

UNIVERSIDADE FEDERAL DE MINAS GERAIS
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO
CURSO DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Observação Dinâmica Cooperativa

José Luiz de Souza Pio

BELO HORIZONTE, MINAS GERAIS
Fevereiro de 2006

UNIVERSIDADE FEDERAL DE MINAS GERAIS
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO
CURSO DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Observação Dinâmica Cooperativa

José Luiz de Souza Pio

Tese apresentada ao Programa de Pós-Graduação em
Ciência Computação da Universidade Federal de Mi-
nas Gerais como requisito parcial para obtenção do
título de Doutor em Ciência da Computação.

Orientador: Mário Fernando Montenegro Campos

BELO HORIZONTE, MINAS GERAIS
Fevereiro de 2006

FEDERAL UNIVERSITY OF MINAS GERAIS
COMPUTER SCIENCE DEPARTMENT
GRADUATE PROGRAM IN COMPUTER SCIENCE

Cooperative Dynamic Observation

José Luiz de Souza Pio

Thesis presented to the Graduate Program in
Computer Science of the Federal University of Mi-
nas Gerais in partial fulfillment of the requirements
for the degree of Doctor in Computer Science.

Advisor: Mario Fernando Montenegro Campos

BELO HORIZONTE, MINAS GERAIS
February, 2006

Aos meus Pais e à
minha Sylvie.

Resumo

Nos dias atuais, devido ao aumento dos conflitos internacionais e da violência urbana, é cada vez mais comum a utilização de sistemas de segurança projetados como uma rede de câmeras com o objetivo de monitorar ou controlar o fluxo de pessoas ou veículos. Nesta tese aborda-se o problema da *observação do movimento de objetos de interesse em um espaço monitorado por sensores visuais*. Este problema é aqui denominado de “Observação Dinâmica Cooperativa”. O objetivo principal é a obtenção da informação visual para o posicionamento das câmeras visando a observação da trajetória de objetos de interesse que se movimentam no ambiente monitorado. Nesse contexto, a *observação* compreende a identificação e o rastreamento dos alvos e suas trajetórias; a *dinâmica* refere-se à capacidade de mobilidade das câmeras, provida a estas por meio de robôs móveis; e a *cooperação* caracteriza-se pelo posicionamento das câmeras promovido pelo planejamento e troca de informações entre estas por meio de um sistema de comunicação em rede.

É proposta uma metodologia cuja idéia principal é a identificação dos objetos e de suas trajetórias, levando ao conseqüente posicionamento das câmeras por meio de um esquema de planejamento de posição e cooperação entre os membros do sistema. A metodologia está dividida em três módulos principais: o rastreamento de alvos, o planejamento de posição e o reconhecimento e associação entre informações visuais obtidas por diferentes câmeras. O rastreamento é realizado por meio de filtro de partículas distribuído que associa, em tempo real, a informação visual da cor dos alvos com os seus movimentos. O planejamento de posição das câmeras é executado por meio da avaliação de uma função de observação sujeita a imposição de restrições visuais e do ambiente. O reconhecimento e associação entre objetos e trajetórias observados por câmeras distintas é conduzido por meio de uma adaptação do Algoritmo Expectativa–Maximização. A metodologia é avaliada sob o aspecto tecnológico e de cenário. Um conjunto de experimentos reais é sistematicamente realizado, garantindo robustez suficiente para a adaptação e implementação da metodologia em sistemas de segurança com visão distribuída.

Abstract

Nowadays, surveillance and security systems based on visual sensors are a very common approach. International terrorism and the growth of urban violence, evoke the new applications of computer vision systems. These systems can be found in international ports, airports, train and subway stations of all great urban centers. A common approach is based on analog closed-circuit television systems with image scene analysis, control and decision centered on human operator. However, modern and sophisticated surveillance based computer vision systems enable the integration of sights from many cameras into a single, consistent scene representation.

This thesis addresses the problem of multi-camera target observation. Here, we name this problem of “Cooperative Dynamic Observation”. We want to find the cameras’ pose that makes possible the observation of moving targets of interest and their trajectories based on visual information shared between cameras. In this problem, we consider that observation means target identification and tracking; dynamic is the cameras’ moving ability provided by mobile robots with navigation and positioning performance; and cooperation refers to the practice of self-organized camera positioning based on shared visual information, provided by a communication network. To address this problem, we developed a framework that finds the cameras’ poses based on visual information acquired and shared by a camera communication network.

