



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

**(11) BR 112014019731-8 B1**



**(22) Data do Depósito: 08/02/2013**

**(45) Data de Concessão: 14/06/2022**

**(54) Título:** MÉTODO DE LOCALIZAÇÃO DE UM VEÍCULO QUE HOSPEDA UM SENSOR, SISTEMA ARRANJADO PARA LOCALIZAR UM VEÍCULO, VEÍCULO TENDO UM SENSOR MONTADO NO MESMO, SERVIDOR ARRANJADO PARA RECEBER UM PEDIDO DE REDE, E, MEIO LEGÍVEL POR MÁQUINA

**(51) Int.Cl.:** G01S 19/48; G06K 9/00; G01C 21/26; G01C 21/32.

**(52) CPC:** G01S 19/48; G06K 9/00664; G01C 21/26; G01C 21/32.

**(30) Prioridade Unionista:** 10/02/2012 GB 1202344.6.

**(73) Titular(es):** OXFORD UNIVERSITY INNOVATION LIMITED.

**(72) Inventor(es):** WINSTON CHURCHILL; PAUL MICHAEL NEWMAN.

**(86) Pedido PCT:** PCT GB2013050299 de 08/02/2013

**(87) Publicação PCT:** WO 2013/117940 de 15/08/2013

**(85) Data do Início da Fase Nacional:** 08/08/2014

**(57) Resumo:** MÉTODO DE LOCALIZAÇÃO DE UM VEÍCULO QUE HOSPEDA UM SENSOR, SISTEMA ARRANJADO PARA LOCALIZAR UM VEÍCULO, VEÍCULO TENDO UM SENSOR MONTADO NO MESMO, SERVIDOR ARRANJADO PARA RECEBER UM PEDIDO DE REDE, E, MEIO LEGÍVEL POR MÁQUINA. Um método de localização de um veículo que hospeda um sensor compreendendo as etapas de capturar dados a partir do sensor provendo uma cena detectada em torno do veículo em um momento atual, processar a cena detectada para extrair um conjunto de características a partir da cena detectada e para determinar, usando as características extraídas, uma posição do veículo a partir da cena detectada em relação à posição do sensor em um momento precedente, comparar o conjunto extraído de características a partir da cena detectada com uma ou mais experiências armazenadas, em que cada experiência armazenada compreende uma pluralidade de conjuntos de características, onde cada conjunto de características foi determinado a partir de uma cena detectada previamente, para determinar se a cena detectada pode ser reconhecida dentro de qualquer das experiências armazenadas; e se o conjunto extraído de características da cena detectada é reconhecido dentro de qualquer uma das experiências armazenadas, então esta experiência armazenada é usada para produzir uma estimativa da posição do (...).

“MÉTODO DE LOCALIZAÇÃO DE UM VEÍCULO QUE HOSPEDA UM SENSOR, SISTEMA ARRANJADO PARA LOCALIZAR UM VEÍCULO, VEÍCULO TENDO UM SENSOR MONTADO NO MESMO, SERVIDOR ARRANJADO PARA RECEBER UM PEDIDO DE REDE, E, MEIO LEGÍVEL POR MÁQUINA”

[0001] Modalidades desta invenção se referem a métodos de localização de um veículo, ou um sensor arranjado para ser usado em um veículo, e aparelho relacionado.

[0002] À medida em que localização (isto é, a determinação da posição) e algoritmos de mapeamento são aperfeiçoados, o problema de navegação verdadeiramente a longo prazo está se tornando mais urgente. Um problema comum na robótica (em algumas modalidades, esta pode ser veículo), a navegação é necessária em ambientes modificáveis e é um grande obstáculo a concretização de autonomia ao longo da vida. As propostas de mapeamento tradicionais frequentemente criam um mapa inicial, que é tipicamente uma única proposta monolítica (como um resultado de processos, tais como fusão ou formação de média), uma vez quando na visita inicial e a esperança que esta será suficientemente próxima em aparência para ser útil nas visitas subsequentes. Recentemente, foram feitas algumas tentativas de combater esses problemas. Para obter autonomia em longo prazo, os sistemas robóticos devem ser capazes de funcionar em ambientes modificáveis, o que é visto na técnica como um grande desafio. Modificação pode provir de muitas fontes: súbita alteração estrutural, condição de iluminação, momento do dia, tempo, mudança sazonal e similar. Para ilustrar, considere o problema de estimativa de ego-movimento com uma câmera montada em um robô que opera ao ar livre ou exterior. Esta é uma rica área de fonte de pesquisa e imediatamente chegamos a um sistema de navegação visual (Simultâneos Localização e Mapeamento - SLAM) que pode mapear e localizar tudo de uma vez. Todavia, quando o local é revisitado, ele pode ter se alterado

drasticamente e, por exemplo, ele pode estar com neve. Quando as estações mudam, do verão para o outono, a aparência de uma cena pode se alterar gradualmente. A técnica anterior sugere que é necessário criar um mapa monolítico contendo informação que permita a localização dentro do ambiente modificável.

[0003] Konolige e Bowman ([1] K. Konolige e J. Bowman, "Towards lifelong visual maps", em IROS, 2009, págs. 1156-1163) desenvolveram seu sistema de mapas baseado em vistas ([2] K. Konolige, J. Bowman, J. D. Chen, P. Mihelich, M. Calonder, V. Lepetit, e P. Fua, "View-based maps", *International Journal of Robotics Research* (IJRR), vol. 29, no. 10, 2010) para adaptação à alteração em um ambiente interno. O sistema cria um gráfico de esqueleto de quadros chaves de um sistema de VO (Odometria Visual). As vistas armazenadas em nós são então atualizadas e apagadas com base em um esquema projetado para preservar a diversidade de vistas enquanto limita o número máximo de vistas. Seu mapa de esqueleto está em um único quadro de referência e é incrementalmente otimizado através de Toro ([3] G. Grisetti, C. Stachniss, S. Grzonka, e W. Burgard, "A Tree Parameterization for Efficiently Computing Maximum Likelihood Maps using Gradient Descent", *Robotics Science and Systems*, 2007). Eles mostram resultados para um escritório que inclui pessoas em movimento, móveis e condições alteráveis de iluminação.

[0004] O sistema RatSLAM de Milford e Wyeth ([4] M. Milford e G. Wyeth, "Persistent navigation and mapping using a biologically inspired SLAM system", *The International Journal of Robotics Research*, 2009) usa a noção de um mapa de experiência que armazena uma coleção de experiências. Suas experiências são locais no mundo que também armazenam informação de transição para outras experiências para combater o problema de deriva de odometria. Ao longo do tempo, o mapa de experiência é relaxado para minimizar a diferença entre o local de experiência absoluto e sua informação de transição para outros. Experiências são também apagadas para manter a

densidade baixa.

[0005] Biber e Duckett ([5] P. Biber e T. Duckett, "Dynamic maps for long-term operation of mobile service robots", em *Proceedings of Robotics: Science and Systems*, Cambridge, EUA, Junho de 2005) sampleiam mapas a laser em uma série de escalas de tempo para criar um único mapa que tem estrutura tanto em longo prazo quanto em curto prazo. Isto permite a eles modelar tanto objetos dinâmicos em curto prazo quanto também mudança estrutural em longo prazo. Eles mostram precisão de localização melhorada sobre o período de semanas para um ambiente de laboratório movimentado.

[0006] Observamos também o trabalho de Furgale e Barfoot ([6] P. Furgale e T. D. Barfoot, "Visual teach and repeat for long-range rover autonomy", *Journal of Field Robotics*, vol. 27, no. 5, págs. 534- 560, 2010) que desenvolveram um sistema de ensino e repetição usando visão em um ambiente ao ar livre ou externo. Durante a fase de ensino, o robô cria uma série de submapas usando VO. Um módulo de localização é então usado para retraçar o trajeto original. Eles mostram resultados sobre vários quilômetros.

[0007] Todavia, nenhuma das técnicas acima cobrem satisfatoriamente o ambiente modificável que pode ser experimentado quando um sensor (ou um veículo portando este sensor) se move através de um ambiente.

[0008] De acordo com um primeiro aspecto da invenção é provido um método de localização de um veículo que hospeda um sensor, compreendendo pelo menos uma das seguintes etapas de:

- a) capturar dados a partir do sensor provendo uma cena detectada em torno do veículo em um momento atual;
- b) processar a cena detectada, onde o processamento determina uma posição do veículo a partir da cena detectada em relação à posição do sensor em algum momento anterior;
- c) comparar a cena detectada com uma ou mais experiências armazenadas para determinar se a cena detectada pode ser reconhecida dentro

de qualquer da, ou de cada, experiência armazenada;

d) se a cena detectada é reconhecida dentro de uma experiência armazenada, então esta experiência armazenada é usada para produzir uma estimativa da posição do veículo no momento atual com relação à cena armazenada;

e) repetir pelo menos algumas das etapas a até d.

[0009] Acredita-se que as modalidades que provêm um tal método são vantajosas, uma vez que elas são capazes de permitir a localização do veículo dentro de um ambiente modificável mais facilmente do que os sistemas da técnica anterior. Ainda, tais modalidades significam que dados relacionados à cena detectada somente precisam ser armazenados se aparecer a necessidade e uma tal estratégia ajudar a reduzir a quantidade de armazenamento que é necessária conjuntamente com a quantidade de processamento que poderia ser necessária para processar os dados armazenados. Como tal, as modalidades da invenção podem ser mais precisas, podem requerer menos memória do que pode ser esperado e podem ser mais rápidas do que a técnica anterior.

[00010] O termo localização de um veículo é destinado a se referir à determinação, a partir dos dados gerados pelo sensor, da posição do veículo em relação a um momento anterior. Localização é a capacidade de tomar dados a partir do sensor, processar esses dados na presença de uma representação anterior (que é tipicamente provida por um mapa ou, em algumas modalidades, por uma experiência armazenada) de uma cena e chegar a uma parametrização da postura (posição) de veículo com relação àquela representação anterior. Tipicamente, nas modalidades da invenção, a representação anterior é provida por uma pluralidade de experiências armazenadas.

[00011] Convenientemente, o método é arranjado de forma que, se a cena atual não é reconhecida dentro de mais do que um número predeterminado das experiências armazenadas, então o método armazena a cena detectada de forma

que a cena detectada torna-se uma parte de uma nova experiência armazenada para comparações futuras. Acredita-se que tais modalidades são vantajosas pelo fato de que elas aprendem a partir de sua própria experiência. Quando o veículo detecta mais cenas, então o número de experiências armazenadas que o veículo pode usar no futuro é provável que aumente.

[00012] A proposta usada pelas modalidades da invenção nota que as características de uma dada árvore no inverno simplesmente não são as características que são notadas na mesma árvore no verão; os detalhes vistos em uma estrada molhada ao meio-dia são diferentes daqueles vistos ao amanhecer quando a estrada está seca. As modalidades da invenção não forçam as coisas para ser coerentes. Se, por exemplo, um ambiente na terça-feira parecer totalmente diferente do ambiente na quarta-feira, então as modalidades provavelmente tratariam esses como duas experiências diferentes, que capturam igualmente a essência deste ambiente. As modalidades somente amarram as mesmas conjuntamente topologicamente. Recentemente, algumas modalidades podem vincular experiências armazenadas através do conceito de locais, discutido abaixo. Em uma visita inicial de um veículo a uma nova área, as modalidades da invenção podem salvar uma constelação de características visuais (tipicamente provida por dados gerados a partir do sensor) como uma nova experiência.

[00013] Todavia, quando o sensor / veículo se move, as modalidades da invenção são arranjadas para tentar localizar a cena detectada dentro de quaisquer experiências salvas. Como tal, ao revisitar a área, o sensor / veículo tenta usar a transmissão ao vivo de dados (que podem ser imagens) para localizar em quaisquer experiências salvas. Em algumas modalidades, o método pode ser arranjado de forma que as experiências armazenadas são geradas pelo veículo / sensor que usa as mesmas.

[00014] Todavia, em outras modalidades, as experiências armazenadas podem ser usadas por veículos e/ou sensores diferentes do veículo/sensor que

gerou as mesmas. Isto pode ser em adição a ser usadas pelo veículo e/ou sensor que gerou as mesmas. Tais modalidades podem ser imaginadas como o agrupamento de experiência que permitem que uma pluralidade de veículos acesse as experiências armazenadas. Assim, um método pode ser vantajoso, uma vez que ele permitiria que dados fossem acumulados e/ou compartilhados por uma pluralidade de veículos.

[00015] À medida que o veículo continua, ainda salvando para esta nova experiência, ele pode também ser arranjado para tentar re-localizar (isto é, reconhecer a cena detectada dentro de uma experiência armazenada) em suas prévia(s) experiência (s) armazenada(s). Se a re-localização teve êxito em qualquer ponto, o salvamento de uma nova experiência pode ser paralisado e o método retorna para localização em sua experiência prévia.

[00016] Um tal método irá tipicamente causar com que as modalidades armazenem mais representações para cenas detectadas que se alteram frequentemente, e menos para cenas detectadas que são mais estáveis.

[00017] Assim, as modalidades podem manipular novas rotas e falhas de localização completas sem problemas - de fato, é tipicamente a falha de localização que impulsiona o salvamento de uma nova experiência. Como tal, a falha de localização pode ser imaginada como sendo a má ou insolúvel associação de dados - o que existiu anteriormente é simplesmente nada agora.

[00018] Em contraste com as propostas prévias da técnica anterior, algumas modalidades podem somente armazenar novas experiências quando as experiências armazenadas atualmente são insuficientes. Um tal método deve, por conseguinte, resultar em cada experiência armazenada representando um diferente modo visual; assim, a deleção ou apagamento não é um problema. As modalidades podem permitir que cada experiência exista em seu próprio quadro de referência, e, como tal, os métodos podem permitir que localização seja paralelizada e permitir o uso de múltiplas experiências em qualquer momento. Isto pode também permitir diferentes modos visuais,

que podem ser significantemente diferentes, para representar o mesmo espaço físico. Em algumas modalidades, os dados a partir do sensor podem fazer parte da transmissão de dados.

[00019] Em uma modalidade, o sensor é um par estereoscópico de câmeras. Assim, a saída a partir do sensor de uma tal modalidade pode ser uma transmissão de imagens e pode ser uma transmissão de pares de imagens.

[00020] O método pode realizar o processamento inicial na saída do sensor, que pode identificar características comuns entre pares de imagens geradas por uma câmera estereoscópica. O processamento inicial pode gerar um modelo em 3D do local do veículo. Assim, algumas modalidades podem extrair um conjunto de características a partir da cena detectada e usar aquelas características extraídas na determinação da posição do sensor, ou veículo, robô, etc. no qual o sensor é montado.

[00021] O modelo em 3D pode prover uma experiência, à qual referência pode ser posteriormente feita. Em algumas modalidades, a experiência pode ser provida por uma pluralidade de conjuntos de características, onde cada conjunto de características é extraído a partir de uma dada cena detectada. Tipicamente, as modalidades causam com que o sensor gere uma transmissão contínua de características em 3D (isto é, pontos dentro dos dados) em relação à trajetória do veículo. Assim, a trajetória pode ser imaginada como uma sequência de pontos em vários momentos no tempo, que têm estéreo pares afixados aos mesmo, a postura relativa a um ponto prévio no tempo (par de imagens) e uma pluralidade de características em 3D.

[00022] O método pode ser arranjado para analisar os dados a partir do sensor para determinar características em 3D (que podem também ser referidas como pontos ou nós) dentro da cena detectada e ainda para vincular aquelas características em 3D dentro de uma imagem (ou outra representação da cena detectada gerada pelo sensor) em um primeiro momento, com uma imagem subsequente em um segundo momento posterior usando uma transformação. Este

vínculo pode prover uma trajetória para o sensor e/ou veículo.

