionada(s). Um estado é estimado com base no subconjunto de partículas selecionado 2110.

Referindo-se agora à figura 22, um método 2200 inclui determinar que uma posição estimada para um objeto em um quadro particular em uma sequência de imagens digitais está ocluída 2205. Uma trajetória é estimada para o objeto 2210. A posição estimada é mudada com base na trajetória estimada 2215.

Referindo-se agora à figura 23, um método 2300 inclui determinar uma trajetória de objeto 2310. O objeto pode estar, por exemplo, em uma imagem particular em uma sequência de imagens digitais, e a trajetória pode ser baseada em uma ou mais localizações anteriores do objeto em uma ou mais imagens anteriores na sequência. O método 2300 inclui determinar um peso de partícula baseado na distância da partícula à trajetória 2320. A partícula pode ser usada, por exemplo, em um quadro baseado em partículas para rastrear o objeto. O método 2300 inclui determinar uma localização de objeto com base no peso de partícula determinado 2330. A localização pode ser determinada usando, por exemplo, um quadro baseado em partículas.

Implementações podem produzir, por exemplo, uma estimativa de localização para um objeto. Uma estimativa como esta pode ser usada na codificação de uma imagem que inclui o objeto, por exemplo. A codificação pode usar, por exemplo, MPEG-1, MPEG-2, MPEG-4, H.264, ou outras técnicas de codificação. A estimativa, ou a codificação, pode ser fornecida, por exemplo, em um sinal ou em uma mídia legível por processador. Implementações também podem ser adaptadas para aplicações que não de rastreamento de objeto, ou aplicações não de vídeo. Por exemplo, um estado pode representar um recurso a não ser uma localização de objeto, e não necessita mesmo se relacionar a um objeto.

As implementações descritas neste documento podem ser implementadas, por exemplo, em um método ou processo, um aparelho ou em um software. Mesmo se discutido somente no contexto de uma única forma de implementação (por exemplo, discutido somente como um método), a implementação de recursos discutida também pode ser implementada em outras formas (por exemplo, um aparelho ou programa). Um aparelho pode ser implementado, por exemplo, em hardware apropriado, software e firmware. Os métodos podem ser implementados, por exemplo, em um aparelho tal como, por exemplo, um processador, o qual se refere a dispositivos de processamento em geral, incluindo, por exemplo,

um computador, um microprocessador, um circuito integrado, ou um dispositivo lógico programável. Dispositivos de processamento também incluem dispositivos de comunicação, tais como, por exemplo, computadores, telefones celulares, assistentes digitais portáteis/pessoais (“PDAs”), e outros dispositivos que facilitam comunicação de informação entre usuários finais.

Implementações dos vários processos e recursos descritos neste documento podem ser incorporadas em uma variedade de diferentes equipamentos ou aplicações, particularmente, por exemplo, equipamentos ou aplicações associados com codificação e decodificação de dados. Exemplos de equipamento incluem codificadores de vídeo, decodificadores de vídeo, codificadores/decodificadores de vídeo, servidores de rede, aparelhos conversores de sinais, computadores portáteis, computadores pessoais, telefones celulares, PDAs e outros dispositivos de comunicação. Tal como deve estar claro, o equipamento pode ser móvel e mesmo instalado em um veículo móvel.

Adicionalmente, os métodos podem ser implementados por meio de instruções sendo executadas por um processador, e tais instruções podem ser armazenadas em uma mídia legível por processador tal como, por exemplo, um circuito integrado, um portador de software ou outro dispositivo de armazenamento tal como, por exemplo, um disco rígido, um disco compacto, uma memória de acesso aleatório (“RAM”), ou uma memória somente de leitura (“ROM”). As instruções podem formar um programa de aplicação incorporado de modo tangível em uma mídia legível por processador. Instruções podem ser, por exemplo, em hardware, firmware, software, ou em uma combinação. Instruções podem ser encontradas, por exemplo, em um sistema de operação, uma aplicação separada, ou em uma combinação dos dois. Portanto, um processador pode ser caracterizado, por exemplo, tanto como um dispositivo configurado para executar um processo quanto um dispositivo que inclui uma mídia legível por computador tendo instruções para executar um processo.

Tal como deve estar evidente para os versados na técnica, implementações também podem produzir um sinal formatado para carregar informação que pode ser, por exemplo, armazenada ou transmitida. A informação pode incluir, por exemplo, instruções para executar um método, ou dados produzidos por uma das implementações descritas. Um sinal como este pode ser formatado, por exemplo, como uma onda eletromagnética (por exemplo, usando uma parte de radiofrequência de espectro) ou como um sinal de banda base. A formatação pode incluir, por exemplo, codificar um fluxo de dados e modular uma portadora com o fluxo de dados codificados. A informação que o sinal leva pode ser, por exemplo, informação analógica ou digital. O sinal pode ser transmitido através de uma variedade de diferentes vínculos com fio ou sem fio, tal como é conhecido.

Diversas implementações foram descritas. Apesar disso, será entendido que várias modificações podem ser feitas. Por exemplo, elementos de diferentes implementações po-

dem ser combinados, suplementados, modificados, ou removidos para produzir outras implementações. Adicionalmente, uma pessoa de conhecimento comum entenderá que outras estruturas e processos podem ser substituídos por esses revelados e as implementações resultantes executarão pelo menos substancialmente a(s) mesma(s) função(s), pelo menos substancialmente no(s) mesmo(s) modo(s), para alcançar pelo menos substancialmente o(s) mesmo(s) resultado(s) tal como as implementações reveladas. Desta maneira, estas e outras implementações são consideradas por este pedido e estão dentro do escopo das reivindicações a seguir.

REIVINDICAÇÕES

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

a determinação de uma trajetória de um objeto em uma imagem particular em uma sequência de imagens digitais, a trajetória sendo baseada em uma ou mais localizações anteriores do objeto em uma ou mais imagens anteriores na sequência;

a determinação de um peso, para uma partícula em uma estrutura de trabalho com base em partícula para o rastreamento do objeto; com base na distância da trajetória para a partícula; e

a determinação da estimativa de localização para o objeto utilizando a estrutura de trabalho com base em partícula, a estimativa de localização sendo baseada no peso da partícula determinada.

2. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de compreender adicionalmente:

a determinação de uma parte de objeto da imagem particular que inclui a localização estimada do objeto;

a determinação de uma parte de não objeto da imagem em particular que é separada da parte de objeto; e

a codificação da parte de objeto e da parte de não objeto, de forma que a parte de objeto seja codificada com maior redundância de codificação do que a parte de não objeto.

3. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de o objeto ser pequeno o suficiente de forma que uma ou mais das localizações anteriores do objeto dentro de uma imagem não se sobreponham uma à outra.

4. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de a determinação do peso para a partícula na estrutura de trabalho com base em particular também ser baseada em um ou mais dentre:

uma extrapolação linear de um ou mais locais anteriores do objeto em uma ou mais imagens anteriores na sequência; e

uma comparação de um gabarito e uma parte da imagem particular correspondendo a uma posição da partícula.

5. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de a trajetória determinada ser não linear.

6. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de uma ou mais localizações anteriores do objeto, que são utilizadas na determinação da trajetória, ser localizações não obstruídas.

7. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de a trajetória ser determinada pelo menos em parte em uma ocorrência ponderada da oclusão do objeto em imagens anteriores na sequência.

8. Método, de acordo com a reivindicação 1, **CHARACTERIZADO** pelo fato de uma localização de objeto em um estado de oclusão em uma das imagens anteriores na sequência se desconsiderada na formação da trajetória.

9. Método, de acordo com a reivindicação 1, **CHARACTERIZADO** pelo fato de uma confiabilidade de uma trajetória estimada ser ponderada pela informação referente à oclusão de um objeto em uma ou mais imagens anteriores.

10. Método, de acordo com a reivindicação 1, **CHARACTERIZADO** pelo fato de o objeto possuir um tamanho de menos de cerca de 30 pixels.

11. Método, de acordo com a reivindicação 1, **CHARACTERIZADO** pelo fato de a estrutura de trabalho com base em partícula compreender um filtro de partículas.

12. Método, de acordo com a reivindicação 1, **CHARACTERIZADO** pelo fato de o método ser implementado em um codificador.

13. Aparelho, **CHARACTERIZADO** pelo fato de compreender:

um dispositivo de armazenamento para o armazenamento de dados com relação a uma sequência de imagens digitais; e

um processador para (1) determinar uma trajetória de um objeto em uma imagem particular em uma sequência de imagens digitais, a trajetória sendo baseada em uma ou mais localizações anteriores do objeto em uma ou mais imagens anteriores na sequência; (2) determinar um peso, para uma partícula em uma estrutura de trabalho com base em partícula para rastrear o objeto, com base na distância da trajetória para a partícula; e (3) determinar uma estimativa de localização para o objeto utilizando a estrutura de trabalho com base em partícula, a estimativa de localização sendo baseada no peso determinado da partícula.

14. Aparelho, de acordo com a reivindicação 13, **CHARACTERIZADO** pelo fato de compreender adicionalmente um codificador que inclui o dispositivo de armazenamento e o processado.

15. Meio legível por processador, **CHARACTERIZADO** pelo fato de ser armazenado no mesmo uma pluralidade de instruções para a realização:

a determinação de uma trajetória de um objeto em uma imagem em particular em uma sequência de imagens digitais, a trajetória sendo baseada em uma ou mais localizações anteriores do objeto em uma ou mais imagens anteriores na sequência;

a determinação de um peso, para uma partícula em uma estrutura de trabalho baseada em particular para o rastreamento do objeto, com base na distância da trajetória para a partícula; e

a determinação de uma estimativa de localização para o objeto utilizando a estrutura de trabalho baseada em partícula, a estimativa de localização sendo baseada no peso de partícula determinado.