The framework is modelled in three principal modules: the tracker, the camera position planner, and the target/trajectory association module. The tracker is based on distributed particle filter that fuses, in real time, the targets’ motion and visual information given by colors. The position planning is done by a observation function with optical and environmental constraints. The targets’ trajectories between cameras with disjoint field of views are matched by an adaptation of classical EM (Expectation-Maximization) Algorithm. The robustness of framework is analyzed and tested in real experiments with a developed systematic experimental protocol.

Agradecimentos

Agradeço ao Senhor meu Deus por ter me conduzido e amparado em todos os momentos da minha vida. Agradeço ainda ao meu pai Azamor Pio pelo apoio e incentivo e a minha mãe Ana Maria por suportar a minha ausência durante esses últimos anos.

O meu doutoramento teve a orientação e o apoio do Professor Mario Fernando Montenegro Campos. O Professor Mario é a quem devo a motivação de deixar Manaus e partir para Belo Horizonte para cursar o doutorado na UFMG.

Durante esse período, o grupo de professores do DCC/UFMNG contribuiu para a minha formação. A estes, agradeço a dedicação e a amizade, em especial aos professores Arnaldo Albuquerque, Claudionor Coelho, Newton Vieira, Antonio Alfredo, José Marcos, Antonio Otávio, Nívio Ziviani, Rodrigo Carceroni, Rodolfo, Alberto Laender, Pacca e Geraldo Robson.

Um agradecimento especial vai para os membros do VeRLab, passados e atuais. Em especial, agradeço aos professores Guilherme Pereira e Luiz Chaimowikz pelo apoio e colaboração sempre presente. Agradeço ainda ao Pedro Shiroma, Rafael, Denilson, Celina, Piti, Daniele, O(s) Eric(k)son(s), Armando, Flávio Cardeal e toda a turma do PAD que passou pelo VeRLab nesse últimos cinco anos. Um agradecimento especial também é extensível ao André Ruas, que ajudou bastante na implementação deste trabalho.

Durante o período de doutoramento tive a oportunidade de fazer amigos para a vida inteira. A estes dedico um agradecimento mais do que especial: a Karla Albuquerque, a Andréa Iabrudi, a Bia e o Camillo. Claro, não posso deixar de citar e de agradecer aos colegas de turma, em particular ao Silvio, Paulo, Lula, Kissia, Silvana, Romaneli Zuim, Júlio, Fernando Sica, Lilian, Joice, Sergião, Humberto e Umberto, Pedrão, Luiz Henrique, Guilherme e Nasif.

A um grupo peculiar de amigos agradeço pelo companheirismo, atenção e pelos laços que se estreitaram aqui na UFMG e se prolongarão até a UFAM. Trata-se da Comunidade Amazonense Residente em BH (CARBH), composta pelo amigo, compadre e colega de trabalho Ruiter Braga Caldas, José Pinheiro, Eduardo e Fábíola Nakamura, Maurício

Figueiredo, Vilarzinho, Marcos Cristo, Josenete, Geraldo e Max (da estatística), Horácio, David e Moisés.

Agradeço às colegas da secretaria Renata, Emília, Túlia, Sheila, Sônia, Gilmara, Luciana, Maristela, Cida, Ilsete e aos da biblioteca, a Belkiz e o Helvécio. Também aos meus amigos da cantina Ana Mara e o Sr. Maurício.

Agradeço a minha esposa Sylvie, pelo carinho, compreensão e por todo o amor que sempre me dedicou.

A CAPES pelo apoio financeiro durante os quatro primeiros anos de trabalho e a UFAM, em particular a todos os colegas do DCC, pelo consideração e amizade. A todos muito obrigado!