[00023] O método pode ser arranjado para determinar pontos de referência dentro das imagens. Um ponto de referência pode ser uma estrutura reconhecível dentro dos dados a partir do sensor e corresponderá tipicamente a um objeto do mundo real. Por exemplo, um ponto de referência pode corresponder a um canto de um edifício, uma árvore ou similar.

[00024] Em algumas modalidades da invenção, o método pode ser arranjado para gerar uma posição do ponto de referência em relação ao sensor e/ou veículo. Em modalidades nas quais o sensor é uma câmera estereoscópica, então a posição do ponto de referência pode ser determinada usando informação de ambas as imagens.

[00025] Tipicamente, o método pode localizar uma pluralidade de pontos de referência dentro de qualquer uma cena detectada.

[00026] Em outras modalidades, o sensor pode ser provido com um scanner a laser, radar, lidar (Light Detection and Ranging), Lidar *flash* em 3D, Microsoft™ Kinect™, ou similar.

[00027] O método pode utilizar um posicionamento grosso do sensor (ou veículo no qual ele é montado) para determinar aproximadamente a localidade do sensor. Em particular, as modalidades da invenção podem usar um Sistema de Posicionamento Global (GPS); um ou mais acelerômetros ou outro tal sistema de odometria; ou similar, para identificar a localidade do veículo.

[00028] Modalidades podem usar a localidade do sensor (ou veículo no qual ele é montado) para determinar quais experiências armazenadas devem ser consideradas para determinar se uma cena detectada pode ser reconhecida dentro da mesma. Como tal, a localidade do sensor usando o posicionamento grosso pode ser imaginada como usando um enlace mais próximo.

[00029] Outras modalidades, que podem ser adicionais ou alternativas ao uso de posicionamento grosso, pode usar uma técnica de correspondência de imagem para determinar experiências armazenadas que são visualmente similares

à cena detectada. Tais modalidades são sentidas vantajosas, uma vez que experiências armazenadas, visualmente similares, podem prover bons candidatos para prover uma correspondência com a cena detectada.

[00030] O método pode ser arranjado para operar em paralelo, de forma que o método tenta reconhecer a cena detectada dentro de uma pluralidade de experiências armazenadas em qualquer momento. Isto pode ser imaginado como operando uma pluralidade de localizadores em qualquer momento. Em adição ao reconhecimento da cena detectada dentro das experiências armazenadas, o método pode calcular uma transformação a partir dos nós posicionados dentro da cena detectada para aqueles nós dentro de uma experiência armazenada.

[00031] O método pode requerer que a cena detectada possa ser reconhecida dentro de um predeterminado número de experiências armazenadas antes de ser determinado que o veículo foi localizado. Uma modalidade provendo uma tal etapa acredita-se que proveja um maior grau de certeza que o veículo foi, de fato, localizado corretamente.

[00032] O método pode ser arranjado para determinar se experiências armazenadas se referem a um local similar a um outro. Caso experiências armazenadas devam ser determinadas como sendo um local similar, então as modalidades podem vincular tais experiências armazenadas. As modalidades vinculando experiências armazenadas para formar tais locais acreditam-se que são vantajosas, uma vez que locais podem ser usados para assistir no reconhecimento da cena detectada dentro das experiências armazenadas.

[00033] A experiência armazenada pode compreender qualquer dos seguintes: a saída de dado bruto a partir do sensor; a posição grosseira do veículo; o momento do dia; a data; a condição do tempo; nós determinados dentro da saída a partir do sensor; o mapa em 3D gerado do local do veículo.

[00034] De acordo com um segundo aspecto da invenção é provido um sistema arranjado para localizar um veículo, o sistema compreendendo um

sensor arranjado para gerar dados representando uma cena detectada em torno do veículo em um momento atual, circuito de processamento arranjado para processar os dados em que o processamento pode ser arranjado para realizar pelo menos um dos seguintes:

determinar uma posição do veículo a partir dos dados relativos à posição do sensor em um momento precedente;

comparar os dados com uma ou mais experiências armazenadas, mantidos dentro de um dispositivo de armazenamento, para determinar se a cena detectada pode ser reconhecida dentro de qualquer das experiências armazenadas; e

caso os dados devam permitir que a cena detectada seja reconhecida dentro de qualquer experiência armazenada, então usar esta experiência armazenada para estimar a posição do veículo.

[00035] O dispositivo de armazenamento pode ser provido por qualquer forma apropriada de dispositivo, tal como uma unidade de disco rígido, um arranjo de unidades de disco rígido ou similar. De fato, o dispositivo de armazenamento pode ser distribuído.

[00036] Em algumas modalidades, pelo menos uma porção do dispositivo de armazenamento pode ser montada no veículo. Em outras modalidades, pelo menos uma porção do dispositivo de armazenamento pode ser montada fora do veículo e arranjada para ser acessível a partir do veículo.

[00037] O sistema pode compreender um servidor arranjado para se comunicar com o circuito de processamento no veículo. O servidor pode ser arranjado para se comunicar com o circuito de processamento através de qualquer rede apropriada, que pode incluir a Internet e/ou redes de telecomunicações sem fio, tais como 3G, 4G, WIFI, ou similares. O circuito de processamento pode ser arranjado para comparar os dados contra um subconjunto das experiências armazenadas disponíveis. Em algumas modalidades, o circuito de processamento pode ser arranjado para obter, a

partir de um dispositivo de armazenamento remoto, o subconjunto de experiências armazenadas disponíveis, que pode então ser mantido localmente em um dispositivo de armazenamento no veículo.

[00038] De acordo com um terceiro aspecto da invenção é provido um veículo tendo um sensor montado no mesmo, em que o sensor é arranjado para gerar dados representando uma cena detectada em torno do veículo no momento atual, e um circuito de processamento arranjado para processar os dados em que o circuito de processamento pode ser arranjado para realizar pelo menos um dos seguintes:

determinar uma posição do veículo a partir dos dados relativos à posição do sensor em um momento precedente;

comparar os dados com uma ou mais experiências armazenadas para determinar se a cena detectada pode ser reconhecida dentro de qualquer das experiências armazenadas; e

caso os dados devam permitir que a cena detectada seja reconhecida dentro de qualquer experiência armazenada, então usar esta cena armazenada para estimar a posição do veículo.

[00039] De acordo com um quarto aspecto da invenção é provido um servidor arranjado para receber um pedido de rede a partir de uma rede para uma ou mais experiências armazenadas, o servidor sendo arranjado para processar o pedido, para recuperar as experiências armazenadas pedidas a partir de um dispositivo de armazenamento e para transmitir as experiências armazenadas pedidas através da rede.

[00040] De acordo com um quinto aspecto da invenção é provido um meio legível por máquina contendo instruções que, quando lidas por uma máquina, fazem com que a máquina execute quando o método do primeiro aspecto da invenção, como pelo menos uma porção do sistema do segundo aspecto da invenção ou como o circuito de processamento do veículo do terceiro aspecto do sistema.

[00041] O meio legível por máquina referido em qualquer dos aspectos acima da invenção pode ser qualquer dos seguintes: um CDROM; um DVD ROM / RAM (incluindo -R/-RW ou +R/+RW); uma unidade de disco rígido; uma memória (incluindo uma unidade de USB; um cartão SD; um cartão flash compacto ou similar); um sinal transmitido (incluindo um descarregamento a partir da Internet, transferência de arquivo ftp ou similar); um fio; etc.

[00042] As características descritas em relação a qualquer um dos aspectos acima da invenção podem ser aplicadas, *mutatis mutandis*, a qualquer um dos outros aspectos da invenção. Agora segue, somente a título de exemplo, uma descrição detalhada de modalidades da invenção, da qual:

a Figura 1 é uma vista esquemática de um veículo utilizando uma modalidade de um veículo para localizar o veículo em relação ao local do veículo;

a Figura 2 é uma série de fotografias que ilustram como o local do veículo pode variar de acordo com o tempo e/ou o tempo sofridos pelo local;

a Figura 3 ilustra esquematicamente um método da n modalidade da invenção;

a Figura 4 ilustra esquematicamente uma saída a partir de um sistema de odometria visual, por exemplo, como usado no veículo da Figura 1;

a Figura 5 ilustra esquematicamente como dados relacionados a um local podem ser atualizados;

a Figura 6 provê outras ilustrações de cenas, nas quais o local muda frequentemente (a e b) ou que permanece relativamente constante (c e d);

a Figura 7 ilustra o número de experiências geradas e armazenadas a partir de múltiplas viagens em torno de uma pista mostrada na Figura;

a Figura 8 ilustra como a cena detectada foi acrescentada às experiências armazenadas for percursos do terreno de um circuito de teste;

a Figura 9 mostra as estatísticas de tempo para as viagens usadas para gerar os resultados das Figuras 7 e 8;

a Figura 10 ilustra o efeito de reordenamento dos dados em como a cena atualmente detectada é armazenada como uma experiência armazenada;

a Figura 11 mostra o efeito de tempo em como a cena detectada é armazenada como uma experiência armazenada;

a Figura 12 ilustra um exemplo de falha de localização durante uma primeira viagem em torno da pista da Figura 7;

a Figura 13 ilustra e exemplo de falha de localização durante uma segunda viagem em torno da pista da Figura 7;

a Figura 14 ilustra o desempenho de temporização do circuito de processamento do veículo, ilustrado na Figura 1, durante uma terceira viagem em torno da pista da Figura 7;

a Figura 15 é um fluxograma delineando uma modalidade de exemplo; e

a Figura 16 mostra uma outra modalidade, distribuída, da presente invenção.

[00043] Modalidades da invenção são descritas em relação a um sensor 100 montado em um veículo 102 e em relação ao fluxograma da Figura 15. O sensor 100 é arranjado para monitorar seu local e gerar dados baseados na monitoração, provendo assim dados sobre uma cena detectada em torno do veículo 1500. Na modalidade sendo descrita, então, uma vez que o sensor é montado em um veículo 102, então o sensor 100 é também arranjado para monitorar o local do veículo. Na modalidade sendo descrita, o sensor 100 é um sensor passivo (isto é, ele não cria radiação e meramente detecta radiação) e em particular é uma câmera estereoscópica; ele comprehende duas câmeras

104, 106. A pessoa especializada apreciará que um tal sensor poderia ser provido por duas câmeras separadas. Em outras modalidades, o sensor 100 pode compreender outra forma de sensor, tal como um scanner a laser ou similar. Como tal, o sensor 100 pode também ser um sensor ativo arranjado para enviar radiação para fora do mesmo e detectar radiação refletida.

[00044] Na modalidade mostrada na Figura 1, o veículo 102 está se deslocando ao longo de uma estrada ou rodovia 108 e o sensor 100 está formando imagens do local (por exemplo, o edifício 110, estrada 108, etc.) à medida em que o veículo 102 se desloca. Nesta modalidade, o veículo 102 também compreende circuito de processamento 112 arranjado para capturar dados a partir do sensor e subsequentemente para processar os dados (neste caso, imagens) gerados pelo sensor 100. Assim, o circuito de processamento captura dado a partir do sensor 100 dado este que provê uma cena detectada a partir de em torno do veículo em um momento atual. Na modalidade sendo descrita, o circuito de processamento 112 também compreende, ou tem acesso a, um dispositivo de armazenamento 114 no veículo.

[00045] A porção inferior da Figura mostra componentes que podem ser encontrados em um típico circuito de processamento 112. Uma unidade de processamento 118 pode ser provida, que pode ser um processador Intel® X86, tal como um processador I5, I7, ou similar. A unidade de processamento 118 é arranjada para se comunicar, através de uma barra coletora de sistema 120, com um subsistema de E/S 122 (e assim com redes externas, exibidores, e similares) e uma memória 124.

[00046] A pessoa especializada apreciará que a memória 124 pode ser provida por uma variedade de componentes incluindo uma memória volátil, uma unidade de disco rígido, uma memória não volátil, etc. De fato, a memória 124 compreende uma pluralidade de componentes sob o controle da unidade de processamento 118. Todavia, tipicamente a memória 124 provê uma porção de armazenamento de programa 126 arranjada para armazenar

código de programa que, quando executado, realiza uma ação, e uma porção de armazenamento de dados 128 que pode ser usada para armazenar dados ou temporariamente e/ou permanentemente.

[00047] Em outras modalidades, pelo menos uma porção do circuito de processamento 112 pode ser provida remotamente a partir do veículo. Como tal, é concebível que o processamento dos dados gerados pelo sensor 100 seja realizado fora do veículo 102 ou parcialmente no, e parcialmente fora do, veículo 102. Em modalidades, nas quais o circuito de processamento é provido tanto no, quanto fora do, veículo, então uma conexão de rede (tal como uma 3G UMTS (sistema de Telecomunicação Móvel Universal) ou Wi-Fi (IEEE 802.11) ou similar).

[00048] É conveniente se referir a um veículo que está se deslocando ao longo de uma estrada, mas a pessoa especializada apreciará que as modalidades da invenção não precisam ser limitadas a veículos terrestres e poderiam ser embarcações de transporte na água, tais como navios, botes ou similares ou, mais especificamente, aeronaves de transporte aéreo, tais como aviões, ou similares. Igualmente, é conveniente na seguinte descrição se referir a dados de imagem gerados por câmeras 104, 106, mas outras modalidades da invenção podem gerar outros tipos dos dados.

[00049] O sensor 100, conjuntamente com o circuito de processamento 112, ao qual o sensor 100 é conectado, conjuntamente com o software rodando no circuito de processamento 112, formam o que é frequentemente denominado um sistema de Odometria Visual (VO). Na modalidade sendo descrita, o sistema de Odometria Visual (VO) continuamente produz um modelo em 3D (possivelmente efêmero) do mundo usando os dados gerados a partir das câmeras (104, 106). Tipicamente, o sistema de VO posiciona características (que podem ser referidas como nós ou pontos) dentro de cada imagem a partir do par de câmeras, que podem ser posicionadas em ambas as imagens. Essas características são então rastreadas entre subsequentes imagens (isto é, entre um primeiro momento ou

momento anterior e um segundo momento, talvez o momento atual) para gerar uma trajetória do veículo 102.

[00050] Este arranjo de câmeras pode ser referido como um par estéreo. Na modalidade sendo descrita, quando o veículo 102 está funcionando, o sistema de VO está sempre ligado, sempre consumindo uma transmissão ao vivo de dados (isto é, imagens) a partir das câmeras (104, 106) e estimando as transformações relativas entre as imagens geradas a partir das duas câmeras 104, 106 (isto é, posturas de câmera) para gerar a trajetória e produzir a posição das características em relação àquelas posturas de câmera. A natureza estereoscópica das câmeras permite que a posição em 3D das características sejam calculadas em relação às câmeras.

[00051] Quando o sensor 100/veículo 102 se move, uma pluralidade de conjuntos de locais de característica é gerada; isto é, um conjunto de características é gerado a partir de cada cena. Uma pluralidade de conjuntos de locais pode ser referida como uma experiência. O circuito de processamento 112 pode ser arranjado para armazenar pelo menos um dos seguintes meta dados para uma experiência: os dados gerados a partir do sensor 100 (que, nesta modalidade, são imagens a partir de duas câmeras); as características (isto é, locais de característica em 3D); o momento e/ou data; o tempo; o local (através do Sistema de Posicionamento Global GPS ou similar); uma medida de luminância; nível de tráfego, cobertura por nuvens, tipo de sensor, ou similar.

[00052] Por conseguinte, quando o veículo 102 se move, uma experiência é armazenada como um conjunto de posturas relativas e locais de característica (isto é, nó). Deve ser notado que é um conjunto de posturas relativas que é armazenado. As modalidades desta invenção geralmente não operarão usando um único quadro global. As modalidades da invenção proporcionam uma ideia metricamente correta do movimento da câmera e locais de característica em 3D na vizinhança da postura atual do sensor 100.