16. Aparelho, **CHARACTERIZADO** pelo fato de compreender:

meios para o armazenamento de dados com relação a uma sequência de imagens digitais;

5 meios para (1) determinar uma trajetória de um objeto em uma imagem particular em uma sequência de imagens digitais, a trajetória sendo baseada em uma ou mais localizações anteriores do objeto em uma ou mais imagens anteriores na sequência; (2) determinar um peso, para uma partícula em uma estrutura de trabalho com base em partícula para rastrear o objeto, com base na distância da trajetória para a partícula; e (3) determinar uma
10 estimativa de localização para o objeto utilizando a estrutura de trabalho com base em partícula, a estimativa de localização sendo baseada no peso de partícula determinado.

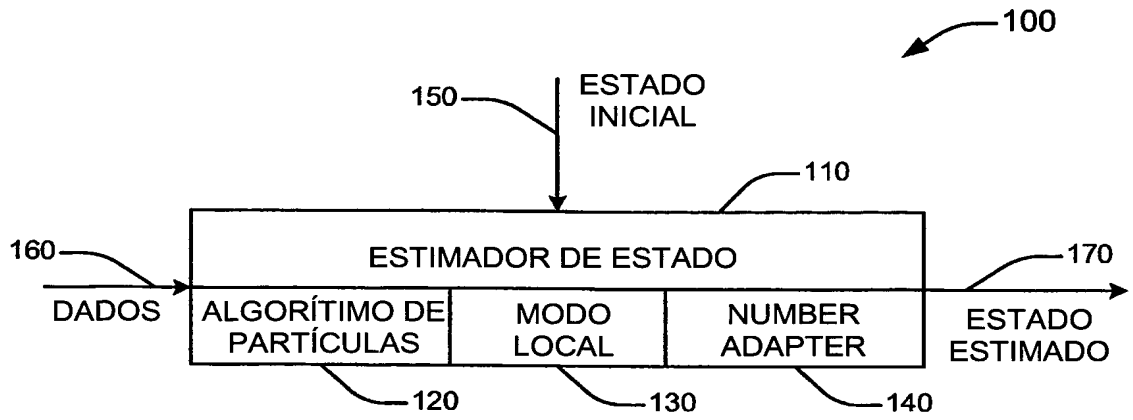


FIG. 1

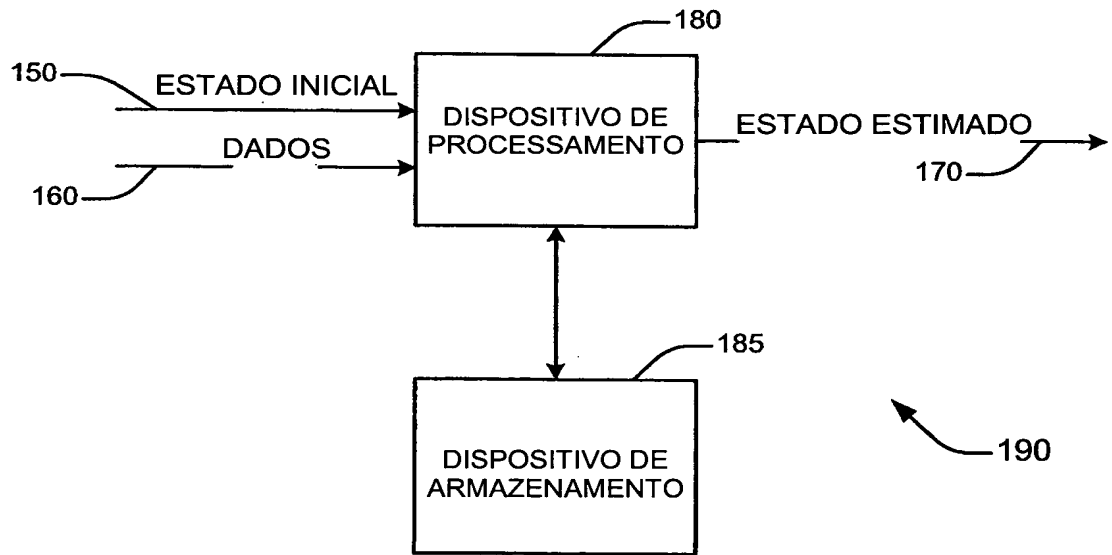


FIG. 2

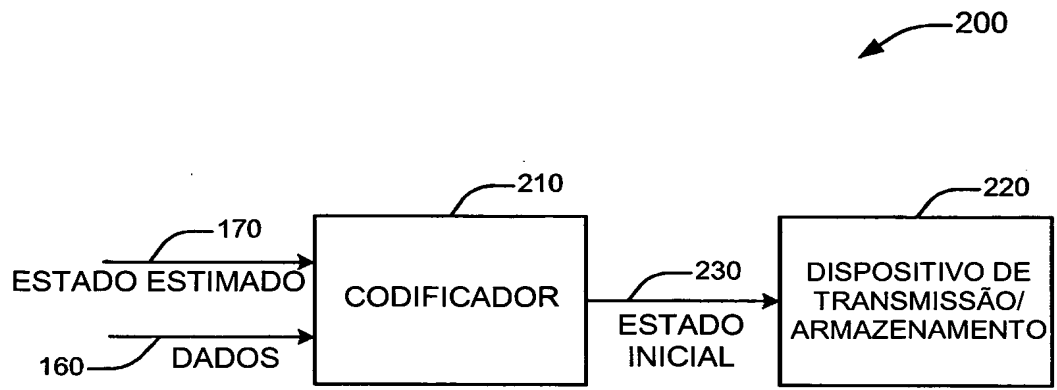


FIG. 3

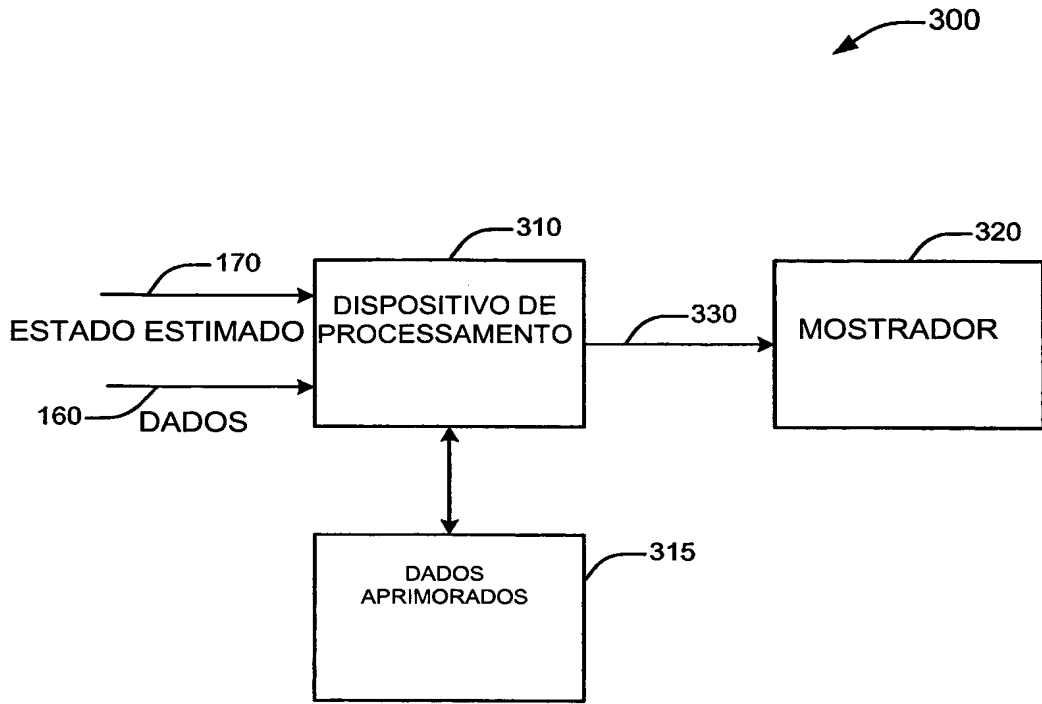


FIG. 4

DISPOSITIVO DE
ARMAZENAMENTO

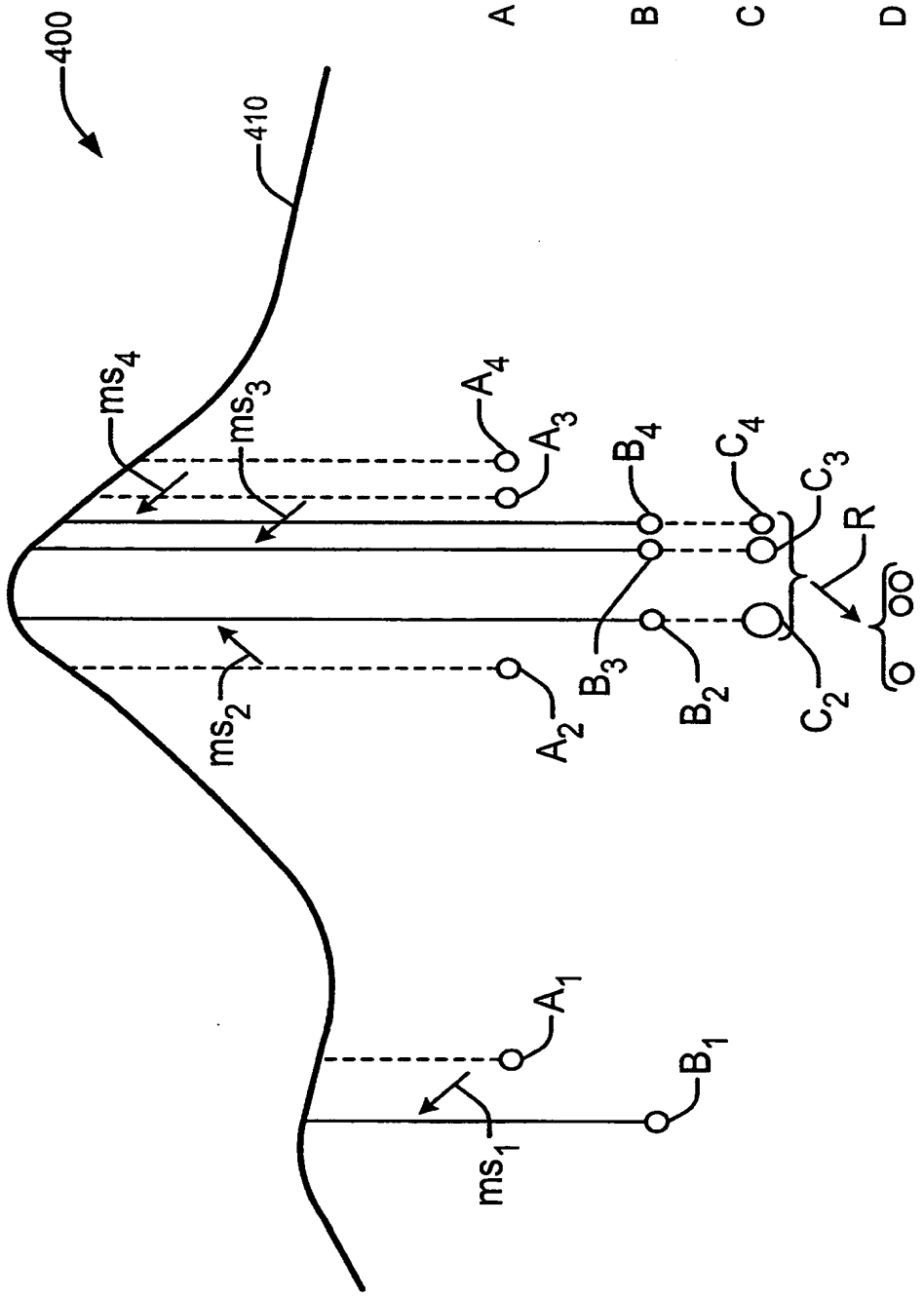