Sumário

| | |
|--|-----------|
| Dedicatória | i |
| Resumo | ii |
| Abstract | iii |
| Agradecimentos | iv |
| 1 Introdução | 1 |
| 1.1 Motivação | 4 |
| 1.2 A Abordagem Desenvolvida | 5 |
| 1.3 Contribuições | 9 |
| 1.4 Organização do Texto | 10 |
| 2 Trabalhos Relacionados | 11 |
| 2.1 Rastreamento Visual de Múltiplos Alvos | 11 |
| 2.1.1 Sistemas com Múltiplas Câmeras | 14 |
| 2.1.2 Sistemas Multi-Câmeras Cooperativos | 23 |
| 2.2 Rastreamento com Robôs | 24 |
| 2.3 Planejamento de Sensores | 28 |
| 3 Metodologia | 31 |
| 3.1 Formalização do Problema | 31 |
| 3.2 Suposições | 33 |
| 3.3 A Metodologia | 34 |
| 3.3.1 O Rastreamento Visual | 35 |

| | | |
|----------|--|------------|
| 3.3.2 | Associação de Trajetórias em Campos de Visão Disjuntas | 46 |
| 3.3.3 | Planejamento do Posicionamento das Câmeras | 52 |
| 3.3.4 | Aplicação com o G_{vis} | 71 |
| 4 | Aplicações Experimentais | 74 |
| 4.1 | Protocolo de Experimentação | 74 |
| 4.1.1 | Projeto de Avaliação Experimental | 75 |
| 4.1.2 | O Plano de Análise de Dados | 76 |
| 4.2 | Avaliação Experimental do Método de Rastreamento | 79 |
| 4.2.1 | Rastreamento e predição do movimento dos alvos | 79 |
| 4.2.2 | Avaliação do Rastreamento Simultâneo de Múltiplos Alvos | 91 |
| 4.2.3 | Discussão | 95 |
| 4.3 | Avaliação do Método de Associação entre Medidas e Trajetórias em uma Rede de Câmeras com Campos de Visão Disjuntas | 96 |
| 4.3.1 | Montagem do Aparato Experimental | 96 |
| 4.3.2 | Experimentação | 97 |
| 4.3.3 | Discussão | 103 |
| 4.4 | Avaliação do Método de Posicionamento | 103 |
| 4.4.1 | Posicionamento do Robô em Relação à Trajetória | 104 |
| 4.5 | Considerações | 106 |
| 5 | Conclusão e Trabalhos Futuros | 109 |
| 5.1 | Limitações | 110 |
| 5.2 | Extensões | 111 |
| 5.3 | Trabalhos Futuros | 112 |
| 5.3.1 | Observação Cooperativa Tridimensional | 113 |
| 5.3.2 | Visão Distribuída em Rede de Sensores Visuais sem Fio | 114 |
| 5.4 | Considerações Finais | 114 |
| A | Métodos Seqüenciais de Monte Carlo | 116 |

| | |
|---|------------|
| B O Rastreamento Visual de Objetos | 123 |
| B.1 Características Visuais | 124 |
| B.1.1 A Cor dos Alvos como Característica Visual | 125 |
| B.2 Detecção do Movimento | 127 |
| C Aspectos Geométricos do Posicionamento de Sensores | 135 |

Lista de Figuras

| | | |
|-----|--|----|
| 1.1 | Sistema de câmeras de controle de acesso ao Campus. As imagens são capturadas pelo conjunto de câmeras e armazenadas em uma central de segurança. Todo o processo de decisão é centralizado no operador humano. | 2 |
| 1.2 | Caminhos, vias e faixas de tráfego formam direções preferenciais para a trajetória. Nota-se nas fotografias que pessoas e veículos tendem em seguir a trajetória definida pelas vias. | 3 |
| 1.3 | Sistema idealizado de câmeras de segurança nas vias de acesso do campus da Universidade Federal de Minas Gerais. As câmeras estão distribuídas pelas principais vias e nas áreas de estacionamento. O sistema extrai as características dos veículos e distribui por todas as câmeras. As câmeras podem correlacionar as características e identificar a chegada dos veículos nas áreas de estacionamento. | 6 |
| 2.1 | Principais áreas envolvidas no trabalho e algumas referências em destaque. | 12 |
| 2.2 | Esquema do sistema de câmeras pan/tilt distribuídas [Woo e Capson (2000)]. As câmeras possuem movimento de rotação em relação ao eixo vertical e capacidade de aproximação visual por meio das lentes zoom. Sinais de vídeo e de controle são coordenados por um microcontrolador associado a uma CPU. | 16 |