As modalidades da invenção não precisam determinar o local dos objetos que estão longe a partir do sensor e que não podem ser detectados (isto é, neste caso, vistos) pelo sensor 100. Como descrito daqui em diante, quando o veículo 102 revisita uma área, localização é tentada nas experiências prévias (armazenadas) que são relevantes para esta área; tal localização pode ser imaginada como a tentativa de determinar a posição do veículo 102 a partir da cena detectada em relação à posição do sensor em um momento precedente no qual a posição do sensor em um momento precedente é provida pela, ou por cada, experiência armazenada. Neste contexto, localização pode ser imaginada como o reconhecimento da cena detectada, provida pelo sensor 100 no momento atual, dentro de pelo menos uma das experiências armazenadas que provê o local do sensor em um momento precedente.

[00053] Como tal, e ao longo do tempo quando o veículo 102 se move em torno de um conjunto de experiências que são armazenadas, cada uma das quais é independente uma da outra. Um localizador pode ser associado com cada uma das experiências armazenadas e ser arranjado para tentar reconhecer a cena detectada dentro de sua experiência armazenada. Em outras modalidades, um localizador pode ser arranjado para processar uma pluralidade de experiências armazenadas.

[00054] Modalidades da invenção são arranjadas para processar experiências armazenadas para tentar localizar o local atual do veículo 102 / sensor 100 dentro de daquelas experiências prévias armazenadas; isto é, as modalidades podem comparar a cena detectada com as cenas armazenadas, em uma tentativa de localizar o veículo 102 (a etapa 1506). Como tal, as modalidades podem ser imaginadas como compreendendo um comparador 134 (isto é, um localizador). Por localização é entendido que o circuito de processamento 112 pode determinar se a imagem ao vivo atual pode corresponder, com êxito, a uma experiência prévia (armazenada). Esta localização é realizada por um localizador 130 funcionando no circuito de

processamento 112 e a localização torna-se possível por manter as experiências prévias armazenadas independentes.

[00055] Modalidades da invenção podem também ser arranjadas para determinar pontos de referência dentro da cena detectada, os quais tipicamente correspondem a objetos do mundo real, tal como o canto de um edifício ou similar. Caso um ponto de referência deva ser reconhecido dentro da área detectada, então a posição do ponto de referência em relação ao veículo é determinada. Na presente modalidade, isto é realizado por um cálculo geométrico, uma vez que o par de câmeras estéreo provê duas imagens nas quais o ponto de referência pode ser encontrado, fornecendo assim suficiente informação para determinar o local do ponto de referência em relação ao sensor. Quando um ponto de referência é criado, então uma correção de dado é armazenada. Na modalidade sendo descrita, esta é dado de pixel e, em um exemplo, uma grade de 9x9 pixels é armazenada. Algumas modalidades podem armazenar quaisquer pontos de referência e/ou correções identificados de dados como parte da experiência armazenada.

[00056] A fim de equiparar a cena detectada às experiências armazenadas, algumas modalidades tentam localizar pontos de referência encontrados nas experiências armazenadas nos pontos de referência posicionados na cena detectada. Isto pode ser obtido por tentativa de localizar a correção de dados (neste caso, as grades de 9x9 de pixels) associada com pontos de referência nas cenas armazenadas com daqueles encontrados na cena detectada. Na modalidade sendo descrita, um localizador é associado com a experiência armazenada e procura por correções de dados associados com os pontos de referência na experiência armazenada com a qual ela está associada na cena detectada.

[00057] Algumas modalidades podem realizar a equiparação das correções em um alto nível, a fim de que correções candidatas apropriadas possam ser posicionadas rapidamente e finalmente realizem uma comparação

mais detalhada para identificar se uma equiparação é atualmente considerada como tendo ocorrido.

[00058] Localização nas experiências prévias pode ser realizada em paralelo e permite ao sistema utilizar relevantes experiências prévias. Na realidade, no tempo de execução é geralmente encontrado que o número de experiências ativas e localizadas com êxito é pequeno. Como descrito daqui em diante, tipicamente, cada nova experiência é somente criada porque os dados podem somente ser equiparados em menos do que um número predeterminado (N) de experiências armazenadas (na modalidade sendo descrita, isto significa que as imagens geradas pelas câmeras 104, 106 que não podem ser equiparadas é 2 ou mais experiências armazenadas e localização não podem ser realizadas). Por conseguinte, subsequentes visitas a uma área deve ser capaz de localizar em somente um pequeno número de experiências armazenadas, pois elas são, por construção, visualmente diferentes. Assim, se o método determinar que a cena detectada não pode ser equiparada em experiências suficientes (isto é, mais do que o número predeterminado), e assim localização não pode ser realizada, a cena detectada é tipicamente armazenada como uma nova experiência armazenada que pode ser usada para a etapa de comparações futuras 1508.

[00059] Assim, as modalidades da invenção provêm uma estrutura para a navegação em longo prazo em um ambiente modificável. Como descrito, brevemente, acima, o veículo 102 constantemente produz um modelo em 3D do ambiente a partir de um sistema de VO que toma a transmissão de imagens ao vivo a partir do sensor 100 como sua entrada. Ao mesmo tempo, o circuito de processamento 112 do veículo 102 oportunisticamente tenta localizar a posição atual nas experiências previamente armazenadas.

[00060] Se a localização é com êxito em uma experiência prévia, as modalidades assumem que a representação do local atual é suficiente e é então possível produzir uma estimativa da posição do veículo no momento

atual com relação à cena armazenada. Todavia, se a localização não puder ser realizada nas experiências prévias, uma nova experiência é criada por salvamento da saída do sistema de VO.

[00061] Experiências são denotadas por  $\epsilon$  e a  $j$ -ésima experiência é referida como  $^j\epsilon$ . Assim, o conjunto de todas as experiências cria um mapa, que é convenientemente referido como um mapa plástico. Este mapa é dado pela equação (1).

$$\mathcal{PM} = \{^j\mathcal{E}\} \quad \forall j \quad (1)$$

[00062] Tipicamente, as modalidades da invenção contam com a capacidade de navegar localmente e de ser capazes de "fechar o enlace"; reconhece quando o veículo 102 / sensor 100 retornou para um local previamente visitado (experiência previamente armazenada).

[00063] Técnicas para navegar localmente são mostradas em vários documentos que provêm sistemas de estimativa de trajetória. Em particular, uma técnica que usa um scanner a laser é discutida em:

[7] M. Bosse e R. Zlot, "Map Matching and Data Association for Large-Scale Two-dimensional Laser Scan-based SLAM", *International Journal of Robotics Research*, vol. 27, págs. 667-691, 2008.

Técnicas para a estimativa de trajetória usando um sistema de visão são mostradas nos documentos:

[8] D. Nister, O. Naroditsky, e J. Bergen, "Visual Odometry for Ground Vehicle Applications", *Journal of Field Robotics*, vol. 23, 2006.

[9] C. Mei, G. Sibley, M. Cummins, P. Newman, e I. Reid, "Real: A system for large-scale mapping in constant-time using stereo", *International Journal of Computer Vision*, págs. 1-17, 2010, edição especial de BMVC.

[00064] O conteúdo desses documentos é aqui incorporado para referência e a pessoa especializada é orientada para ler e compreender os mesmos, particularmente com referência acerca de como uma trajetória pode

ser determinada para o veículo a partir do sistema de VO.

[00065] Outras técnicas de fechamento de enlace, que podem ser usadas para determinar se o dado de sensor atual foi experimentado em uma experiência prévia, são ensinadas, em relação a um sistema de visão em:

[10] M. Cummins e P. Newman, "Highly Scalable Appearance-Only SLAM FAB- MAP 2.0", em *Robotics Science and Systems*, 2009.

[00066] Novamente, os conteúdos deste documento são aqui incorporados para referência e a pessoa é orientada para ler e compreender o mesmo, particularmente com referência à implementação de um sistema de fechamento de enlace a partir do sistema de VO. Um tal sistema tipicamente tenta localizar imagens (isto é, prover técnicas de correspondência de imagem) dentro das experiências armazenadas que são visualmente similares àquelas da cena detectada. Como tal, FAB MAP pode ser imaginada como um enlace mais próximo que pode ser usado para disparar um localizador associado com a experiência armazenada, para tentar localizar esta experiência armazenada com a cena detectada.

[00067] A pessoa especializada apreciará também que Sistemas de Posicionamento Globais (GPS) são disponíveis, que são capazes de prover um local para uma precisão de poucos metros. Assim, as modalidades da invenção podem usar a saída de um sistema de GPS 116, ou outro sistema similar, para fornecer uma posição aproximada, ou grosseira. A saída do sistema similar 116 será tipicamente conectada ao circuito de processamento para permitir ao circuito de processamento identificar uma posição grosseira do veículo 102. O termo posição grosseira é destinado a transportar um significado de relativamente imprecisa, quando comparado com o posicionamento que pode ser obtido usando o processamento dos dados a partir do sensor 100. Todavia, as modalidades da invenção, usando a saída a partir do sensor 100, podem ser capazes de determinar o movimento do veículo 102 em relação à experiência prévia armazenada até a ordem de

milímetros, ou pelo menos poucos centímetros.

[00068] Assim, como descrito acima em termos amplos, o sistema de VO opera em uma sequência de quadros estéreo  $\mathcal{F}^k = \{\mathcal{F}_0, \dots, \mathcal{F}_k\}$ .

[00069] No momento  $k$ , um quadro estéreo é processado e um nó de câmera  $n_k$  é vinculado ao nó  $n_{k-i}$  por uma transformada de 6 graus de liberdade  $t_k = [x, y, z, \theta_r, \theta_p, \theta_q]^T$ , onde  $\theta_r$ ,  $\theta_p$  e  $\theta_q$  são volta, inclinação e guinada, respectivamente. Assim, o nó  $n_{k-i}$  pode ser imaginado como ocorrendo em um primeiro momento e o nó  $n_k$  pode ser imaginado como ocorrendo em um segundo momento posterior. Como tal, o circuito de processamento 112 identifica nós dentro da cena detectada 1502. Esses nós podem ser rastreados sobre uma série de quadros, como descrito daqui em diante, para gerar uma trajetória para o veículo 1504. As modalidades podem ser imaginadas como compreendendo um localizador de nó para realizar esta função 132.

[00070] Se novos pontos de referência em 3D são inicializados como  $\mathcal{F}_k$ , então esses são também afixados ao nó de câmera  $n_k$ . Denotamos o  $i$ -ésimo de tal ponto de referência afixado a  $n_k$ , onde  $i$  é um índice global (de forma que cada ponto de referência tem uma única ID), como  $l_{i,k} = [x, y, z]^T$  - um vetor no quadro de câmera no momento  $k$ . O nó de câmera  $n_k$  é também arranjado para conter uma lista de todos os pontos de referência observados em  $\mathcal{F}_k$ , muitos dos quais serão afixados aos outros nós - aqueles nos quais eles foram inicializados.

[00071] Como tal, um ponto de referência descrito no quadro  $p$  pode ser vinculado a um diferente quadro  $q$  por transformação de  $l^*,_{p}$  para  $l^*,_{q}$  pela operação representada por  ${}^p\Pi_q$ , de forma que:

$$l^*,_{q} \leftarrow {}^p\Pi_q(l^*,_p) \quad (2)$$

[00072] Na modalidade sendo descrita, o sistema de VO funciona continuamente na transmissão de quadros ao vivo (isto é, dados) gerada pelas câmeras 104, 106. Se uma nova experiência está sendo criada (ver a discussão abaixo de quando isto ocorre), então uma nova experiência  ${}^j\varepsilon$  é criada e a

saída a partir do sistema de VO é armazenada nesta experiência. Isto então, pode, por conseguinte, simplesmente ser uma cadeia de nós de câmera  $n_k$ , as transformadas inter-nós e associadas características em 3D. Referimo-nos a nós nas experiências como  $^j\varepsilon_m$ . Posteriormente explicaremos como essas cadeias são relacionadas (topologicamente) para formar em conjunto um mapa plástico.

[00073] Assim, a Figura 4 ilustra como a experiência é construída a partir da série de quadros que são os nós  $n$ . Cada quadro da saída a partir do sistema de VO (representado pelos círculos) é denotado por  $n_{k-2}$  através de  $n_k$  até  $n_{k+1}$ ; com o quadro atual  $n_k$  referenciado com o numeral 400. Cada um dos nós  $n$  é vinculado pelos 6 graus da transformação de liberdade que descreve o movimento da câmera entre  $\mathcal{F}_{k-1}$  e  $\mathcal{F}_k$ .

[00074] Ainda, uma série de pontos de referência 402, 404, 406, 408 é notada com o local do veículo 102 dentro do quadro atual  $\mathcal{F}_k$ . Cada um desses pontos de referência 402-408 pode ser transformado de forma que eles estão em relação a  $k$  usando a equação 2 acima.

[00075] Como descrito em termos amplos acima, um localizador é usado para tentar localizar o local atual do veículo 102 dentro de uma experiência prévia. Em termos amplos, isto pode ser imaginado como o reconhecimento da cena detectada dentro da experiência prévia (que pode ser a experiência armazenada). Mais que um localizador pode ser provido e o, ou cada, localizador corre sobre a experiência prévia (isto é, a experiência armazenada). Dado um quadro ao vivo  $\mathcal{F}_k$ , a tarefa do, ou de cada, localizador é a de calcular uma transformação do quadro  $\mathcal{F}_k$  para um nó  $n$  na experiência prévia.

[00076] Os localizadores operam de uma maneira similar a o sistema de VO ao vivo, exceto que o ajuste do ponto de referência proposto provém da experiência salva, não do quadro prévio  $\mathcal{F}_{k-1}$ . Os pontos de referência são tomados a partir da região local circundando a posição prévia na experiência.

Adicionalmente, na modalidade sendo descrita, o localizador não tenta acrescentar ou atualizar pontos de referência em qualquer da saída de VO atual ou da experiência prévia. ele é completamente passivo em termos de seu impacto sobre ambas; isto é, ele não modifica os dados da saída de VO atual ou da experiência prévia.

[00077] Localizadores são arranjados para saber se ele está "perdido"; isto é, quando o quadro de chegada  $\mathcal{F}_k$  não mais pode ser localizado na experiência prévia. Existem muitas maneiras nas quais isto pode ser calculado. Possibilidades incluem o número de pontos de referência encontrados e/ou classificados como formações envoltas por outras mais recentes, e comparações com a saída de VO atual. A saída de cada localizador em cada etapa de tempo é um resultado binário indicando se ele é ainda localizado com êxito:

$$L(j\mathcal{E}, \mathcal{F}_k) = \begin{cases} 1 & \text{se localizado} \\ 0 & \text{se perdido.} \end{cases} \quad (3)$$

[00078] Se com êxito, o localizador pode ser consultado com respeito ao nó na experiência, ao qual  $\mathcal{F}_k$  foi o mais próximo.

$$m \leftarrow {}^j\bar{\mathcal{E}}(\mathcal{F}_k) \quad (4)$$

[00079] Uma vez quando um localizador é declarado perdido, ele fica neste estado até receber assistência do exterior, discutida abaixo e que fecha o enlace novamente com a experiência prévia.