FIG. 5

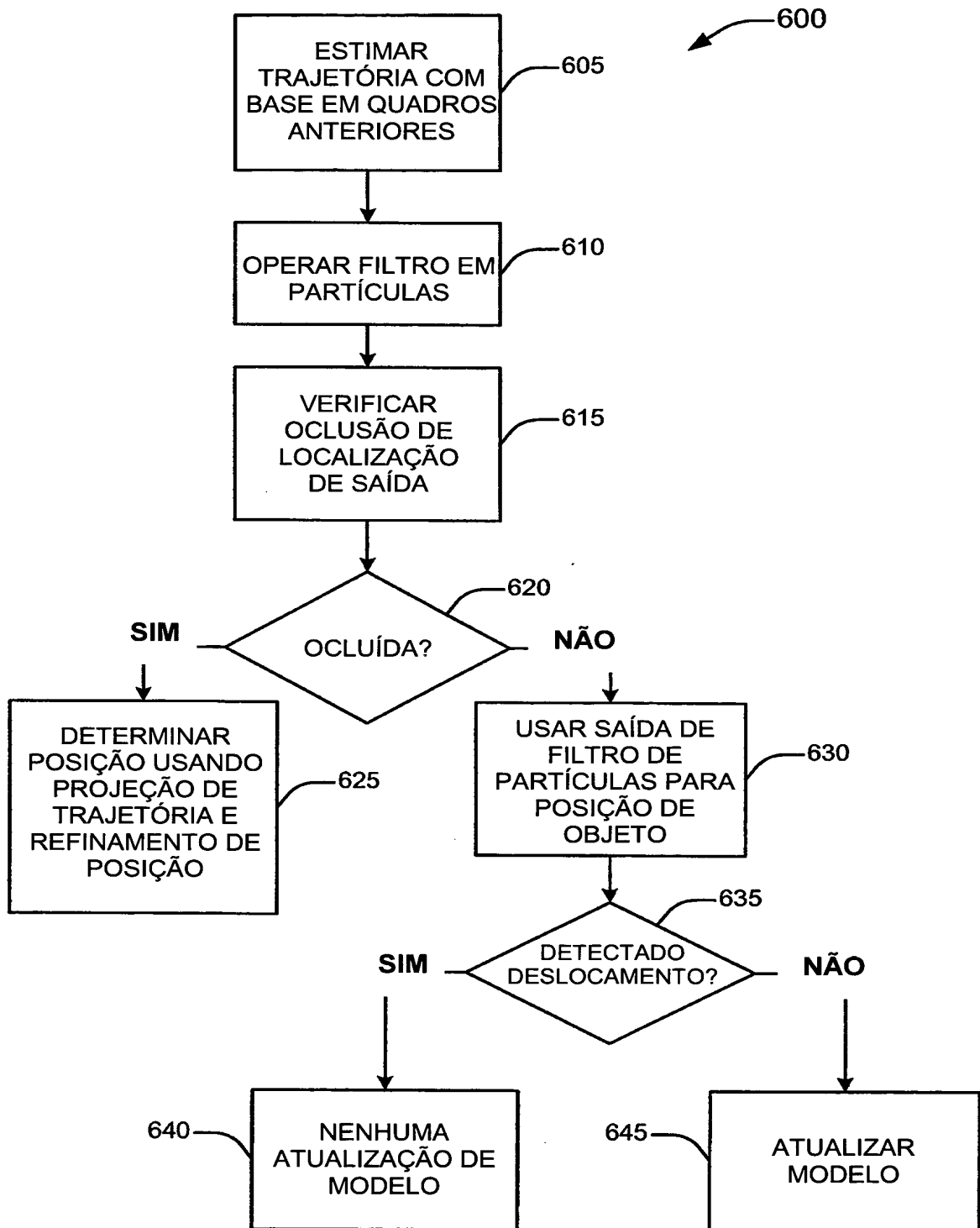


FIG. 6

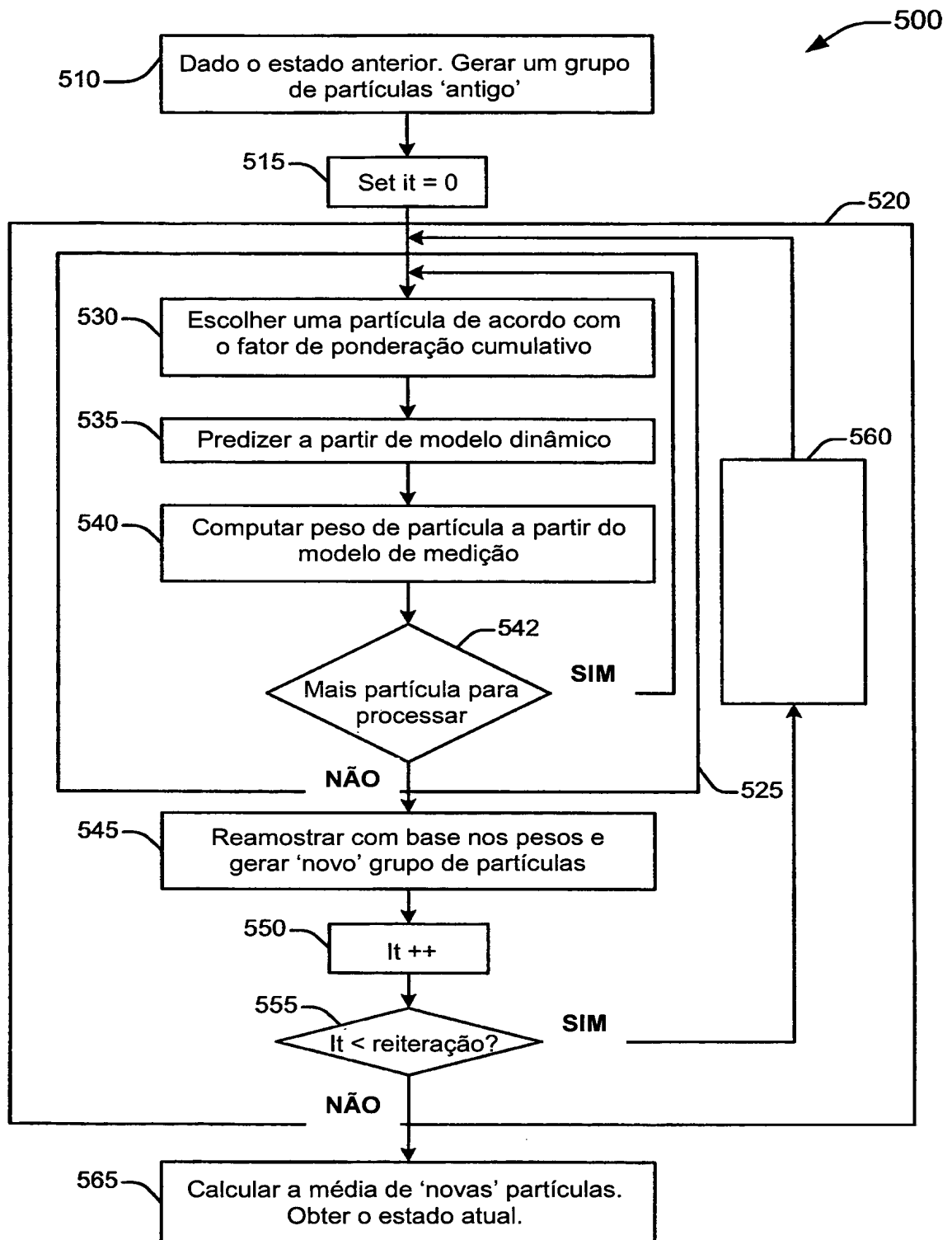


FIG. 7

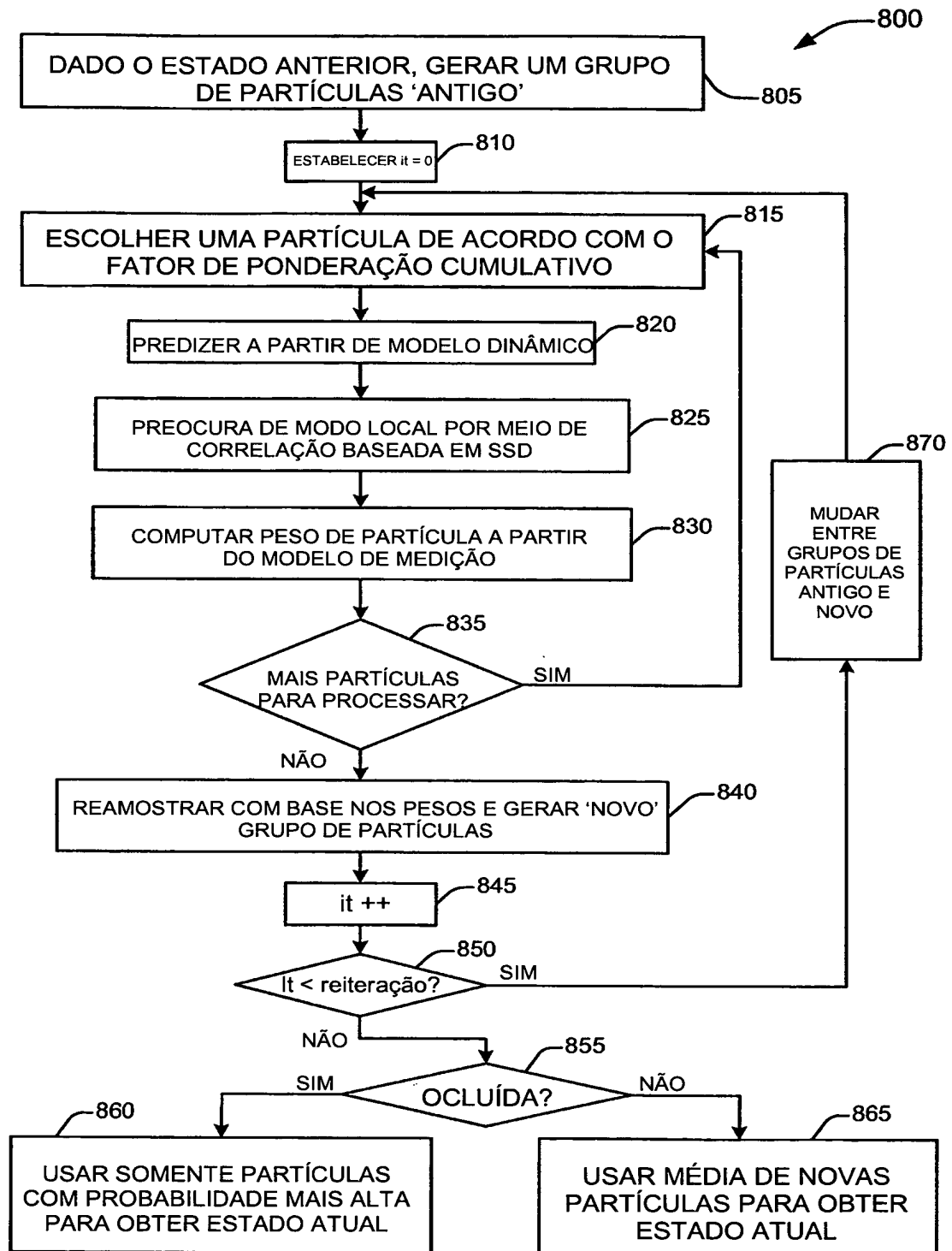


FIG. 8

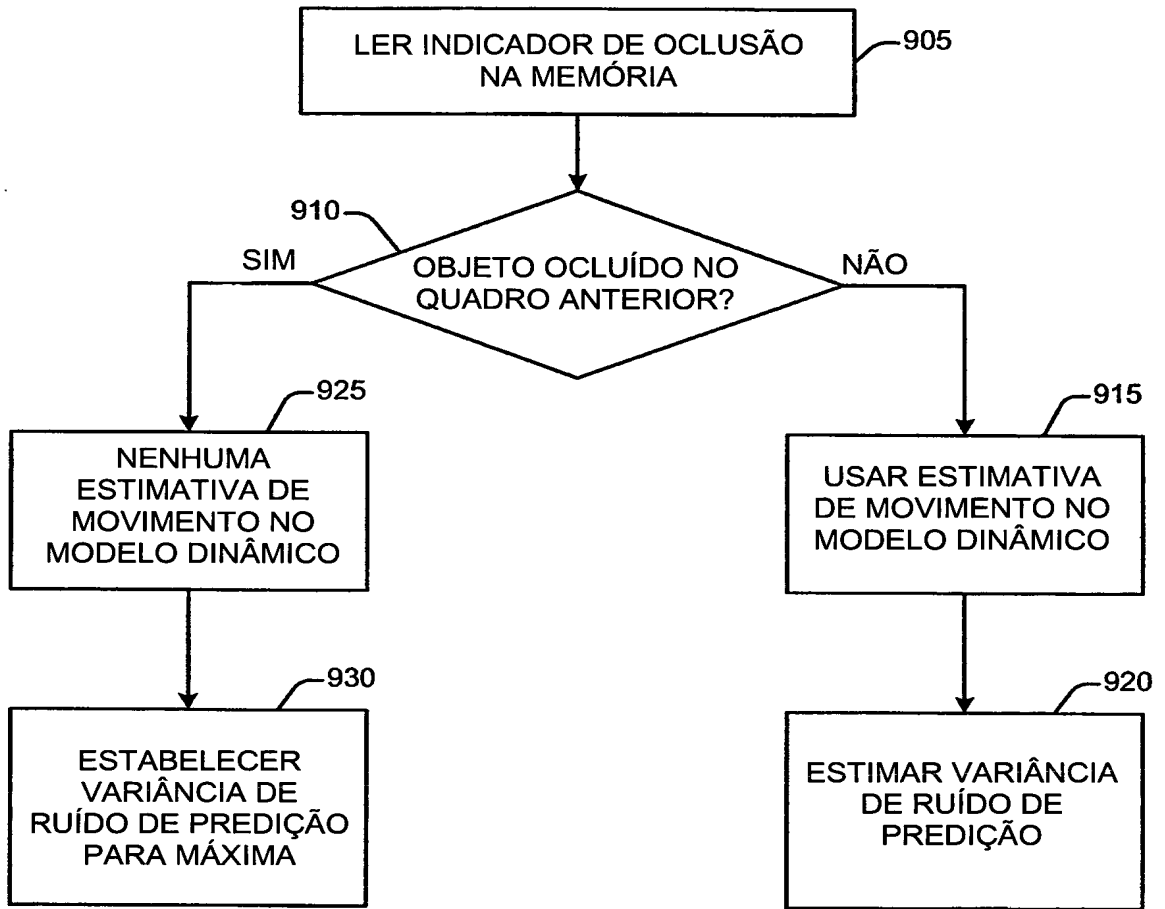


FIG. 9

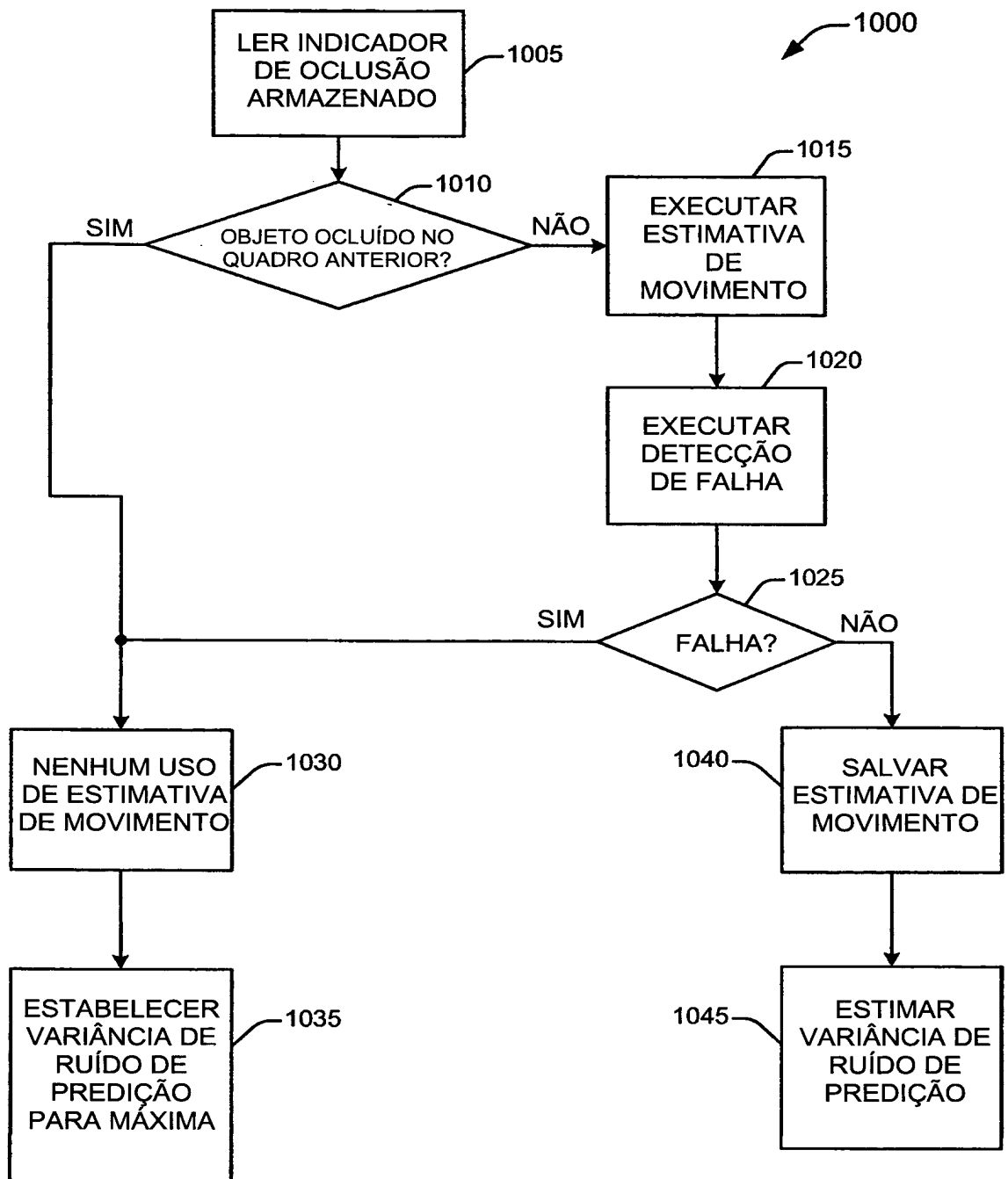


FIG. 10

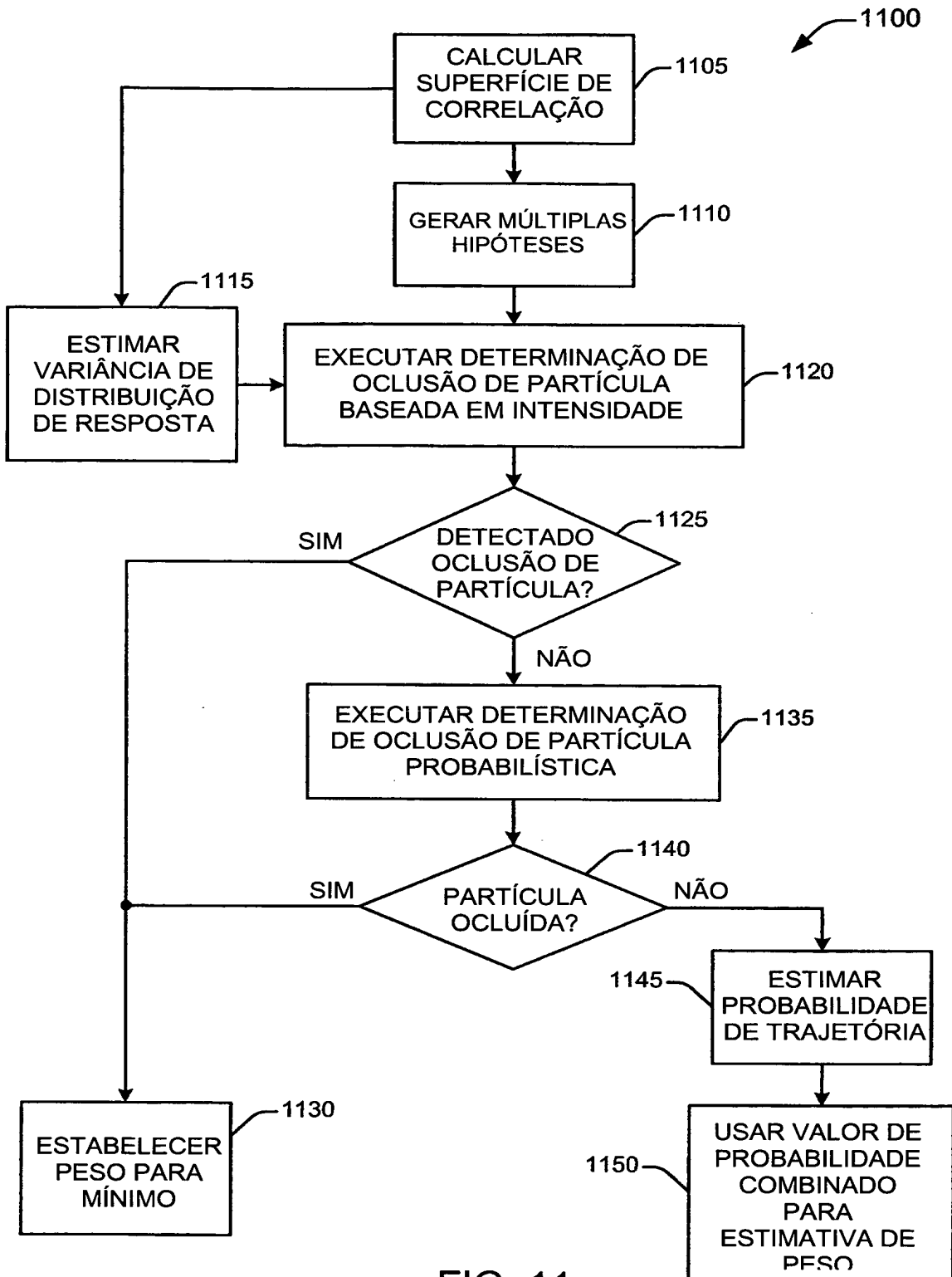


FIG. 11