| | | |
|-----|--|----|
| 2.3 | Esquema do modelo lógico das tarefas do sistema de segurança de Collins et al. (2000). O sistema é dividido em duas estruturas lógicas denominadas de espaço da imagem e espaço de serviço. No espaço da imagem, tarefas como a detecção do movimento, rastreamento e classificação são realizadas. No espaço de serviço realizam-se tarefas como a fusão de dados sensoriais e adequação das informações a um modelo do ambiente. | 23 |
| 3.1 | Arquitetura funcional do sistema de observação dinâmica. As caixas representam os principais módulos. As setas indicam o fluxo de informações entre os módulos do sistema e a área de monitoramento. As câmeras C_i estão representadas pelo círculo posicionado na parte inferior a esquerda da figura. | 35 |
| 3.2 | Diagrama de blocos do sistema de rastreamento visual. O sistema é decomposto funcionalmente em três módulos principais: processamento de imagens, rastreamento e fusão de dados. | 37 |
| 3.3 | Diagrama funcional do subsistema de associação de trajetórias. O sistema recebe de uma câmera C_i informações sobre o movimento e da trajetórias de alvos que devem ser associados a movimentos e trajetórias de alvos vistos por outra câmera C_j do sistema. | 48 |
| 3.4 | O processo de mapeamento. Um ponto com características no plano de imagem está associado a um alvo no sistema referência global. | 49 |
| 3.5 | As coordenadas $C(x, y)$ e o ângulo θ definem o ponto de visada da câmera. | 52 |
| 3.6 | O vetor \mathbf{v}_p é formado entre a posição do robô e o ponto de referência. O ângulo de visada é formado por \mathbf{v}_p e \mathbf{n} | 54 |
| 3.7 | O encurtamento do alvo no espaço da imagem. Dependendo da posição da câmera, o alvo pode ser encurtado em uma de suas dimensões devido a transformação perspectiva. | 55 |
| 3.8 | Parâmetros que influenciam a capacidade de observação: distância entre a trajetória e a câmera e os ângulos de visada α e de rastreamento ϕ | 56 |

| | | |
|------|--|----|
| 3.9 | Cobertura da câmera. A altura da câmera e tamanho do ângulo de visada determinam a área de cobertura da câmera. | 57 |
| 3.10 | A distância entre a trajetória e a câmera e os ângulos de visada α e de rastreamento ϕ | 58 |
| 3.11 | Determinação dos pontos base. São pontos base os vértices contidos na região de cobertura definida pela rotação da câmera e visíveis a partir de uma vértice de referência (ponto mais escuro). | 60 |
| 3.12 | Formação das arestas do grafo de visibilidade restrita. As arestas são formadas por todo ponto base visível pelo vértice de referência. | 63 |
| 3.13 | Formação de um grafo de visibilidade restrita $G_{vis}(E, V)$. A partir de um ponto de referência inicial, como mostrado no primeiro desenho no canto superior esquerdo da figura, o algoritmo determina os pontos base e inclui arestas entre estes e os pontos determinados. O processo se repete para todos os vértices de todos os polígonos. | 65 |
| 3.14 | Determinação de pontos base em regiões com grande comprimento. Pontos podem ser incluídos para fornecer posições de visibilidade para as câmeras. | 66 |
| 3.15 | O custo da visualização está associado aos vértices do grafo e as suas arestas. O custos dos vértices leva em conta os as rotações dos ângulos de rastreamento ϕ_i e ao ângulo de visada α . Em relação as arestas, está associado o custo relativo às distâncias. | 69 |
| 3.16 | Exemplo de um grafo de visibilidade. Os vértices 1 e 2 (mais escuros) representam os pontos de observação sobre a trajetória. | 72 |
| 4.1 | Metodologia de experimentação. A avaliação tecnológica é executada, separadamente, sobre cada um dos principais módulos do sistema. A avaliação de cenário usa um ou mais módulos em conjunto. | 76 |
| 4.2 | Cálculo da distância média entre as trajetórias T_1 e T_2 . As duas trajetórias estão separadas em pontos comuns tomados ao mesmo instante de tempo. | 77 |