[00080] N é definido para ser o número mínimo aceitável de localizadores com êxito em qualquer momento (isto é, uma tentativa com êxito para determinar a posição atual dentro de uma experiência prévia ou, em outros termos, N é um número mínimo predeterminado de experiências armazenadas, em que a cena detectada deve ser reconhecida antes da localização ser considerada que tenha ocorrido). Para cada quadro  $\mathcal{F}_k$ , o conjunto de experiências ativas A é computado. Para cada experiência em A, seu localizador associado funciona e o resultado é um vetor binário S

indicando o sucesso ou falha de cada. Se  $|S|$  cair abaixo de N, uma nova experiência  ${}^j\mathcal{E}$  é criada e a saída de VO é salva até  $|S|$  retornar para maior que ou igual a N. Este processo é descrito pelo Algoritmo 1:

[00081] Algoritmo 1 Mapeamento plástico

---

**Algoritmo 1** Mapeamento plástico

---

```

bool Saving
while 1 do
     $\mathcal{F}_k = GetFrame()$ 
     $\{n_k, t_k\} \leftarrow \mathbf{VO}(\mathcal{F}_k)$ 
     $\mathcal{A} = ComputeActiveLocalisers(\mathcal{PM})$ 
     $S = Localise(\mathcal{A}, \mathcal{F}_k)$ 
    if  $|S| < N$  then
        Saving = True
         ${}^j\mathcal{E} \leftarrow \{{}^j\mathcal{E}, {}^j\mathcal{E}_m\}$ 
    else if  $|S| >= N$  && Saving==True then
        Saving = False
         $\mathcal{PM} \leftarrow \{\mathcal{PM}, {}^j\mathcal{E}\}$ 
    end if
end while

```

---

[00082] Como tal, se a cena detectada pode ser localizada em mais do que N cenas armazenadas, é possível determinar a posição do veículo 102 em relação àquelas experiências armazenadas - etapa 1512.

[00083] Este algoritmo pode ser explicado com referência à Figura 3 e, em uma região superior desta Figura 3, 'Épocas' separadas são mostradas, onde cada Época representa um conjunto de quadros  $\mathcal{F}_k$  a partir da saída dos sistemas de VO em três momentos separados. Nesta modalidade, o número mínimo de localizadores com êxito (N) foi ajustado em dois.

[00084] Para a primeira Época A, na esquerda da Figura, cada um dos quadros foi designado como  $F_{k-2}$ , de  $F_k$  até  $F_{k+3}$ . Cada das outras Épocas (B e C) poderia ser similarmente designada, mas não foi por razões de clareza.

[00085] Abaixo de cada dos quadros da Época A até C, estão três representações: A fileira superior de círculos representa a saída atual do sistema de VO; a segunda fileira, a fileira central de círculos representam dados armazenados em uma experiência 1; e a terceira fileira, fileira inferior

de círculos representam dados armazenados em uma experiência 2.

[00086] Na Época A, pode ser visto que as imagens criadas pelo sensor são uma combinação de duas diferentes tendências de dados que podem ser imaginados como sendo 'pretos/cinzas' 300 e linhas sólidas horizontais 302. Olhando para a Experiência 1, pode ser visto que existe uma tendência de linha sólida horizontal nos dados e, como tal, um localizador é capaz de localizar as imagens atuais com a Experiência 1. Olhando para a Experiência 2, pode ser visto que existe uma tendência de 'pretos/cinzas' nos dados e, como tal, um localizador é capaz de localizar as imagens atuais com a Experiência 2. Observando que  $N = 2$ , não existe agora nenhuma necessidade de criar uma nova experiência, uma vez que já estamos localizados no número mínimo e a saída de VO pode simplesmente ser esquecida.

[00087] Movendo-se para a Época B, pode ser visto que as imagens criadas pelo sensor são uma combinação de duas diferentes tendências de dados que podem ser imaginados como sendo pretos/cinzas 300 e linhas tracejadas horizontais 304. Olhando para a Experiência 1, a tendência de linha horizontal sólida 392 não pode ser localizada nas imagens atuais. Todavia, olhando para a Experiência 2, pode ser visto que existe uma tendência de pretos/cinzas como nos dados atuais e, como tal, um localizador é ainda capaz de localizar os dados atuais dentro de Experiência 2. Todavia, uma vez que  $N$  é ajustado em 2 e somente uma única localização ocorreu, a saída do sistema de VO é agora armazenada como uma nova experiência 306.

[00088] então, se movendo para a Época C, a tendência nas imagens atuais retorna para pretos/cinzas e linhas horizontais sólidas que podem ser localizadas tanto na Experiência 1 quanto na Experiência 2. Como tal, a saída a partir do sistema de VO pode novamente ser esquecida.

[00089] A pessoa especializada apreciará que às modalidades podem armazenar a saída de VO mesmo se mais que  $N$  localizações ocorreram, mas é provavelmente vantajoso que isto não ocorra. o armazenamento dos dados

desnecessariamente irá conduzir a um maior acúmulo de dados e ao possível reduzido desempenho do sistema.

[00090] O mapa plástico (identificado na equação (1)), para tanto, armazena muitas experiências que cobrem uma área espacial ilimitada, em adição à captura de diferentes aparências da mesma área, assim as experiências não serão todas relevantes todas as vezes; isto é, o sistema de VO pode estar fornecendo informação a partir de uma diferente área espacial e/ou diferente aparência visual. A posição espacial do veículo em seu ambiente não será capturada por todas as experiências. Por conseguinte, é desejável calcular quais experiências são aplicáveis para  $\mathcal{F}_k$ . As experiências não são armazenadas em um único quadro de referência, assim não é possível integrar transformadas locais para a estimativa da posição de uma na outra.

[00091] Para superar a incapacidade da estimativa da posição em um quadro a partir de outro, introduzimos a ideia de locais, que criam vínculos topológicos entre as experiências. Os locais são referidos como  $P$  e o  $z$ -ésimo local como  $P^z$ .

[00092]  $P^z = \{ {}^j \varepsilon_m \}$ , é o conjunto de todos os nós (tomados de várias experiências) que são conhecidos que foram simultaneamente localizados contra. Assim,  $P^z$  é um conjunto de quadros de câmera visualizando o mesmo local físico.

[00093] Quando  $\mathcal{F}_k$  é simultaneamente localizado em mais do que uma experiência, podemos criar um local. Por indagação de cada localizador com êxito por seu nó mais próximo, podemos criar um local com o conjunto  $\{ {}^j \varepsilon_m \}_k$ . Todavia é também possível que um local previamente criado contenha alguns desses nós, em cujo caso fundimos os conjuntos criados a partir de  $\mathcal{F}_k$  e o local prévio.

$$\mathcal{P}^z = \mathcal{P}^z \cup \underbrace{\left[ \bigcup_j {}^j \bar{\mathcal{E}}(\mathcal{F}_k) \right]}_{\mathcal{F}_k} \quad (5)$$

[00094] Onde  $P^z$  no lado direito é potencialmente um conjunto vazio,

dependendo de se os nós gerados a partir de  $\mathcal{F}_k$  estão nos locais prévios. Um exemplo simples de criação e fusão de local é mostrado na Figura 5 e também explicado com referência à Figura 3.

[00095] Com referência inicialmente à Figura 3, pode ser visto que existem, com efeito, 3 experiências referidas com nesta Figura: a experiência denotada pela linha horizontal sólida; a experiência denotada pelos pretos/cinzas; e a experiência denotada pela linha horizontal tracejada. A Figura 5 também mostra 3 experiências: Experiências 1, 2 e 3 denotadas pelos círculos de preto, cinza e branco, respectivamente.

[00096] A Figura 5 mostra que em um quadro prévio  $F_{\text{prévio}}$  foi possível localizar este quadro tanto na Experiência 1 quanto na Experiência 2 e assim gerar um local A. Todavia, no quadro atual  $F_k$  é possível localizar na Experiência 2 e na Experiência 3 e assim gerar um local B. Consequentemente, pode ser concluído que os locais A e B atualmente se referem ao mesmo local (devido à inclusão da Experiência 2 em ambos) e, como tal, os locais A e B podem ser fundidos para criar o local C.

[00097] Assim, com referência novamente à Figura 3, poderia ser inferido que todas das três das experiências referidas aqui (Experiência 1; Experiência 2 e a experiência armazenada a partir do VO na Época B), todas atualmente se referem ao mesmo local, uma vez que a Experiência 2 é comum a todas as três Épocas A até C.

[00098] Assim, e como discutido em relação à Figura 3, locais podem ser usados para vincular experiências. Usamos a capacidade de consultar todos os locais com um conjunto de nós de experiência, e receber o conjunto de locais que contêm esses nós. Esta consulta é definida através de  $\Omega$ .

$$\{\mathcal{P}^z\} \leftarrow \Omega(\{^j\mathcal{E}_m\}) \quad (6)$$

[00099] 1) *Experiências Relevantes à Computação*: Como notado acima, nem todas experiências armazenadas serão relevantes para  $\mathcal{F}_k$  quando as experiências armazenadas não irão necessariamente cobrir a posição atual

do veículo. Por conseguinte, as modalidades permitem aos localizadores assumir dois estados, ativo e inativo. Antes do processamento  $\mathcal{F}_k$ , o conjunto de posições de localizador com êxito a partir do quadro prévio,  $\{\mathcal{E}_m\}_{k-1}$ , é usado para gerar o conjunto de locais relacionados:

$$\{\mathcal{P}^z\}_{relacionado} \leftarrow \Omega(\{\mathcal{E}_m\}_{k-1}) \quad (7)$$

[000100] Cada localizador inativo então consulta  $\{\mathcal{P}^z\}_{relacionado}$  quanto a presença de seus próprios nós, e se alguns estão presentes, ele ativa si próprio na posição recuperada. Quando o veículo 102 se move além da faixa de uma experiência (isto é, o localizador fica perdido, como descrito em relação à Figura 3), o localizador associado ajusta-se ele próprio para inativo.

2) *Reinicialização de Localizadores Perdidos*: Localizadores ficam perdidos porque eles não mais podem rastrear o quadro atual  $\mathcal{F}_k$ , todavia este pode somente ser um estado temporário e o localizador pode ser capaz de re-localizar si próprio dentro de uma experiência. Como tal, este localizador / experiência pode se tornar relevante novamente em um curto tempo depois do localizador inicialmente se tornar perdido, e, por conseguinte, as modalidades tentam reinicializar localizadores perdidos. Falha em reinicializar localizadores perdidos, onde possível, provavelmente conduziria ao sistema salvar mais experiências do que necessárias, com os problemas associados de as exigências de armazenamento elevadas para os dados e reduzido desempenho de processamento em vista dos dados extras.

[000101] Como tal, as modalidades da invenção podem avaliar se, dado um local conhecido em uma experiência, a posição é conhecida de outra experiência através de um dos locais. Por conseguinte, algumas modalidades podem ser arranjadas, que localizadores perdidos periodicamente consultam  $\{\mathcal{P}^z\}_{relacionado}$  para ver se eles podem ser reinicializados. Às vezes, todos dos processos de localização tornam-se perdidos, em cujo ponto o sistema não sabe onde ele está no mapa plástico. Isto pode ser a razão pela qual o local atual se alterou显著mente (por exemplo, está nevado), ou porque o

veículo 102 está explorando uma nova rota. Por qualquer que seja o motivo, o sistema de VO continuará a processar a transmissão ou corrente de quadros (dados) ativos e estará salvando a saída para uma nova experiência, como descrito em relação à Figura 2 com respeito à Época B. Todavia, é possível que o veículo 102 subsequentemente retorne para um local, no qual ele pode se localizar subsequentemente com êxito, e retomar sua posição no mapa plástico; isto é, o veículo 102 pode ser capaz de, mais uma vez, localizar em um número de experiências maior que N. Para uma solução para este problema, que deve se tornar cada vez menos frequente sobre repetidos percursos do terreno, usamos um enlace externo mais próximo, que pode reinicializar localizadores perdidos.

[000102] O fechamento de enlace externo é destacado em relação à Figura 7 discutida abaixo na discussão de um veículo que reentra em um circuito previamente explorado a partir de um novo circuito.

[000103] Todavia, em algumas modalidades, sistemas, tais como GPS, ou similares, podem ser usados para fornecer uma posição grosseira do veículo 102. In tais modalidades, a posição grosseira do veículo pode ser usada pelo circuito de processamento para destacar quais experiências armazenadas são candidatas promissoras para considerar a localização do veículo com as mesmas. Como tal, o posicionamento grosseiro do veículo 102 pode ser usado para ajudar a identificar quais experiências armazenadas podem ser usadas para localizar o veículo 102 - etapa 1510. As modalidades podem ser imaginadas como compreendendo um enlace mais próximo 136.

[000104] Em modalidades da invenção, existe geralmente o caso em que, uma vez quando o conjunto de localizadores relevantes foi computado, cada um se torna independente e assim pode ser operado em paralelo. Dado que a associação de dado (isto é, equiparação dos pontos de referência a partir de uma experiência armazenada para cada um dos quadros de câmera estéreos da cena detectada) e etapas de estimativa de trajetória (isto é, a determinação

da alteração em postura do veículo entre o quadro atual e o prévio) dominam o tempo de computação, por paralelismo deles, as modalidades são atualmente capazes de processar quadros em 15 Hz. A pessoa especializada apreciará que, caso outra potência de processamento deva ser acrescentada, então a frequência em que dados podem ser manipulados irá aumentar.

[000105] Embora a pessoa especializada possa apreciar como implementar um sistema que é capaz de processar e gerar trajetórias a partir de uma transmissão de dados de sensor, os seguintes documentos são providos como uma referência. Esses documentos são incorporados para referência e a pessoa especializada é orientada para ler as seções relevantes desses documentos.

[000106] Para obter robusta associação de dados, modalidades da invenção podem usar Descritores de Características Elementares Independentes Robustas Binárias (BRIEF).

[11] M. Calonder, V. Lepetit, C. Strecha, e P. Fua, "BRIEF: Binary Robust Independent Elementary Features", em *European Conference on Computer Vision*, setembro de 2010.

[000107] As características de descriptor BRIEF são rápidas de computar e vincular usando somente uma CPU do circuito de processamento 112 e são capazes de prover desempenho comparável ao SURF [12].

[12] H. Bay, A. Ess, T. Tuytelaars, e L. V. Goo, "Surf: Speeded up robust features", *Computer Vision and Image Understanding (CVIU)*, vol. 110, págs. 346-359, 2008.

[000108] Modalidades da invenção que podem usar implementações de GPU (Unidade de Processamento de Gráfico) de descritores como o SURF são disponíveis e permitem desempenho de taxa de quadro:

[13] N. Cornells e L. V. Gool, "Fast Scale Invariant Feature Detection e Matching on Programmable Graphics Hardware", em *Computer Vision and Pattern Recognition*, 2008.

[000109] Todavia, a etapa de equiparação em cada localizador e o sistema de VO ao vivo requerem o acesso a uma GPU (extração de característica no quadro de chegada é independente e pode ser realizada uma vez quando no início). Como a maioria dos sistemas somente têm uma GPU na melhor das hipóteses, a exigência de usar uma GPU torna a paralelização difícil, em comparação com a rodagem de um programa somente em CPU em um sistema multi-núcleo ou multiprocessor. Como tal, as modalidades tendem a usar métodos baseados em CPU.

[000110] As modalidades podem usar o extrator de canto FAST para computar pontos de interesse para BRIEF:

[14] E. Rosten, G. Reitmayr, e T. Drummond, "Real-time video annotations for augmented reality", em *Advances in Visual Computing. LNCS 3840*, Dezembro de 2005, págs. 294-302.

[000111] Em algumas modalidades, pontos de referência equiparados podem ser refinados para a precisão de sub-pixel usando a eficiente equiparação de segunda ordem, descrita por Mei:

[15] C. Mei, S. Benhimane, E. Malis, e P. Rives, "Efficient homography- based tracking and 3-d reconstruction for single-viewpoint sensors", *IEEE Transactions on Robotics*, vol. 24, no. 6, págs. 1352-1364, Dec. 2008.

[000112] Dado foi coletado a partir de 53 percursos do terreno de duas rotas de 0,7 km semi-superpostas em torno de Begbroke Science Park. dado foi coletado sobre um período de três meses em diferentes momentos do dia e com diferentes condições do tempo usando o veículo de pesquisa do grupo, o Wildcat.