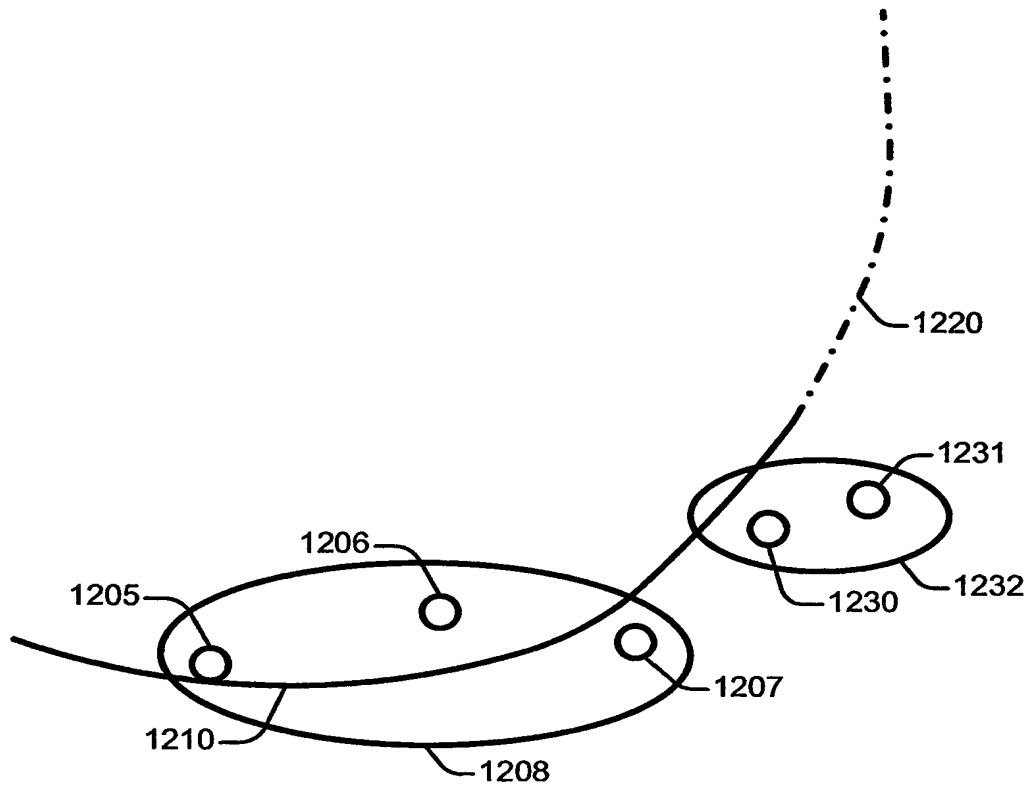


FIG. 12

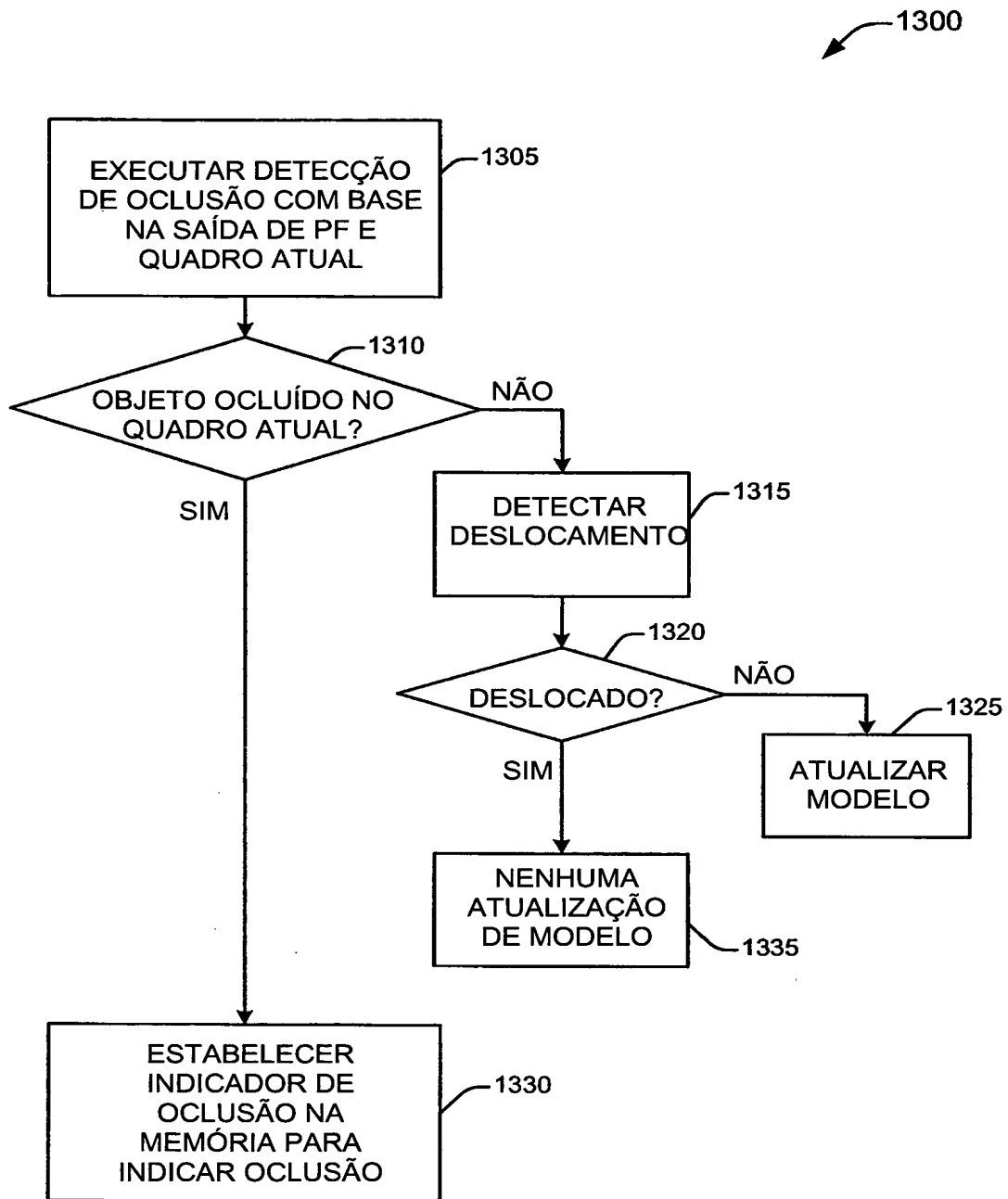


FIG. 13

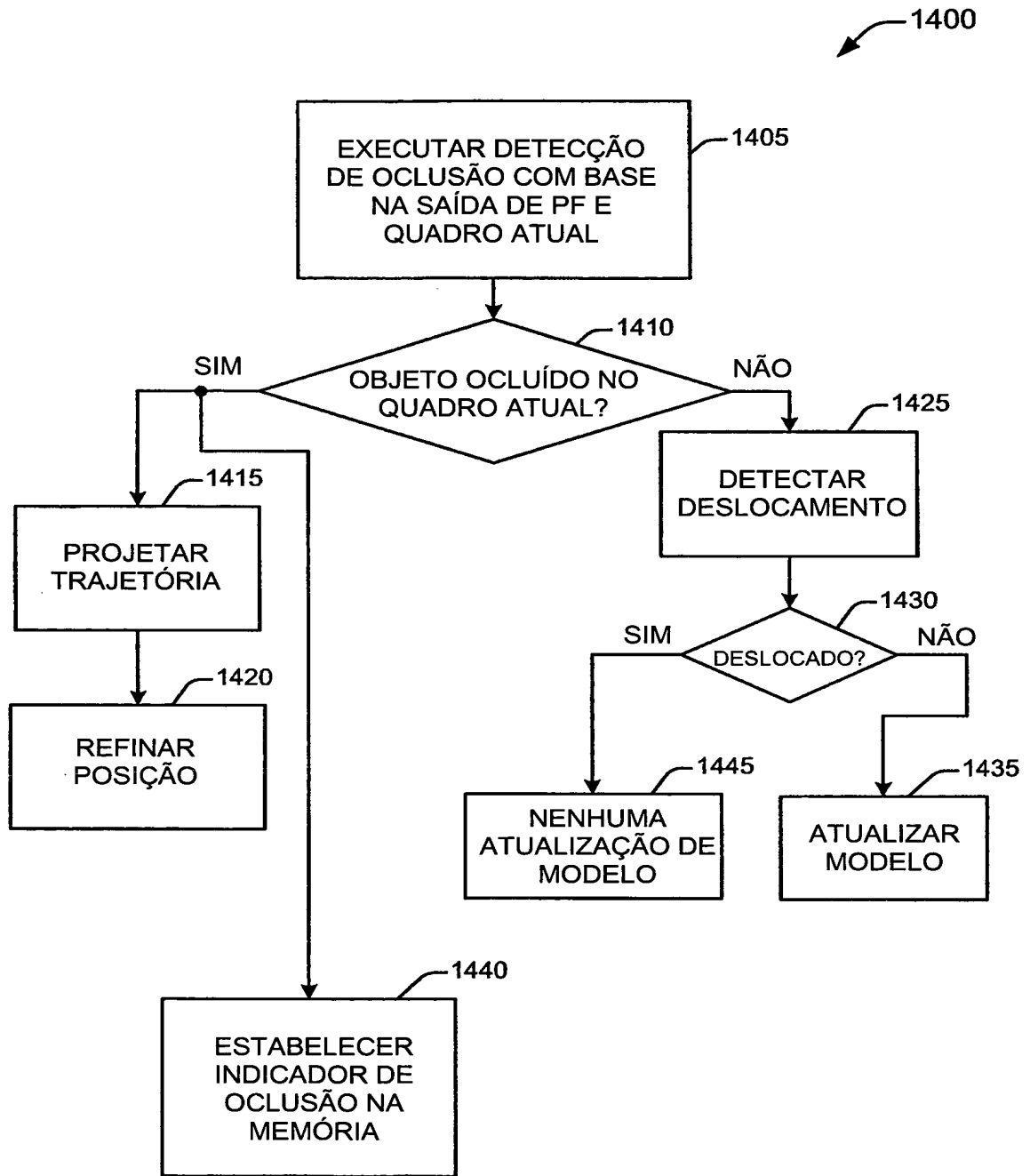


FIG. 14

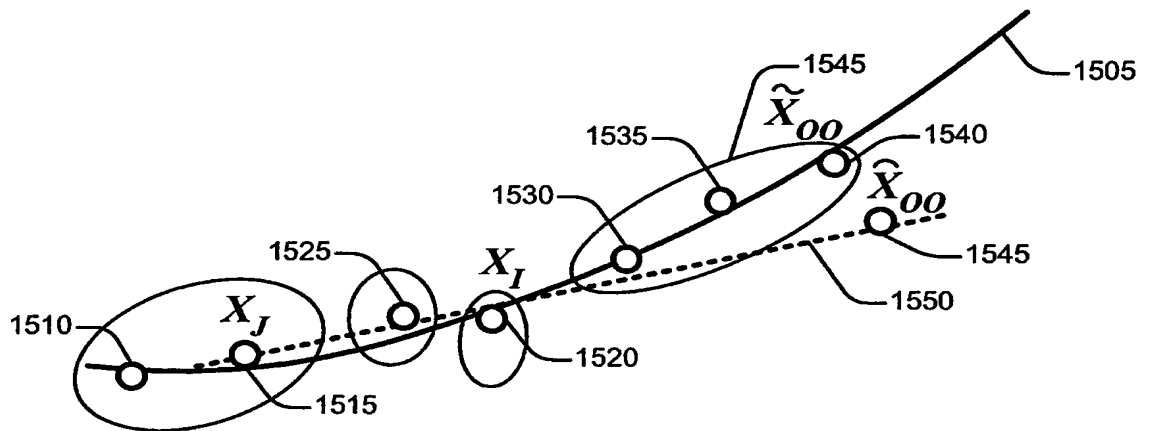


FIG. 15

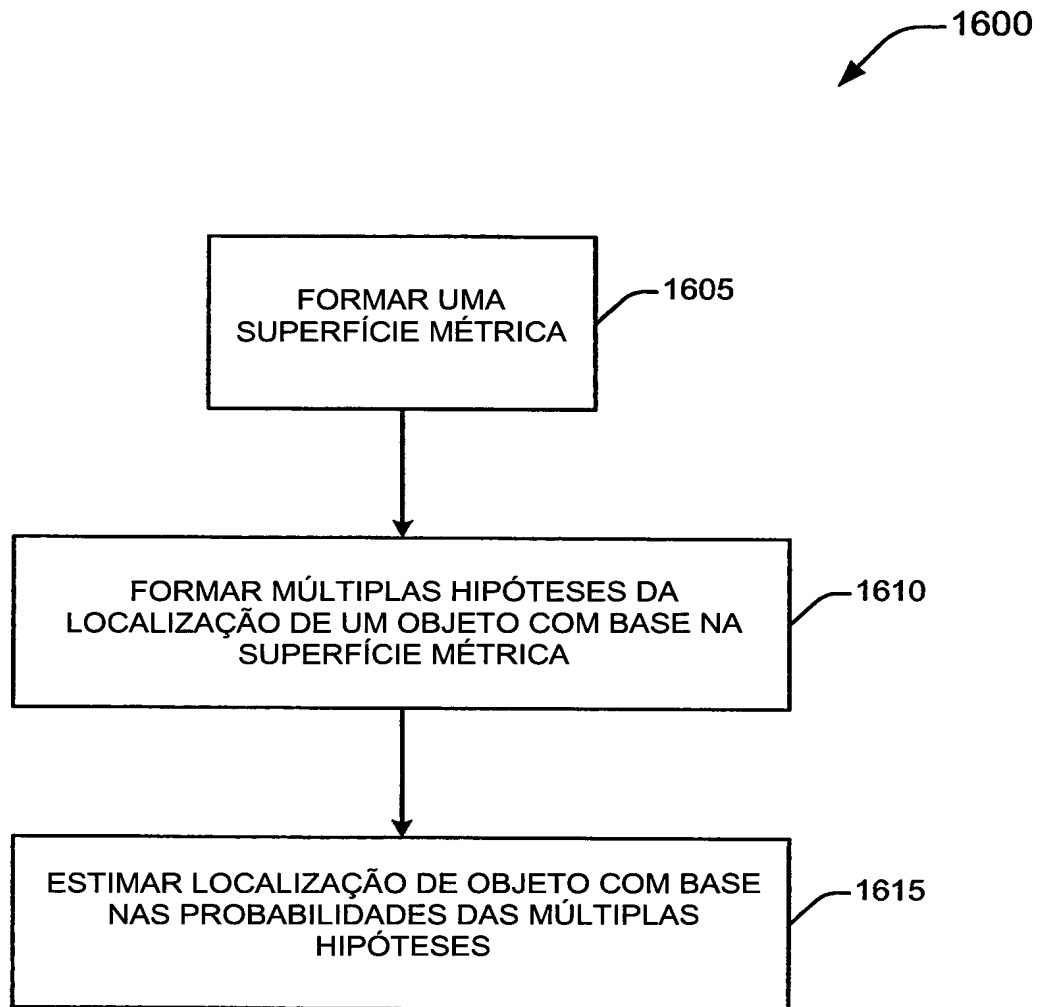


FIG. 16

1700

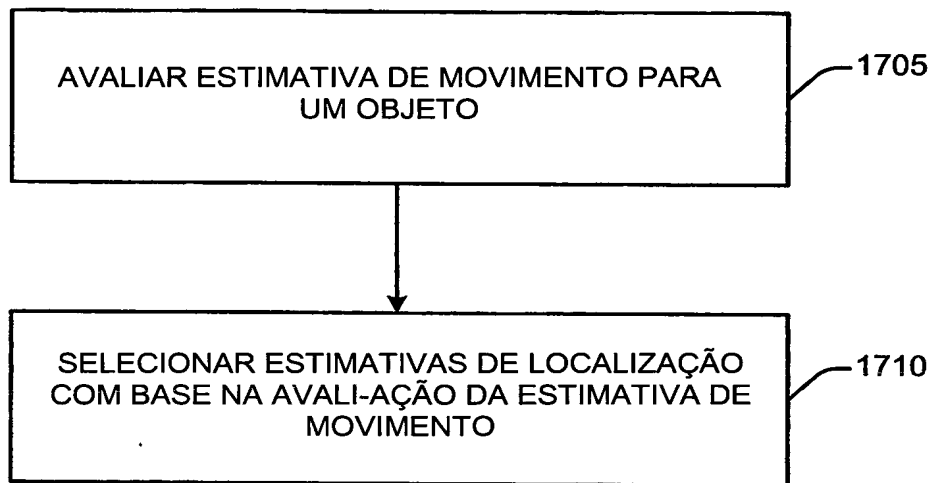
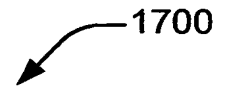