| | | |
|-----|--|----|
| 4.3 | Cálculo do deslocamento horizontal e vertical entre as trajetórias T_1 e T_2 . Os pontos em comum são obtidos no mesmo instante de tempo. | 78 |
| 4.4 | Um trilho de trem de brinquedo é montado sobre uma base fixa, servindo como a trilha de referência. O alvo é envolto em uma capa plástica de cor uniforme. | 80 |
| 4.5 | Montagem para a avaliação da precisão do método de rastreamento. A câmera está ortogonal ao plano da mesa, posicionada a uma altura de $1,50m$ em relação a esse plano. A mesma altura encontra-se posicionada a fonte de luz incandescente. | 81 |
| 4.6 | A segmentação do alvo e o posicionamento deste sobre a trilha de referência fornecem as medidas para a estimativa dos parâmetros da trajetória no espaço da imagem. As medidas são obtidas a partir do valor central do blob. | 82 |
| 4.7 | Círculo de melhor ajuste e pontos medidos no espaço da imagem. | 83 |
| 4.8 | Valores estimados e medidos do raio e o erro em relação ao círculo de referência pelos modelos de predição e de medidas com 50 partículas. Os gráficos da coluna da direita da figura mostram a variação do valor do raio no modelo de predição (na parte de cima) e no modelo de medida ou observação (em baixo). Os gráficos da coluna da esquerda mostram o erro absoluto no modelo de predição (na parte de cima) e no modelo de medida (em baixo). | 85 |
| 4.9 | Valores estimados e medidos do raio e o erro em relação ao círculo de referência pelos modelos de predição e de medidas com 100 partículas. Os gráficos da coluna da direita da figura mostram a variação do valor do raio no modelo de predição (na parte de cima) e no modelo de medida ou observação (em baixo). Os gráficos da coluna da esquerda mostram o erro absoluto no modelo de predição (na parte de cima) e no modelo de medida (em baixo). | 85 |

- 4.10 Valores estimados e medidos do raio e o erro em relação ao círculo de referência pelos modelos de predição e de medidas com 200 partículas. Os gráficos da coluna da direita da figura mostram a variação do valor do raio no modelo de predição (na parte de cima) e no modelo de medida ou observação (em baixo). Os gráficos da coluna da esquerda mostram o erro absoluto no modelo de predição (na parte de cima) e no modelo de medida (em baixo). 86
- 4.11 Valores estimados e medidos do raio e o erro em relação ao círculo de referência pelos modelos de predição e de medidas com 500 partículas. Os gráficos da coluna da direita da figura mostram a variação do valor do raio no modelo de predição (na parte de cima) e no modelo de medida ou observação (em baixo). Os gráficos da coluna da esquerda mostram o erro absoluto no modelo de predição (na parte de cima) e no modelo de medida (em baixo). 86
- 4.12 Valores estimados e medidos do raio e o erro em relação ao círculo de referência pelos modelos de predição e de medidas com 1000 partículas. Os gráficos da coluna da direita da figura mostram a variação do valor do raio no modelo de predição (na parte de cima) e no modelo de medida ou observação (em baixo). Os gráficos da coluna da esquerda mostram o erro absoluto no modelo de predição (na parte de cima) e no modelo de medida (em baixo). 87
- 4.13 Valores estimados e medidos do raio e o erro em relação ao círculo de referência pelos modelos de predição e de medidas com 2000 partículas. Os gráficos da coluna da direita da figura mostram a variação do valor do raio no modelo de predição (na parte de cima) e no modelo de medida ou observação (em baixo). Os gráficos da coluna da esquerda mostram o erro absoluto no modelo de predição (na parte de cima) e no modelo de medida (em baixo). 87