[000113] A Figura 7 mostra uma visão aérea do local, juntamente com as duas rotas percorridas por um veículo, tal como o veículo 102. O enlace externo 700, denotado pela linha mais grossa, foi percorrido nos primeiros 47 percursos do terreno, enquanto os últimos 6 percursos do terreno foram através do enlace interno 702, indicado pela linha mais fina. A intensidade do

traçado indica quantas experiências foram previstas em cada ponto e uma escala é mostrada à direita da Figura. Para finalidades ilustrativas, controlamos o sinal do enlace externo mais próximo, assim ele somente ligou em 14 pontos pré-definidos em cada enlace. Os pontos foram espaçados aproximadamente uniformemente ao longo de cada enlace. Como tal, caso os localizadores devam ter se perdido, então eles poderiam ser re-localizados nesses 14 pontos pré-definidos. A pessoa especializada apreciará que outras modalidades poderiam implementar isto diferentemente; talvez aleatoriamente, substancialmente continuamente, em predeterminados intervalos de tempo ou similares.

[000114] Algumas regiões das rotas requerem mais experiências do que outras. Na Figura 6 mostramos exemplos de locais que exibem tanto baixa quanto alta alterações visuais. Um exemplo de uma região que tem alta alteração visual é um estacionamento de automóveis, Figura 6(a). Os conteúdos e configuração deste espaço varia diariamente, assim as experiências armazenadas em dias prévios são improváveis que sejam úteis. Outro é uma seção de estrada coberta por árvores frondosas, como mostrada na Figura 6(b). A luz solar causa fortes e intrincados efeitos de sombreamento. Embora esses sejam úteis para que o sistema de VO ao vivo proveja pontos de referência nos mesmos, os pontos de referência posicionados não são frequentemente encontrados novamente, significando que experiências prévias não são úteis. imagens que têm um alto grau de alteração visual encorajam o uso do sistema de VO ao vivo em todos os momentos. Em contraste, as Figuras 6c e 6d mostram trechos abertos de estrada que, em similares condições do tempo, têm uma aparência muito constante.

[000115] A Figura 8 mostra quanto de cada percurso do terreno é salvo, juntamente com o momento do dia em que foi capturado. Aqui, os números de visita estão na ordem em que o dado foi coletado. Como tal, pode ser visto que, à esquerda da Figura, quando o mapa plástico é relativamente não

extenso e não muitas experiências foram armazenadas, muito da saída de VO é armazenada. Quando o número de percursos do terreno aumenta, a quantidade de saída de VO que é armazenada reduz quando o veículo 102 é capaz de para localizar si próprio em mais das experiências prévias (como pela discussão em relação à Figura 3, na qual, quando a localização em mais que N experiências ocorre, nova saída de VO não mais pode ser armazenada).

[000116] O grande salto em torno de percursos do terreno 35-38 acontece porque, para o primeiro momento, dado foi coletado ao anoitecer. A estradas também tiveram poças de água estagnada e estava chovendo fracamente, algo que o sistema não encontrou anteriormente. A Figura 13 provê imagens geradas pelo sensor no percurso do terreno 38 e as imagens b) e d) são geradas no percurso do terreno 38, enquanto que imagens a) e c) foram geradas em percursos precedentes do terreno e, como tal, se referem a experiências armazenadas. Assim, pode ser visto que as imagens geradas sobre o percurso do terreno 38 são diferentes daquelas de percursos anteriores do terreno.

[000117] O segundo pico na visita 47 é causado pela condução no enlace interno 702 pela primeira vez. Subitamente, localizações não são com êxito e toda a secção é salva até o enlace mais próximo disparar. O enlace mais próximo deve ser capaz de re-localizar a posição do veículo 102 quando, ou logo depois, o enlace interno 702 se reunir ao enlace externo 704 e os localizadores são novamente capazes de localizar dentro das experiências armazenadas.

[000118] Descrevendo a Figura 8 ligeiramente em mais detalhe, o gráfico mostra (ver o eixo à esquerda) a fração da saída de VO que é salva para o uso futuro. Também, os resultados destacam o efeito de variação de N (o número mínimo de localizações que são requeridas a fim de decidir que o veículo 102 não está perdidos) e resultados são mostrados para o número mínimo de localizadores N = 1, N = 2 e N = 3 por variação do símbolo usado para cada linha do gráfico.

[000119] Para cada percurso do terreno computamos também a fração da saída de VO salva quando cada outra visita foi usada como um única experiência disponível e o resultados colocados na média. Referimo-nos a isto como Média com Única Experiência Prévia (AwSPE) e isto é representado pela linha tracejada marcada com triângulos.

[000120] A Figura 12 mostra exemplos de falhas de localização sobre percurso do terreno 4, onde fortes efeitos de sombreamento são encontrados pela primeira vez: as imagens a) e c) foram geradas a partir de percursos precedentes do terreno e armazenadas como experiências, enquanto que as imagens b) e d) foram geradas sobre o percurso do terreno 4. Assim, como visto em relação às Figuras 12 e 13, onde situações apareceram que causaram uma grande modificação no ambiente (tal como a condução no crepúsculo ou uma diferente rota), é encontrado que, quando o veículo 102 revisita a rota, é tipicamente necessário lembrar cada vez menos da saída de VO sobre cada percurso do terreno.

[000121] A fim de avaliar a vantagem de salvar múltiplas experiências para elaborar um mapa plástico, como descrito em relação às modalidades da invenção como descritas acima, experimentos foram feitos, os quais tentaram para localizar os dados de VO atuais em uma única visita anterior, isto é, não fazendo tentativa para capturar o ambiente modificável. Para esses experimentos, a única experiência armazenada foi limitada a ser uma visita de exemplo selecionada e, em cada percurso do terreno da rota, a quantidade de saída de VO que precisou ser salva foi determinada em cada caso, indicando falha de localização, e os resultados foram colocados na média. Assim, como um primeiro passe, o percurso do terreno número 1 foi tomado para ser o exemplo selecionado e usado como a única experiência armazenada para os percursos do terreno números 2 a 53. Em um segundo passe, percurso do terreno de número 2 foi tomado para ser o exemplo selecionado e usado como a única experiência armazenada para os percurso do terreno de números 1 e 3

a 53. Este padrão foi repetido até cada um dos percursos do terreno ter sido usado como o exemplo selecionado.

[000122] Esses são mostrados na Figura 8, referidos como Média com Única Experiência Prévia (AwSPE). Através de todas as visitas, a saída de VO média salva é 65,8%, e nenhuma visita é particularmente "boa" como uma anterior para outras visitas. Isto motiva o uso de múltiplas experiências para representar o ambiente. Ao longo do tempo, descobrindo novas rotas, as modalidades que geram um mapa plástico, como descrito acima, produzem um decaimento 1 / (Contagem de Percurso do terreno) quando o sistema captura a modificação típica da rota. A ordem em que os dados foram coletados é uma sequência privilegiada. Todavia, ela poderia ser ordenada em 53 ! maneiras para produzir diferentes gráficos de desempenho. Para fazer o ponto, um reordenamento ganancioso foi realizado, em que os 10 mais surpreendentes percursos do terreno da rota externa e os 6 percursos do terreno internos para o início e re-operação do sistema. O gráfico de desempenho resultante é mostrado na Figura 10. Como tal, os mais "interessantes" e "surpreendentes" percursos do terreno foram movidos para o início da criação de mapa plástico e é notado que acúmulo de experiências fica alto (enquanto que os percursos altamente variáveis do terreno estão ocorrendo) até o percurso do terreno 18 antes de cair significantemente.

[000123] É também possível classificar cada percurso do terreno ou como céu encoberto ou ensolarado (o tempo para cada visita é mostrado na Figura 9). Em um outro experimento, o dado foi reprocessado usando somente experiências de céu encoberto ou ensolarado e os resultados das quais são mostrados na Figura 11. (Remove os 4 crepúsculo e 6 percursos de terreno do enlace interno.). De forma interessante, com o céu encoberto, somente percursos do terreno (o gráfico superior da Figura 11), o sistema rapidamente acumula suficientes experiências para gerir condições nubladas, enquanto que percursos do terreno ensolarados (o gráfico inferior da Figura

11) têm uma ligeira tendência de deslocamento. Acredita-se que a ligeira tendência de deslocamento é causada por efeitos de sombreamento que tornam a localização nas experiências prévias difícil; tais sombras podem naturalmente ser vistas na Figura 6b.

[000124] Finalmente, o desempenho do sistema funcionando no veículo de teste é mostrado na Figura 14. É mostrado o número de localizadores com êxito e desempenho de temporização para cada quadro sobre o percurso do terreno 47, que é o primeiro percurso do terreno do enlace interno. Localização é com êxito até o quadro 1296, em cujo ponto o veículo faz a volta para o enlace interno. Neste ponto, uma nova experiência começa e não existe localização, uma vez que não existem experiências relevantes armazenadas. No quadro 2239, o enlace externo mais próximo dispara e resulta em localização com êxito, então salvando as novas paradas de experiência. A despeito de números de localizadores ativos variáveis, a temporização por quadro tipicamente fica situada inferior a 100 ms, enquanto que a média para a parte de localização com êxito da sequência (isto é, não incluindo os quadros 1296-2239) é 53 ms.

[000125] A Figura 16 mostra uma modalidade do sistema que é modificado quando comparado com o sistema na Figura 1. Todavia, por facilidade, as mesmas partes são referidas com os mesmos números de referência. Nesta modalidade, o circuito de processamento 112 é provido com um dispositivo de comunicação sem fio 1600, tal como um MODEM 3G ou similar. O dispositivo de comunicação sem fio 1600 permite ao circuito de processamento 112 se comunicar, através da rede 1602, com um servidor mais remoto 1604. A rede pode ser qualquer rede apropriada, mas pode tipicamente ser a Internet. A pessoa especializada apreciará que, embora o servidor 1604 seja representado como um único dispositivo na Figura, ele pode atualmente ser um número de dispositivos separados, que podem ser diferentes entre si. Em uma modalidade, tal como aquela mostrada na Figura

16, é concebível que o servidor remoto 1604 pudesse compartilhar alguma da carga de processamento do processamento dos dados, gerada pelo sensor 100 (e/ou a saída do sistema de VO). Todavia, o servidor pode também armazenar experiências que podem ser usadas para localizar o veículo 102.

[000126] Como foi discutido acima, a saída de dado a partir do sensor 100 pode variar substancialmente de acordo com o momento do dia, condições do tempo, etc. Como tal, é concebível que somente um subconjunto de experiências armazenadas irá ser útil para localizar o veículo 120 em seu local atual.

[000127] Como tal, o servidor remoto 1604 pode ser arranjado para armazenar, em um dispositivo de armazenamento 1606, um conjunto de experiências, e o veículo 102 pode ser arranjado para baixar um subconjunto das experiências armazenadas. Assim, será apreciado que o dispositivo de armazenamento 1606 provê um dispositivo de armazenamento montado fora do veículo 102. O subconjunto de experiências pode ser baixado para o veículo 102, dependendo dos vários parâmetros incluindo qualquer um ou mais dos seguintes: momento do dia, condições do tempo, rota pretendida, local, ou similar.

[000128] De fato, caso o veículo 102 deva gerar novas experiências durante um percurso de terreno que ele percorre, o veículo 102 pode ser arranjado para carregar esta nova experiência para o servidor remoto 1604. De fato, o conjunto de experiências mantido no servidor 1604 pode ser gerado por uma pluralidade de veículos. Como tal, o servidor 1604 pode prover um repositório de experiências, que pode ser usado por qualquer número de veículos que desejam localizar a si próprios usando aquelas experiências.

[000129] Como tal, as modalidades da invenção podem demonstrar a localização contínua do veículo 120. Esta localização pode ser em condições de iluminação e/ou condições do tempo cambiantes sobre um período de tempo, o qual, nos experimentos, foi um período de 3 meses. As modalidades usadas aqui podem ser denominadas de mapeamento plástico. Como tal, o foco das modalidades não é sobre a formação de um único mapa monolítico

ou dedução um estado subjacente latente que explica todas as observações do espaço de trabalho, mas sim na criação de uma representação compósita construída a partir de múltiplas experiências que se superpõem. Esta representação (isto é, o mapa plástico) pode somente ser tão complexa e rica quanto precisa ser. As modalidades tipicamente manipularão alterações tanto drásticas quanto lentas da mesma maneira - tão logo as experiências prévias falhem em descrever adequadamente a presente, uma nova experiência é salva para referência futura. Experimentos têm mostrado as vantagens de mapas plásticos no desempenho de localização (robustez) e, usando os dados de 3 meses que foram obtidos, foi possível demonstrar o comportamento assintótico da manutenção de mapa plástico.

[000130] A pessoa especializada apreciará que muitas quaisquer das características descritas acima poderiam ser implementadas em software, firmware ou hardware ou de fato, por uma combinação desses. Como tal, pretende-se que sistemas implementados por qualquer combinação dessas técnicas sejam cobertos.

## REIVINDICAÇÕES

1. Método de localização de um veículo (102) que hospeda um sensor (100), caracterizado pelo fato de que compreende as etapas de:

a) capturar (1500) dados a partir do sensor (100) provendo uma cena detectada em torno do veículo (102) em um momento atual;

b) processar (1502) a cena detectada para extrair um conjunto de características a partir da cena detectada e para estimar, usando as características extraídas, uma posição do veículo (102) a partir da cena detectada em relação à posição do sensor (100) em um momento precedente;

em que a estimativa é realizada por:

c) comparar (1506) o conjunto extraído de características a partir da cena detectada com uma pluralidade de experiências armazenadas, para verificar se a cena detectada pode ser reconhecida em qualquer uma das experiências armazenadas, em que cada experiência armazenada compreende uma pluralidade de conjuntos de características, onde cada conjunto de características foi determinado a partir de uma cena detectada previamente, e as experiências armazenadas cobrem uma área espacial não ligada, além de capturar diferentes aparências da mesma área;

d) se o conjunto extraído de características da cena detectada é reconhecido dentro de qualquer uma das experiências armazenadas, então esta experiência armazenada é usada para produzir uma estimativa (1512) da posição do veículo (102) no momento atual com relação à experiência armazenada; e

e) se a cena detectada não for reconhecida dentro de mais do que um número predeterminado de experiência armazenada, então a cena detectada torna-se parte de uma nova experiência armazenada para comparações futuras.

2. Método de acordo com a reivindicação 1, caracterizado pelo fato de que cada experiência armazenada existe em seu próprio quadro de

referência.

3. Método de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de que cada experiência armazenada representa pelo menos uma porção de um ambiente através do qual o veículo (102) se move, e que cada experiência armazenada refere-se a pelo menos um dos seguintes em relação às outras experiências armazenadas:

- (i) uma área espacial diferente dentro do ambiente;
- (ii) uma aparência visual diferente do ambiente; ou
- (iii) um tempo e/ou data de captura diferente.

4. Método de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que as experiências armazenadas estão arranjadas em subconjuntos segundo um ou mais dos seguintes parâmetros, e que um subconjunto de experiências armazenadas é selecionado para uso nas etapas (c) e (d) do método segundo um ou mais parâmetros;

- (i) momento do dia;
- (ii) dia da semana;
- (iii) estação;
- (iv) condições do tempo;
- (v) rota do veículo pretendida; e
- (vi) localização geográfica.

5. Método de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que pelo menos um dos seguintes aplica-se:

- (i) pelo menos algumas das experiências armazenadas são baixadas a partir de um servidor remoto (1604) para o veículo, em que a estimativa utiliza as experiências baixadas;
- (ii) pelo menos algumas das experiências armazenadas foram capturadas em uma determinada área espacial em momentos diferentes; e
- (iii) as experiências armazenadas fornecem uma representação composta de um ambiente através do qual o veículo (102) se move,

construída a partir de múltiplas experiências armazenadas que se superpõem.