FIG. 17

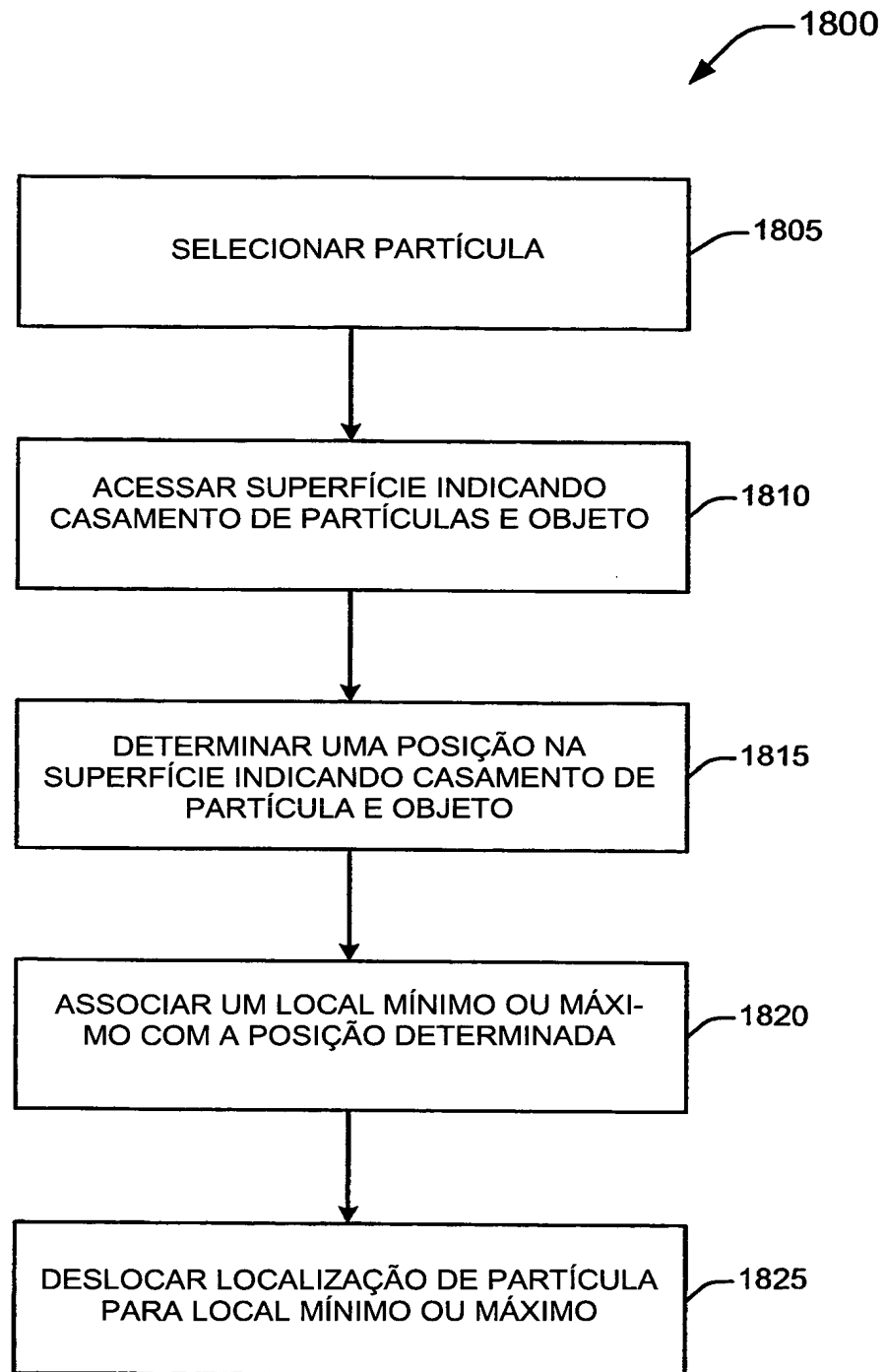


FIG. 18

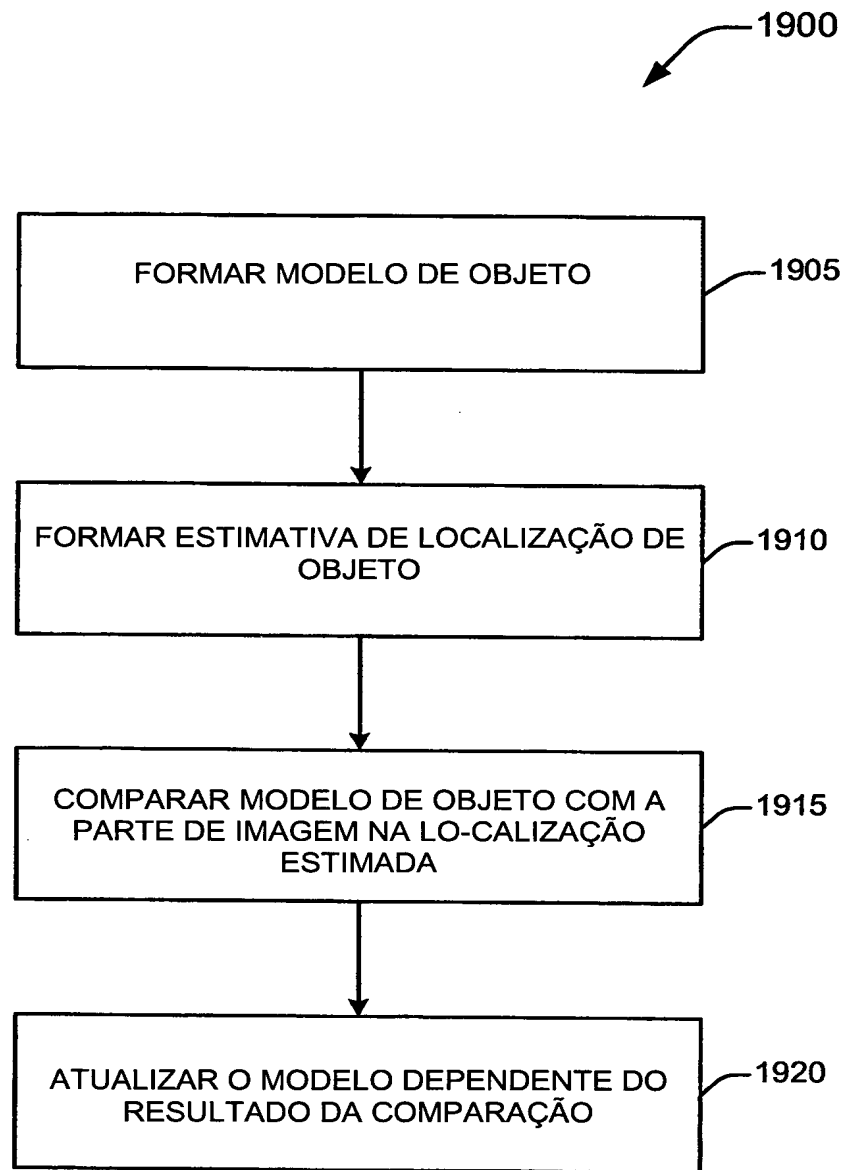


FIG. 19

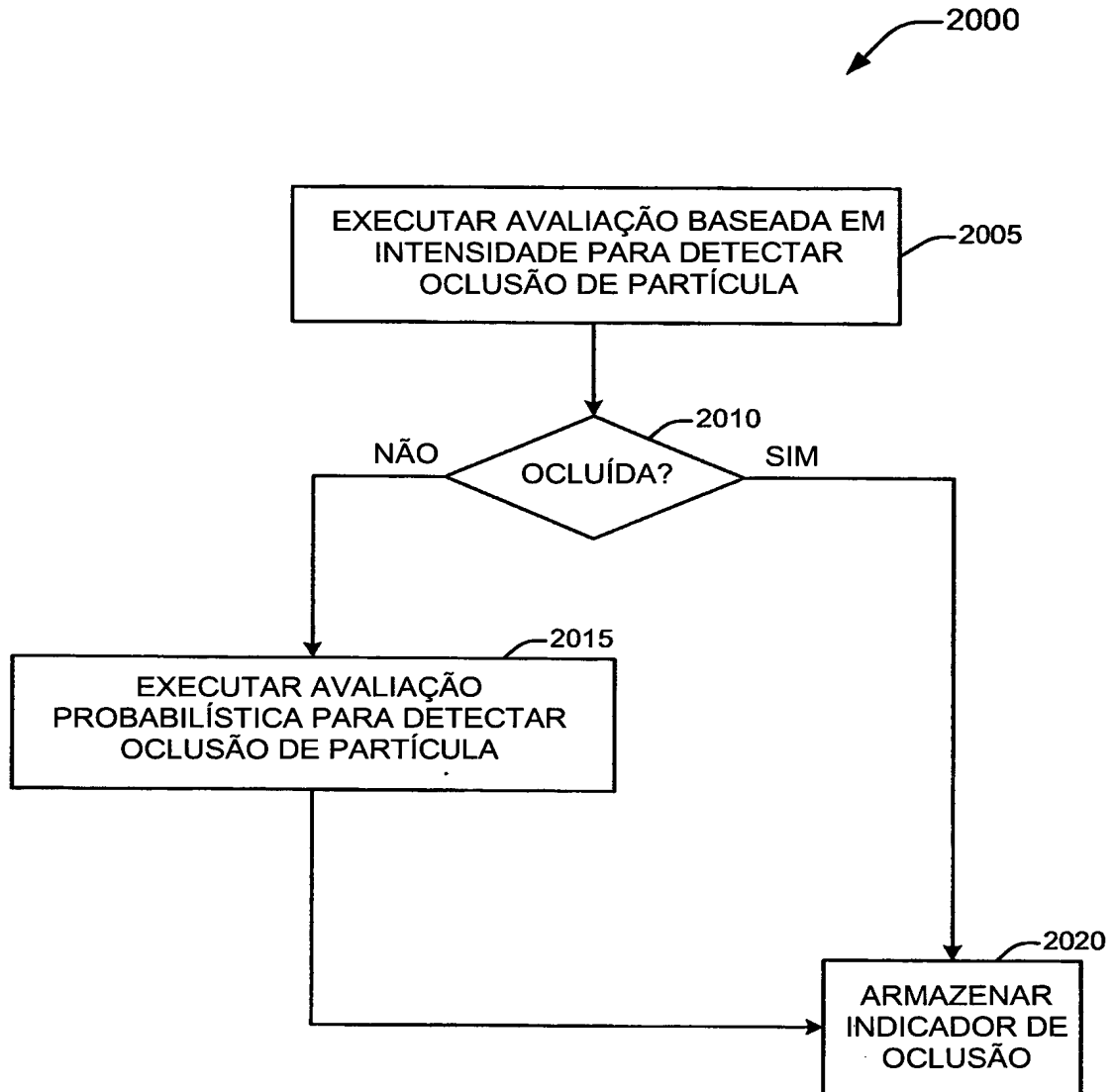


FIG. 20

2100

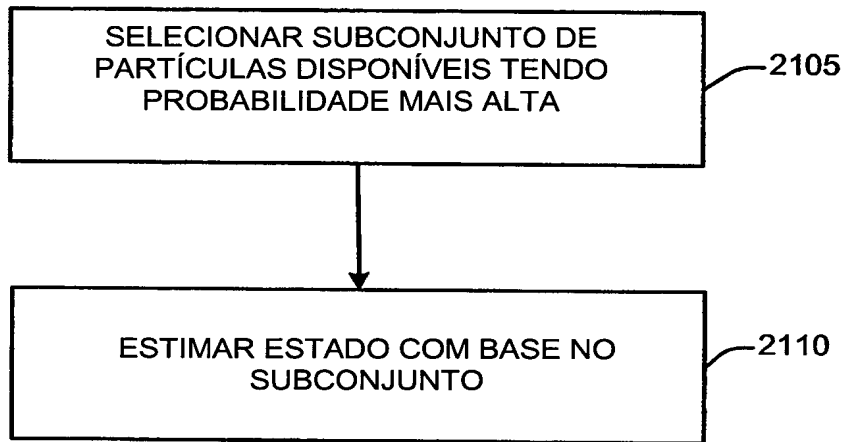
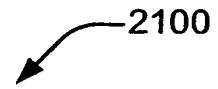


FIG. 21

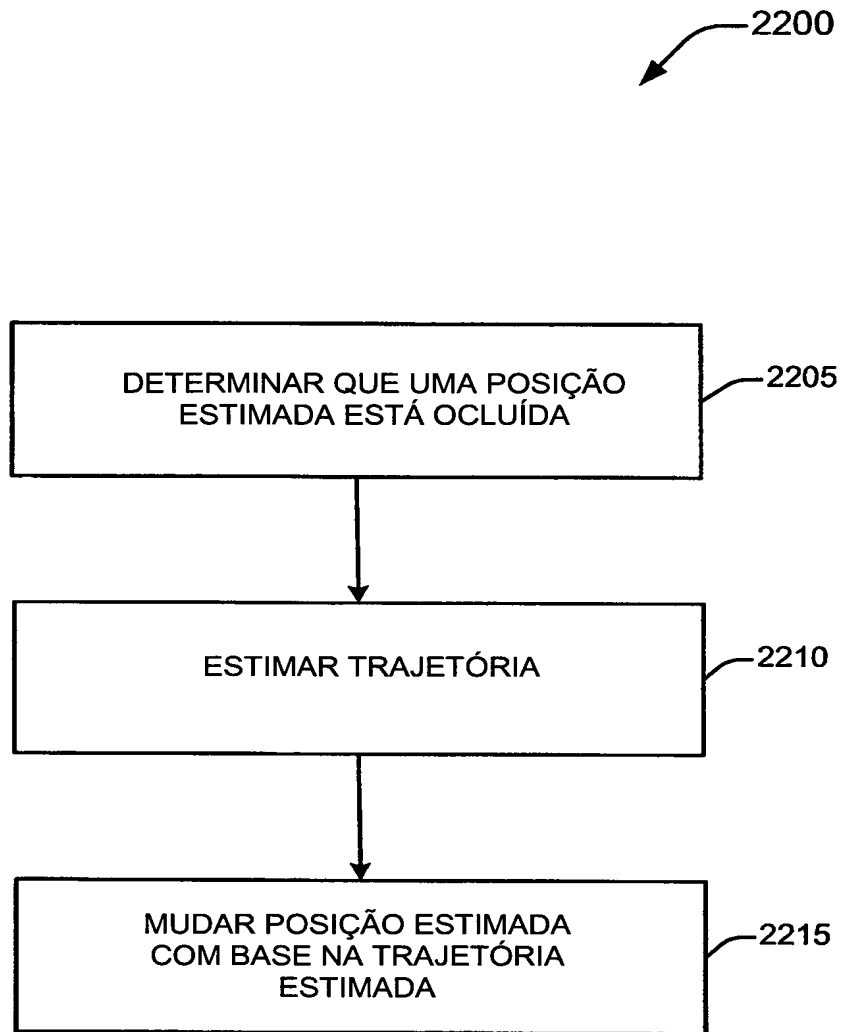


FIG. 22

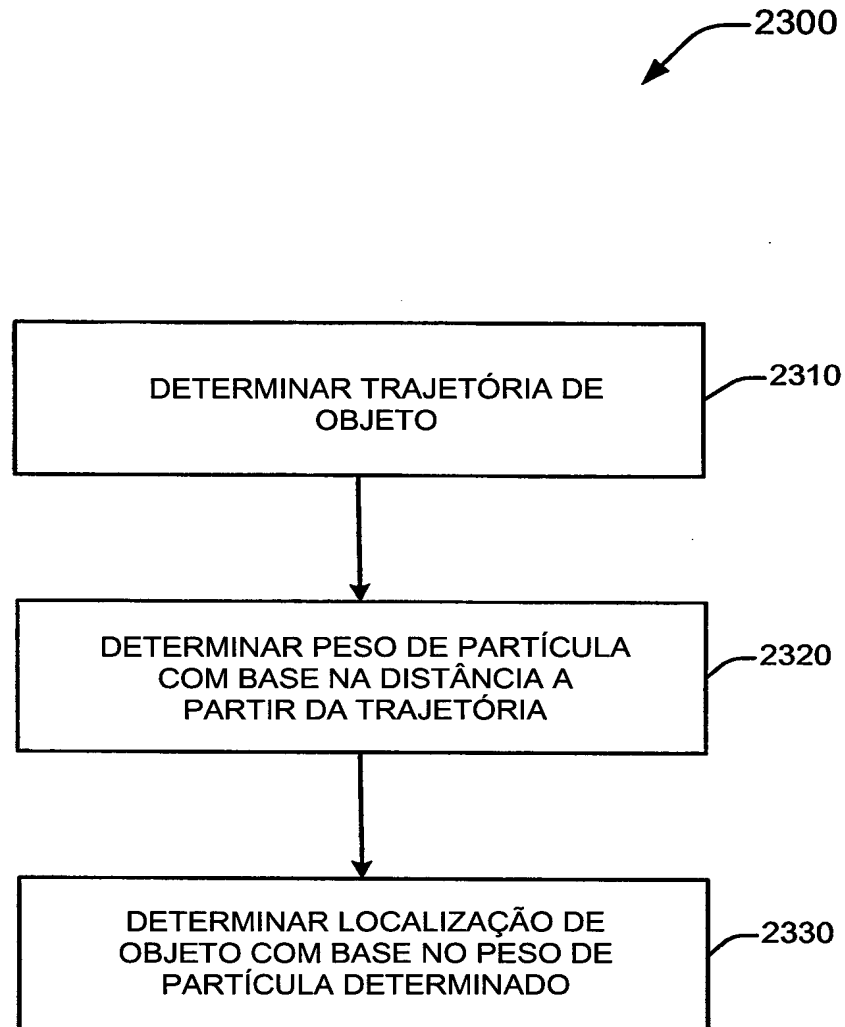


FIG. 23

RESUMO

“ESTIMATIVA DE UMA LOCALIZAÇÃO DE UM OBJETO EM UMA IMAGEM”

Uma implementação fornece um método para determinar uma trajetória de um objeto em uma imagem em particular em uma sequência de imagens digitais, a trajetória sendo baseada em um ou mais locais anteriores do objeto em uma ou mais imagens anteriores na sequência. Um peso é determinado, para uma partícula em uma estrutura de trabalho com base em partícula para o rastreamento do objeto, com base na distância da trajetória para a partícula. Uma estimativa de localização é determinada para o objeto utilizando a estrutura de trabalho com base em particular, a estimativa de localização sendo baseada no peso determinado da partícula.