| | | |
|------|--|-----|
| 4.14 | Variação do valor do erro médio na estimativa do raio. O erro decai com o aumento do número de partículas. | 89 |
| 4.15 | Erro na translação horizontal e vertical nas medidas de predição e observação do modelo de rastreamento para diferentes números de partículas geradas. As altas frequências tendem a se manter mais próximas da média com o aumento do número de partículas. | 90 |
| 4.16 | Erro médio nas medidas horizontais (a esquerda) e verticais (a direita). A maior variação do erro nas medidas verticais é devido a maior dimensão do pixel nesta direção. | 92 |
| 4.17 | Rastreamento de múltiplos alvos. O quadro 1 mostra a entrada de um objeto em cena. Os outros quadros mostram a propagação das partículas sobre quatro objetos, rastreados simultaneamente. | 92 |
| 4.18 | Detecção de falsos positivos. As setas indicam regiões de ocorrências de falsos positivos. | 94 |
| 4.19 | Disposição física das câmeras C_1, C_2 e C_3 no ambiente do VeRLab. A rede é formada por quatro computadores P_1, P_2, P_3 e P_4 | 97 |
| 4.20 | Conjunto de dados da trajetória estimada obtidos pela câmera 1 (a). Trajetórias e alvos identificados na câmera 1 (b). Pontos da trajetória estimada obtidos pela câmera 2 (c). Trajetórias e alvos identificadas na câmera 2 (d). | 98 |
| 4.21 | Trajetórias sobre uma região de monitoramento. As retas são as aproximações da trajetória em cada campo de visão das câmeras. Os retângulos tracejados representam uma aproximação grosseira do campo de visão das câmeras. | 100 |

| | | |
|------|--|-----|
| 4.22 | Fotografias do posicionamento do robô em relação à trajetória. Como trajetória de referência e guia para lançamento dos alvos, usou-se a fita preta marcada sobre o chão do laboratório. A bola colorida é lançada e o seu movimento capturado pela câmera (foto 1(a)). Os parâmetros obtidos pela câmera são transmitidos ao robô que planeja sua posição final, executa o movimento de rotação (foto 1(d)) e em seguida a translação para a posição final (foto 1(c)). A seqüência de fotos de 2(a) a 2(c) ilustra o mesmo procedimento com o robô em outra posição. | 108 |
| B.1 | Projeção do movimento de um ponto no plano de imagem. | 127 |
| C.1 | Exemplo de um polígono simples (à esquerda da figura) e um polígono ortogonal (à direita da figura). | 136 |
| C.2 | Polígono de visibilidade de um ponto. | 136 |
| C.3 | Um polígono simples e seu grafo de visibilidade. | 137 |
| C.4 | A triangulação e a colorização dos vértices de um polígono. | 138 |
| C.5 | Em um polígono pente com $n = 3m$ vértices se verifica a condição de necessidade do número de guardas. | 138 |

Lista de Tabelas

| | | |
|-----|---|-----|
| 4.1 | Estatísticas dos erros na estimação e na medida do raio. | 84 |
| 4.2 | Estatísticas dos erros na translação horizontal e vertical. | 91 |
| 4.3 | Valor Médio Observado da Cor. | 93 |
| 4.4 | Ocorrência de falsos positivos em seqüências de vídeo. | 95 |
| 4.5 | Resultados da minimização quadrática (M-Step). | 99 |
| 4.6 | Similaridade das medidas obtidas entre as trajetórias. | 99 |
| 4.7 | Experimentos para análise da similaridade entre observações de trajetórias. | 102 |
| 4.8 | Valor Médio Observado da Cor. | 102 |
| 4.9 | Planejamento da posição e orientação das câmeras. | 105 |

Lista de Algoritmos

| | | |
|---|---------------------------------|-----|
| 1 | MULTITARGETSMCM | 45 |
| 2 | FUSE | 45 |
| 3 | GVIS | 64 |
| 4 | BUSCAGRAFO | 67 |
| 5 | DIJKSTRA | 70 |
| 6 | PLANNING | 71 |
| 7 | AMOSTRAGEMIMPORTANCIA | 120 |
| 8 | REAMOSTRAGEM SIR | 121 |

Capítulo 1

Introdução

If a man will begin with certainties, he shall end in doubts, but if he will be content to begin with doubts, he shall end in certainties.

Francis Bacon (1561–1623), filósofo inglês.

O avanço tecnológico da Visão Computacional e das redes de computadores tornaram viáveis a utilização de sistemas de câmeras distribuídas convenientemente interconectadas. A finalidade geral de tais sistemas é monitorar grandes regiões observando automaticamente o fluxo de pessoas, veículos ou qualquer outro objeto de interesse que se desloque sobre a região monitorada [Ishiguro (1997); Matsuyama (1998); Bowdene Kaew-Tra-Kul-Pong (2005)]. Cenários típicos são demandados por meio dos sistemas de segurança em agências bancárias, lojas de conveniência, aeroportos, portos, grandes estacionamentos [Collins et al. (1999, 2000); Black et al. (2003); Micheloni et al. (2005)], eventos esportivos [Saito et al. (1999); Júnior e de Oliveira Anido (2004)], controle do tráfego de veículos em rodovias [James M. Ferryman (2000); Haag et al. (2001); Jung e Ho (2001); Lou et al. (2002)] e no movimento de grupos de pessoas [McKenna et al. (2000); Boulton et al. (2001); Plankers e Fua (2001); Siebel e Maybank (2002); Siebel (2003); Gasser et al. (2004)].