6. Método de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que as experiências armazenadas são usadas por um ou ambos dos seguintes:

- (i) veículo (102) e/ou sensor (100) que gerou as mesmas; e
- (ii) veículos (102) e/ou sensores (100) em adição ao veículo e/ou sensor que gerou as mesmas.

7. Método de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que o método é arranjado para executar pelo menos uma das seguintes etapas:

(i) analisar as características extraídas e vincula características dentro de uma cena detectada no momento atual, com características em uma cena subsequente detectada em um segundo momento posterior usando uma transformação, e que opcionalmente o método é ainda arranjado para gerar (1504) uma trajetória do veículo (102) usando uma transformação que liga nós dentro das cenas detectadas;

(ii) requerer que a cena detectada seja reconhecida dentro de um número predeterminado de experiências armazenadas antes de ser determinado que o veículo (102) foi localizado; e

(iii) determinar se experiências armazenadas se referem a um local similar a outro e, caso deva ser determinado que experiências armazenadas se referem substancialmente ao mesmo local, ligar tais experiências armazenadas.

8. Método de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que pelo menos algumas das experiências armazenadas são associadas com uma localidade, e que o método usa posicionamento grosso do veículo (102) para determinar aproximadamente a localidade do veículo, e que opcionalmente o método usa a localidade do veículo para determinar quais experiências armazenadas devem ser

consideradas para determinar se uma cena detectada pode ser reconhecida dentro da mesma.

9. Método de acordo com qualquer reivindicação precedente, caracterizado pelo fato de que os dados capturados e as experiências armazenadas compreendem imagens, e que o método usa uma técnica de correspondência de imagem para determinar experiências armazenadas que parecem similares à cena detectada e usa subsequentemente as experiências armazenadas que parecem similares para determinar se uma cena detectada pode ser reconhecida dentro da mesma.

10. Sistema arranjado para localizar um veículo (102), o sistema caracterizado pelo fato de compreender um sensor (100) arranjado para gerar dados representando uma cena detectada em torno do veículo em um momento atual, e circuito de processamento (112) arranjado para processar os dados, em que o processamento é arranjado para:

extraír um conjunto de características a partir da cena detectada e estimar, utilizando as características extraídas, uma posição do veículo (102), em que a estimativa é realizada por:

comparar o conjunto de características extraídas da cena detectada com uma pluralidade de experiências armazenadas mantidas dentro de um dispositivo de armazenamento (128), para determinar se a cena detectada pode ser reconhecida dentro de qualquer das experiências armazenadas em que cada experiência armazenada compreende uma pluralidade de conjuntos de características, onde cada conjunto de características foi determinado a partir de uma cena detectada anteriormente e as experiências armazenadas cobrem uma área espacial ilimitada, além de capturar diferentes aparências da mesma área; e

se o conjunto de características extraídas da cena detectada for reconhecido dentro de qualquer experiência armazenada, então esta experiência armazenada é usada para produzir uma estimativa da posição do

veículo no momento atual com relação à experiência armazenada; e

se a cena detectada não for reconhecida dentro de mais do que um número predeterminado de experiências armazenadas, então a cena detectada torna-se parte de uma nova experiência armazenada para comparações futuras.

11. Sistema de acordo com a reivindicação 10, caracterizado pelo fato de que pelo menos um dos seguintes aplica-se:

(i) uma porção do dispositivo de armazenamento (128) é montada fora do veículo (102) e arranjada para ser acessível a partir do veículo;

(ii) o sistema ainda compreende um servidor (1604) arranjado para se comunicar com o circuito de processamento (112) no veículo (102); e

(iii) o circuito de processamento (112) é arranjado para comparar os dados contra um subconjunto das experiências armazenadas disponíveis, e que opcionalmente o circuito de processamento (112) é arranjado para obter, a partir de um dispositivo de armazenamento remoto (128), o subconjunto de experiências armazenadas disponíveis, que pode ser mantido localmente em um dispositivo de armazenamento (128) no veículo (102).

12. Veículo (102) tendo um sensor (100) montado no mesmo, caracterizado pelo fato de que o sensor é arranjado para gerar dados representando uma cena detectada em torno do veículo no momento atual, e um circuito de processamento arranjado (112) para processar os dados, em que o circuito de processamento é arranjado para:

extrair um conjunto de características a partir da cena detectada e estimar, utilizando as características extraídas, uma posição do veículo (102), em que a estimativa é realizada por:

comparar o conjunto de características extraídas da cena detectada com uma pluralidade de experiências armazenadas mantidas dentro

de um dispositivo de armazenamento (128), para determinar se a cena detectada pode ser reconhecida dentro de qualquer das experiências armazenadas em que cada experiência armazenada compreende uma pluralidade de conjuntos de características, onde cada conjunto de características foi determinado a partir de uma cena detectada anteriormente e as experiências armazenadas cobrem uma área espacial ilimitada, além de capturar diferentes aparências da mesma área; e

se o conjunto de características extraídas da cena detectada for reconhecido dentro de qualquer experiência armazenada, então esta cena armazenada é utilizada para produzir uma estimativa da posição do veículo (102) no momento atual com relação à experiência armazenada; e

se a cena detectada não for reconhecida dentro de mais de um número predeterminado de experiências armazenadas, então a cena detectada torna-se parte de uma nova experiência armazenada para comparações futuras.

13. Servidor (1604) arranjado para receber um pedido de rede (1602) a partir de uma rede para uma pluralidade de experiências armazenadas, o servidor caracterizado pelo fato de ser arranjado para processar o pedido, para recuperar as experiências armazenadas pedidas a partir de um dispositivo de armazenamento (128) e para transmitir as experiências armazenadas pedidas através da rede (1602).

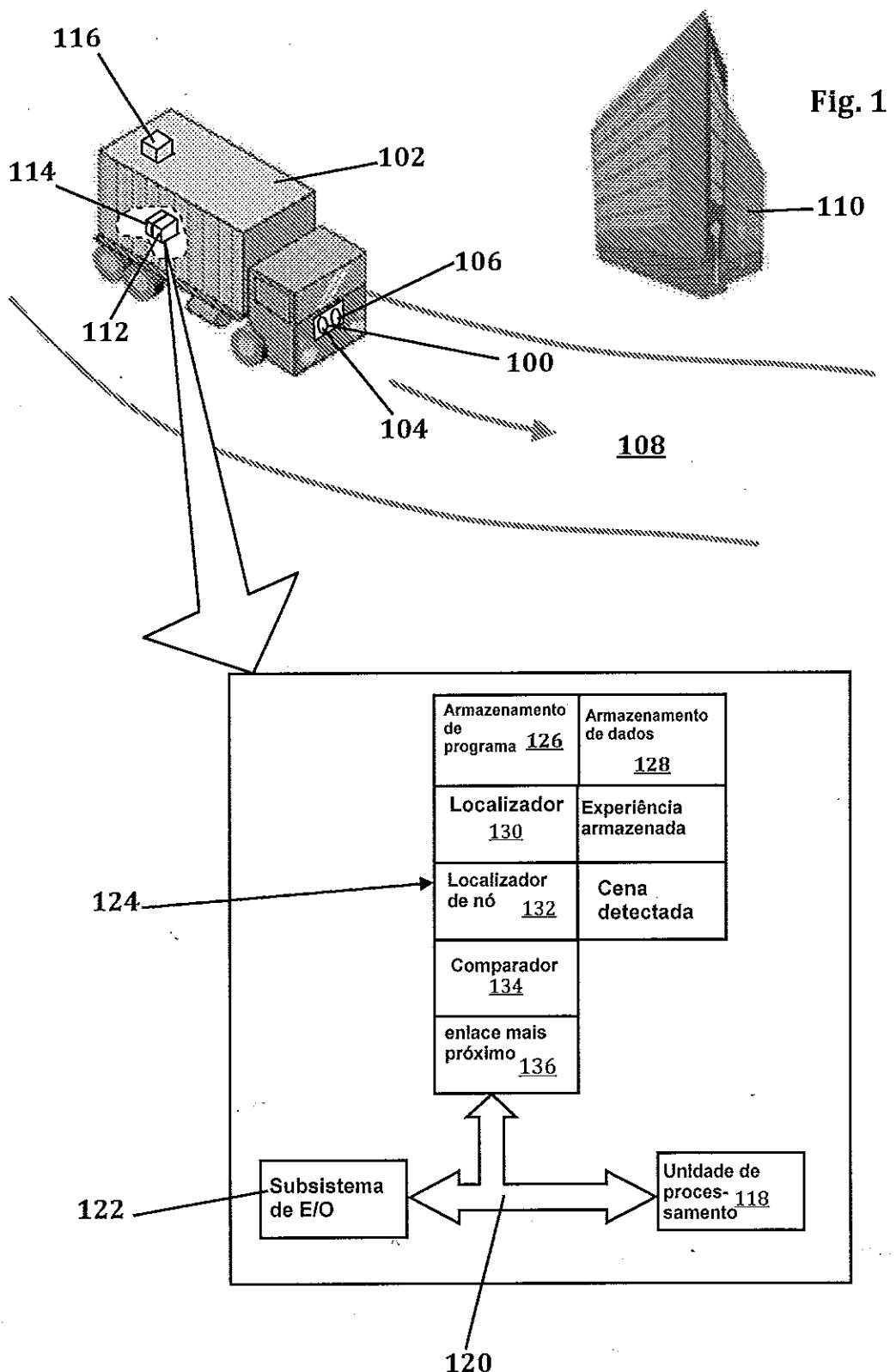
14. Meio legível por máquina contendo instruções, caracterizado pelo fato de que, quando lidas por uma máquina, fazem com que esta máquina execute pelo menos um dos seguintes:

(i) o método como definido em qualquer uma das reivindicações 1 a 9;

(ii) pelo menos uma porção do sistema como definido na reivindicação 10 ou 11;

(iii) o veículo como definido na reivindicação 12; e

(iv) o servidor como definido na reivindicação 13.



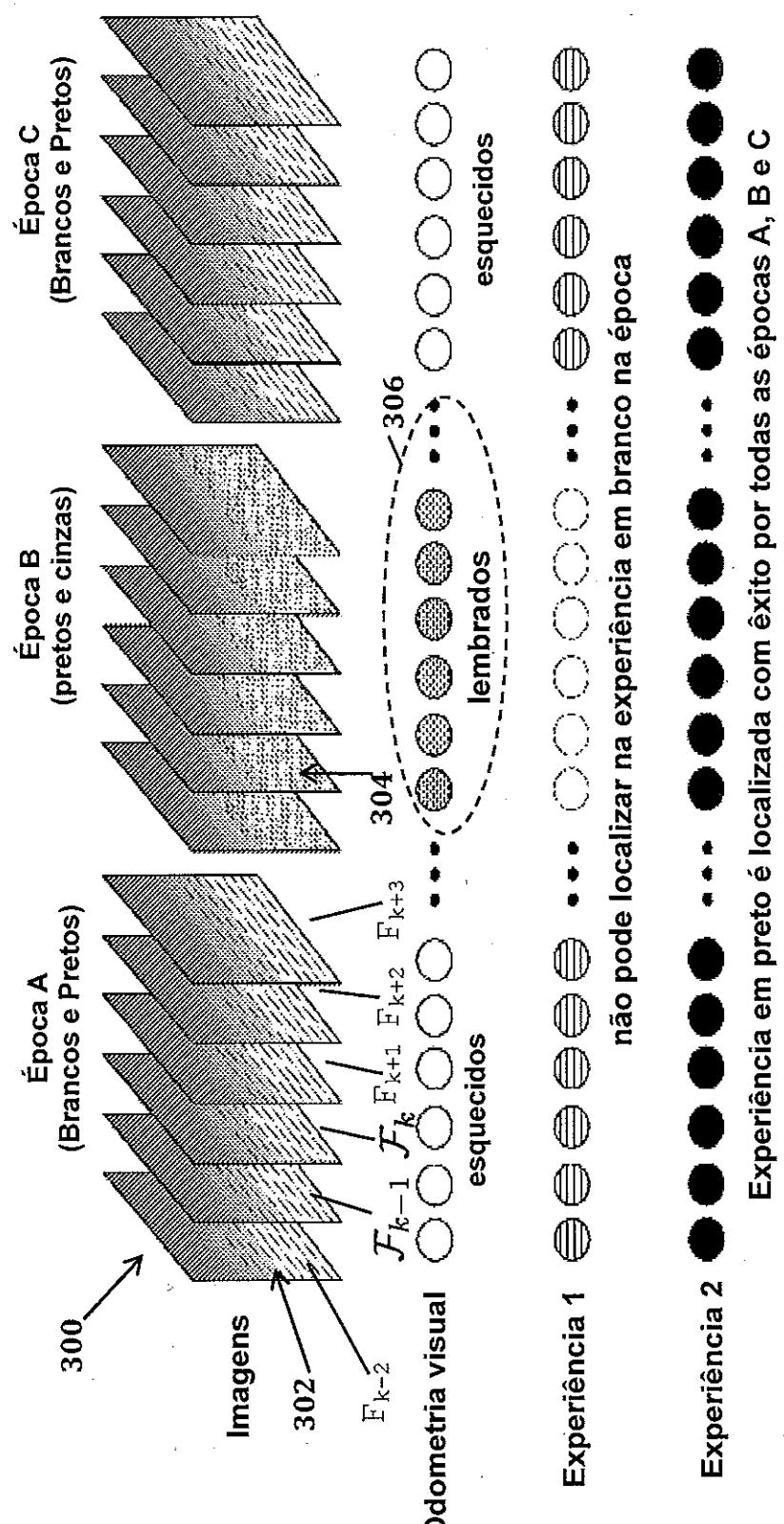


Fig. 3

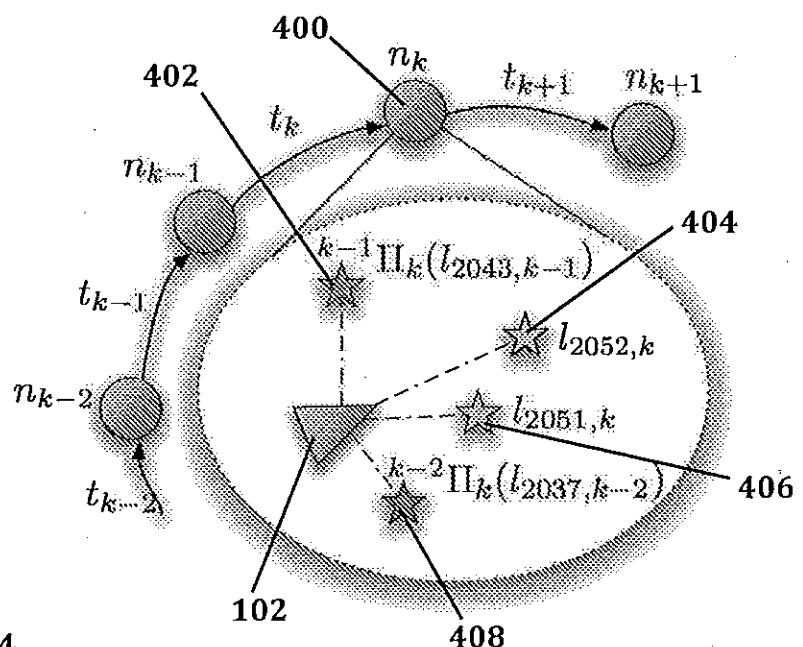


Fig. 4

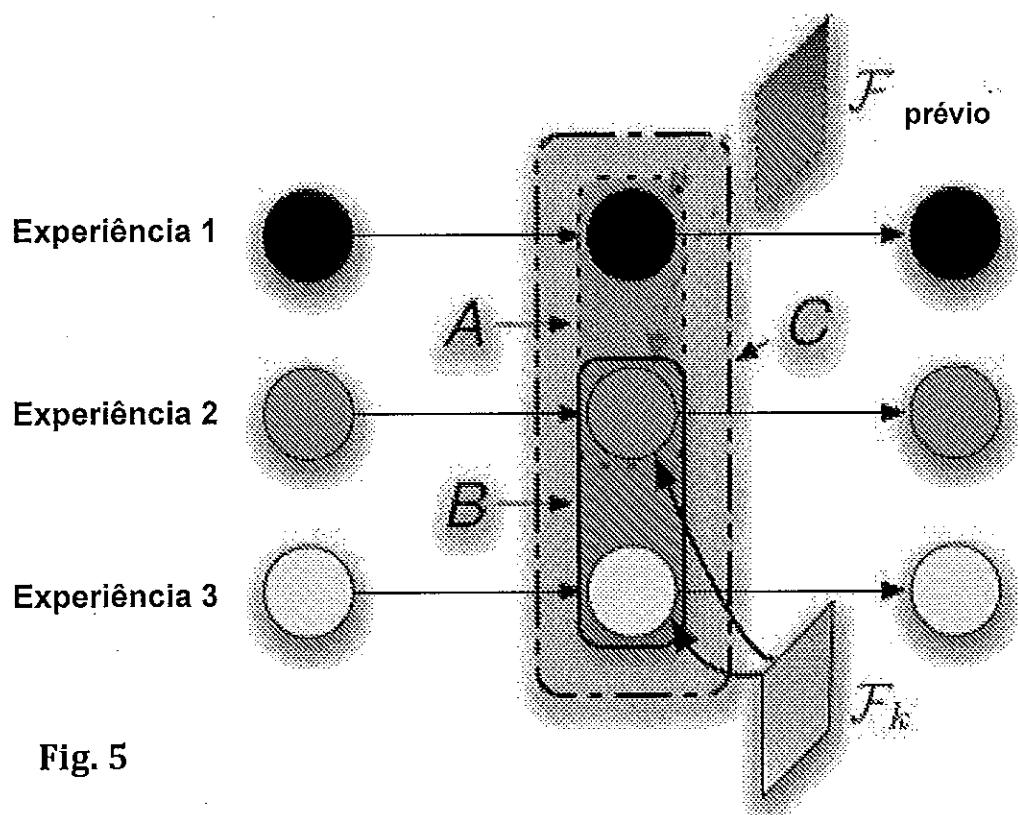
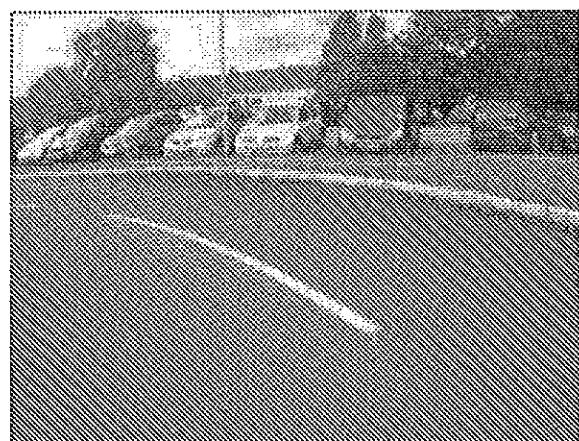
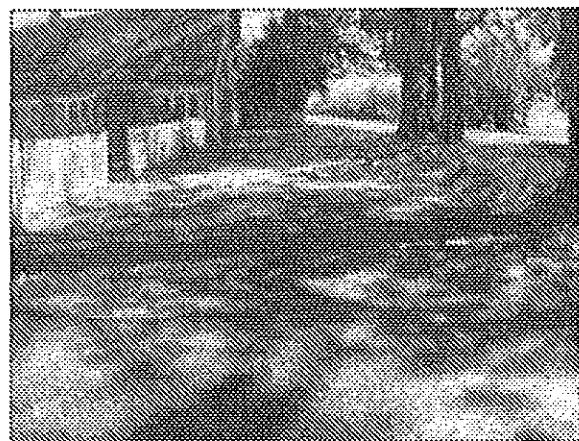


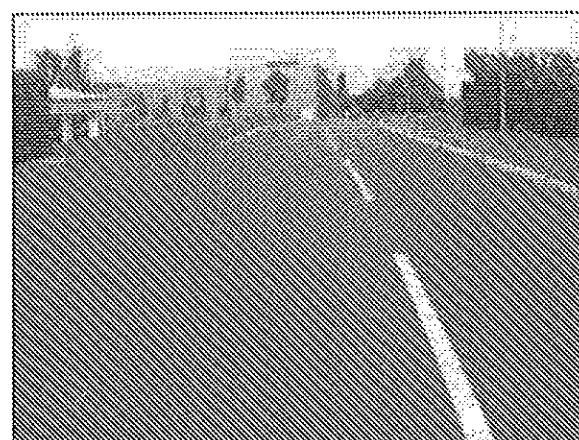
Fig. 5



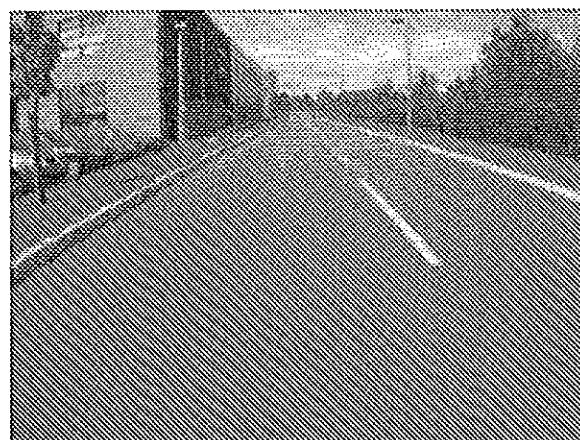
(a)



(b)



(c)



(d)

Fig. 6

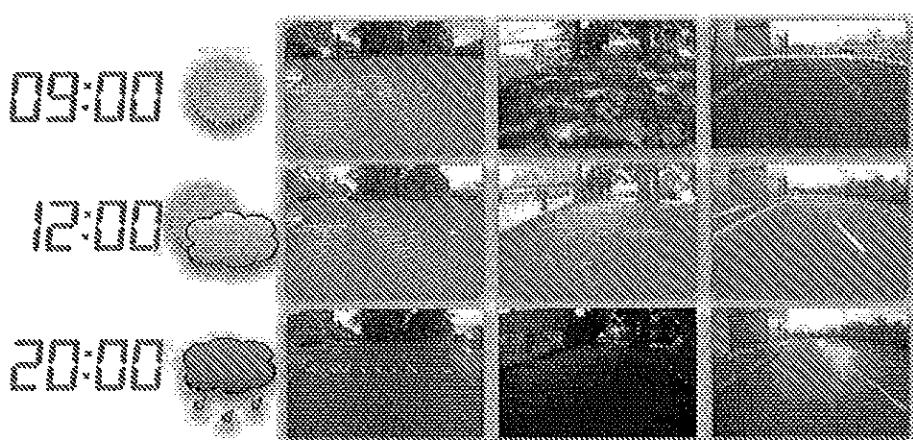


Fig. 2

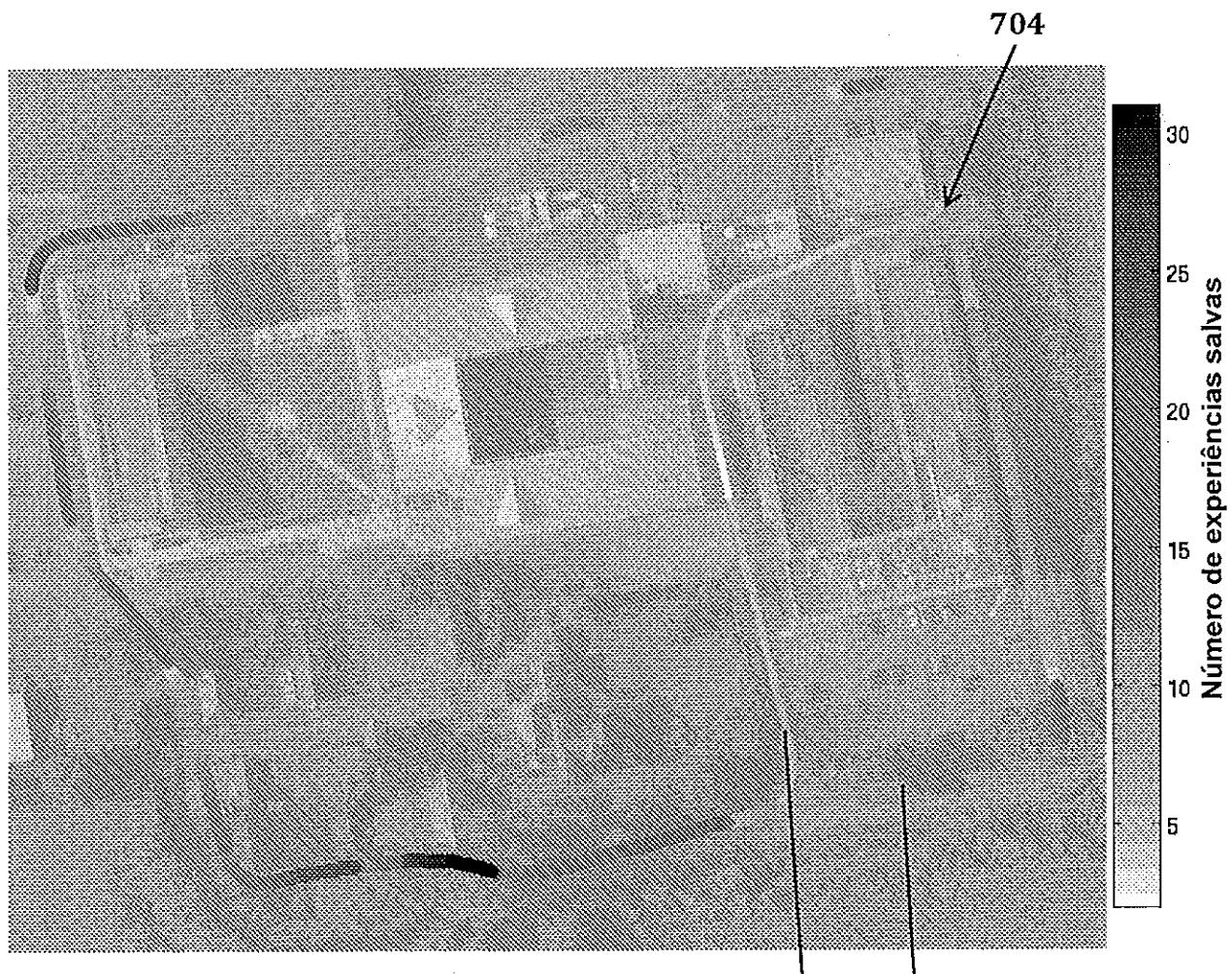


Fig. 7

Número de visita vs Saída de VO Normalizada, acrescentada ao Mapa Plástico

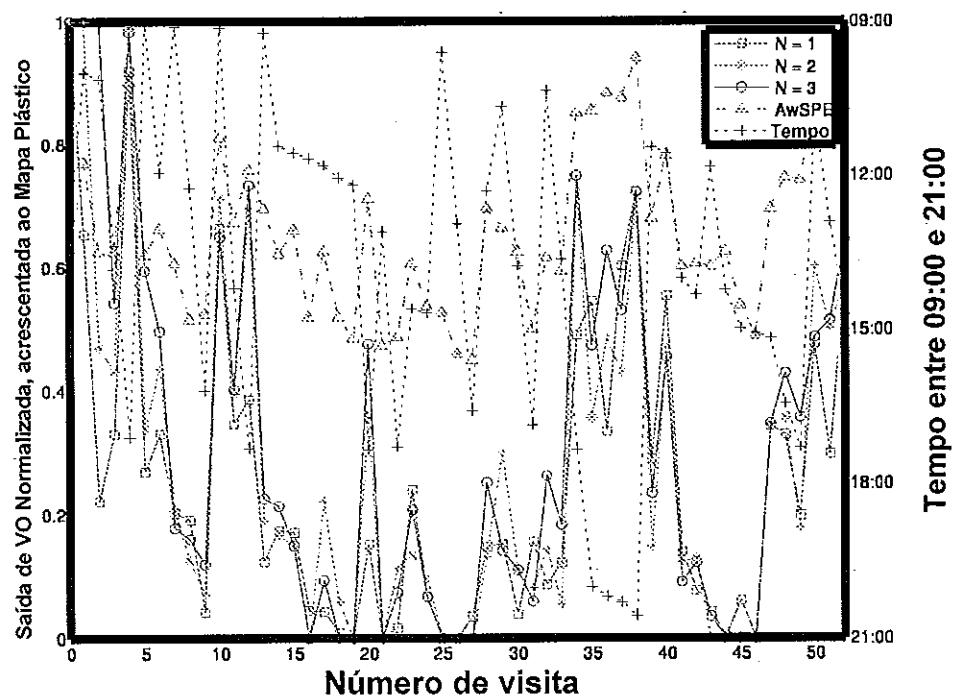


Fig. 8

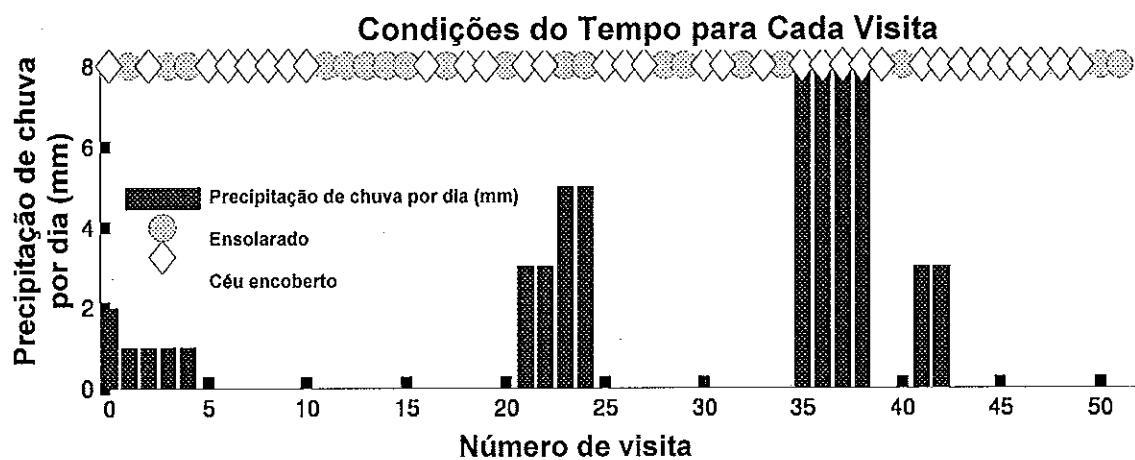


Fig. 9

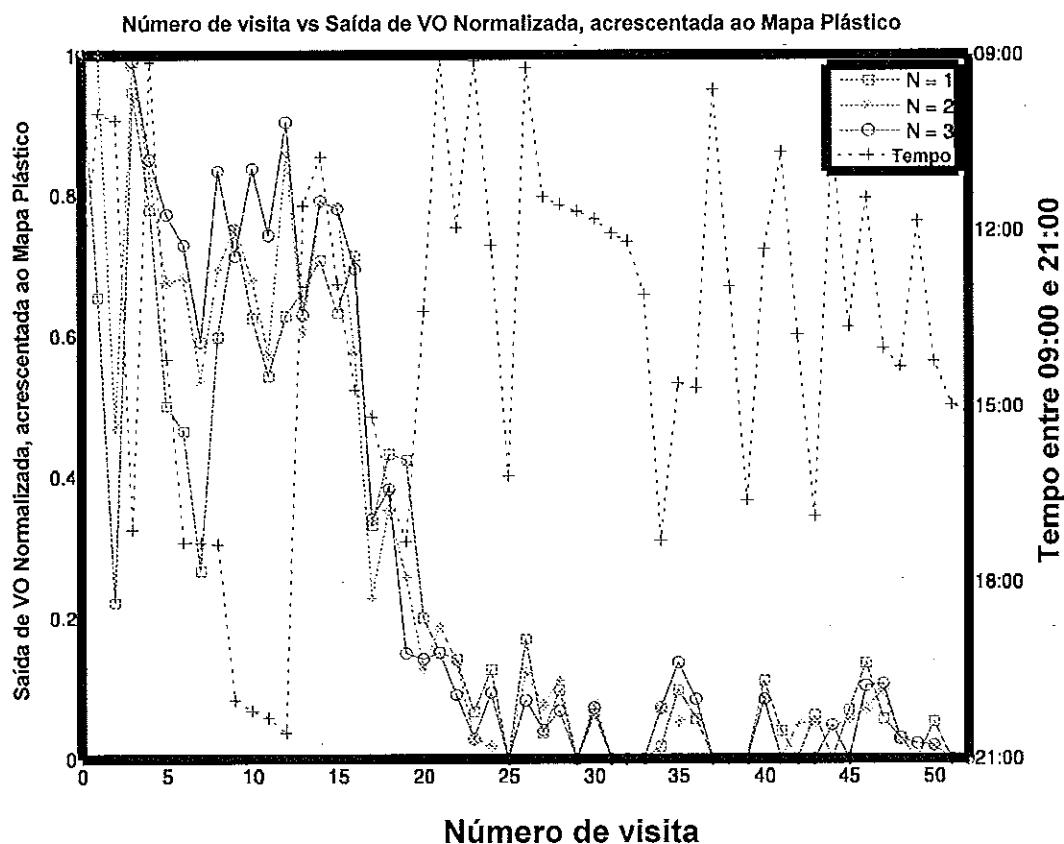


Fig. 10

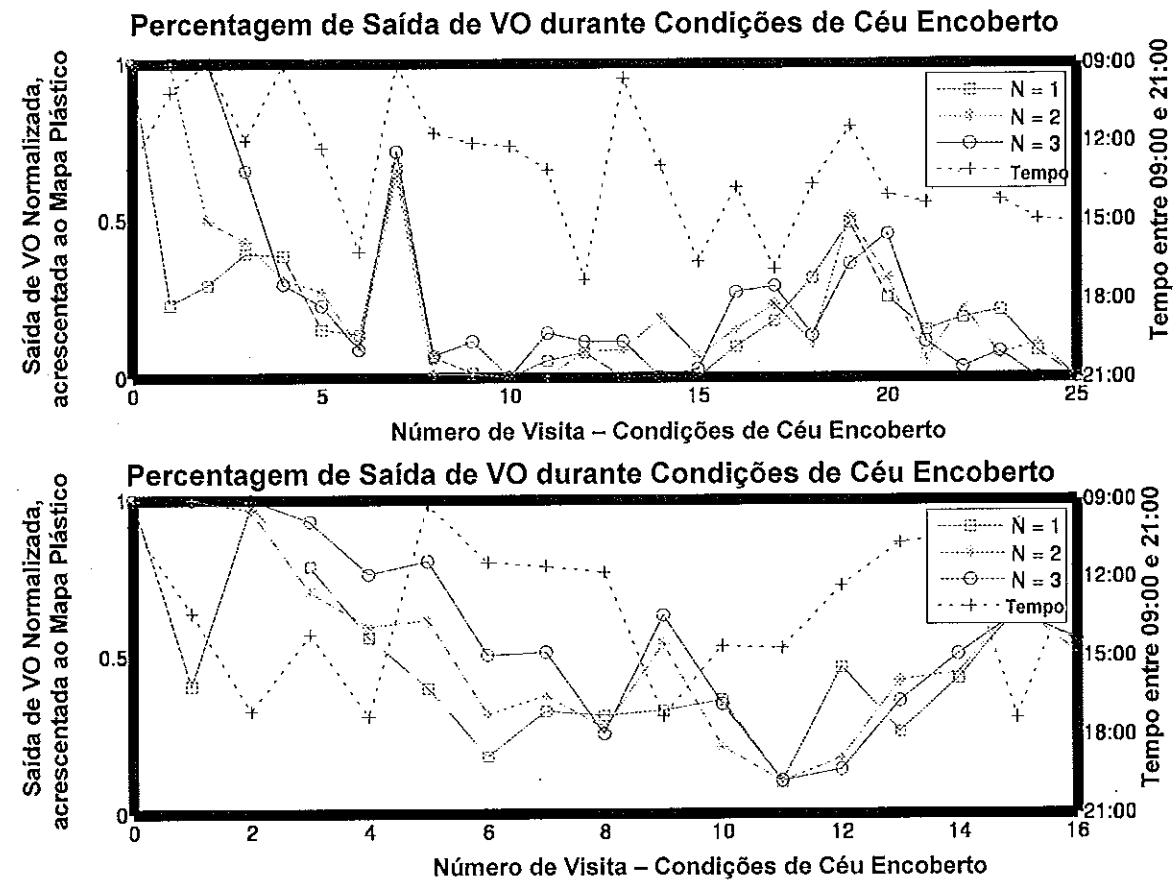
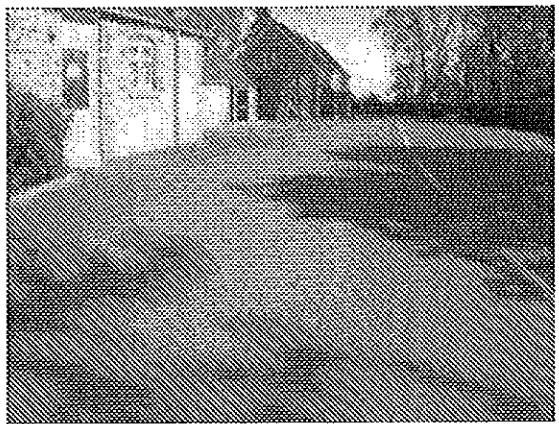


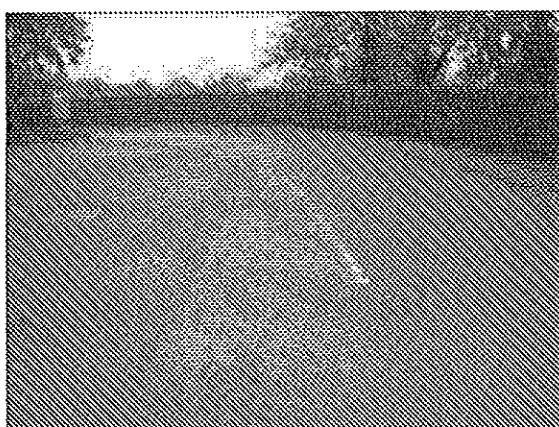
Fig. 11



(a)



(b)

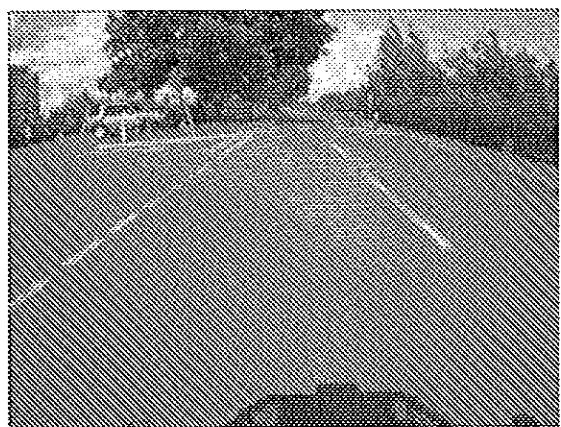


(c)

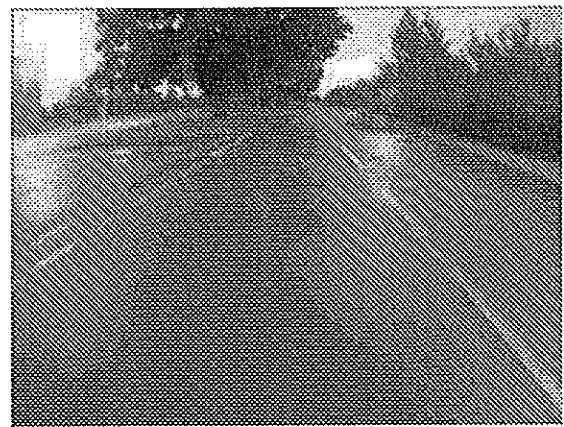


(d)

**Fig. 12**



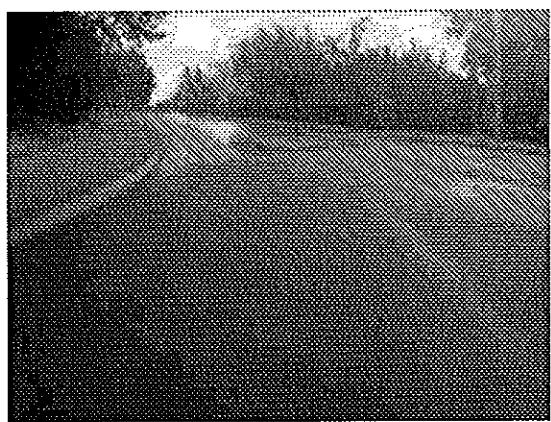
(a)



(b)

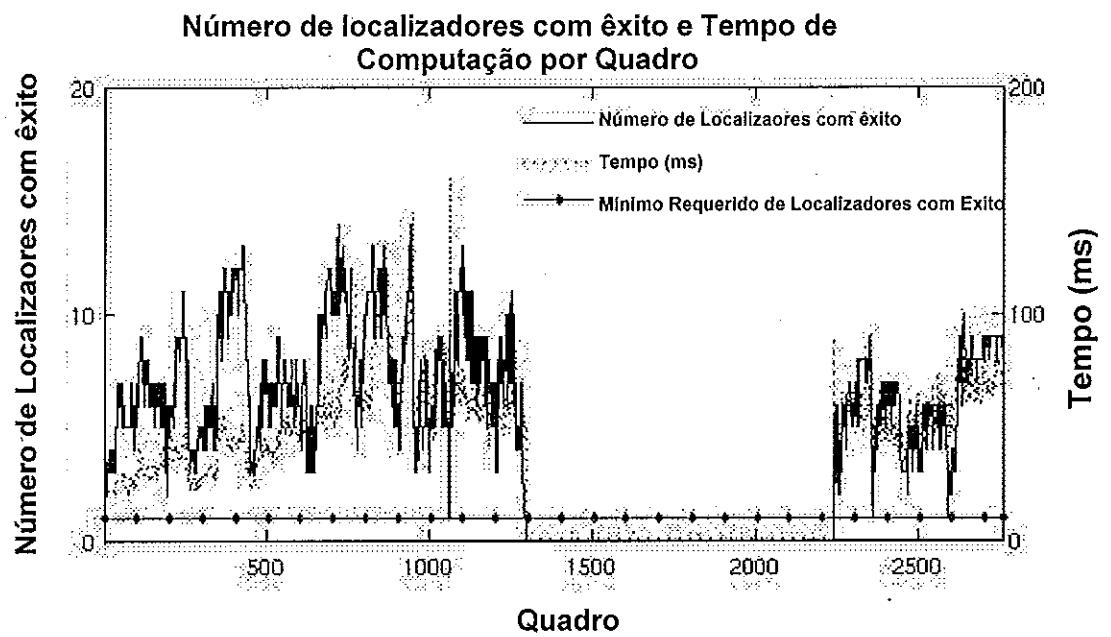


(c)



(d)

**Fig. 13**



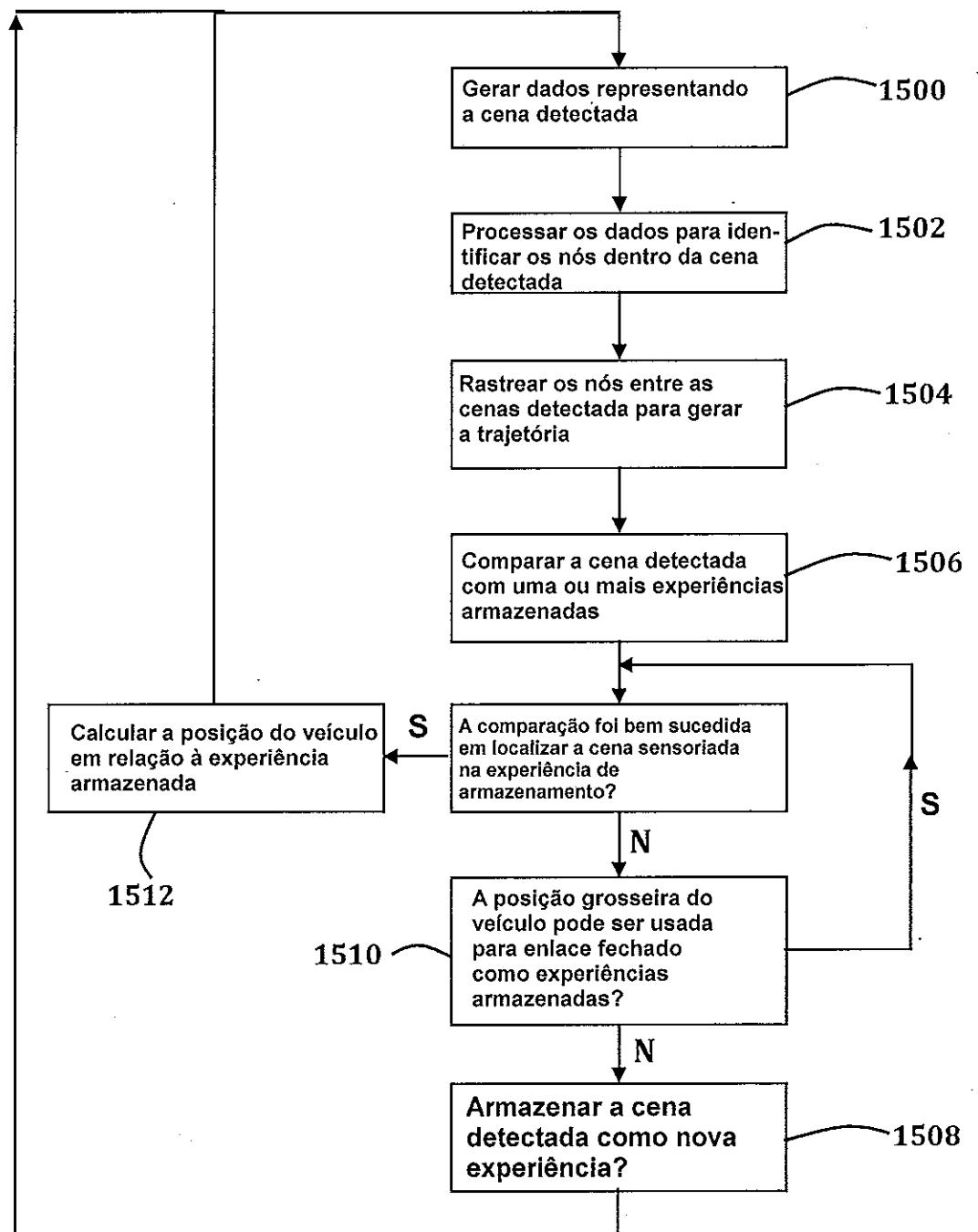


Fig. 15