Embora o desenvolvimento tecnológico tenha contribuído consideravelmente para aquisição, processamento e transmissão da informação visual, a natureza dos sistemas de monitoramento não sofreu mudanças que estendam o alcance e a confiabilidade dos resultados obtidos pelos sistemas de visão computacional atuais. Em geral, a segurança visual ainda é provida por meio sistemas de circuito fechado de TV (CFTV), não apresentando nenhuma característica “inteligente”. Operadores humanos ainda são centrais no processo de decisão e permanecem com a responsabilidade de avaliar, planejar e executar todos os procedimentos.

A Figura 1.1 mostra o esquema de um sistema típico de segurança baseado em CFTV.

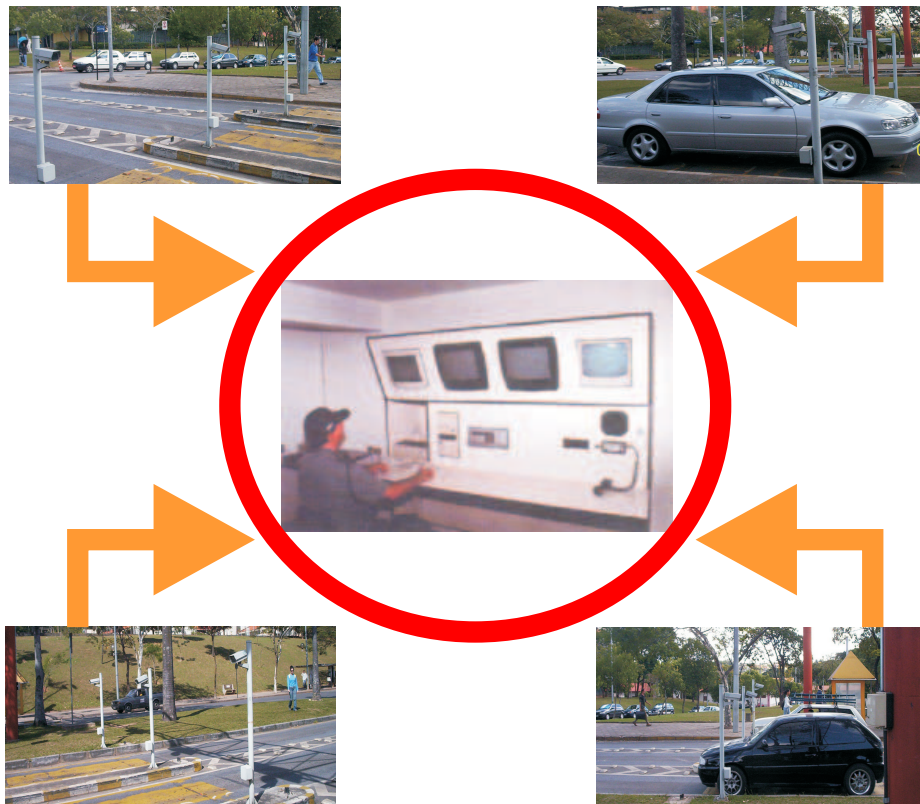


Figura 1.1: Sistema de câmeras de controle de acesso ao Campus. As imagens são capturadas pelo conjunto de câmeras e armazenadas em uma central de segurança. Todo o processo de decisão é centralizado no operador humano.

Nesse sistema, imagens de veículos são capturadas pelas câmeras e enviadas para uma central de segurança, onde funcionários responsáveis monitoram o fluxo de veículos e executam os procedimentos de gravação e controle visual dos dados e informações observadas.

Em geral, pessoas ou veículos que trafegam em regiões monitoradas por câmeras tendem a seguir uma direção natural, obedecendo o curso e o sentido das vias sobre as quais trafegam (Figura 1.2). Porém, mesmo obedecendo uma via pré-definida, monitorar de forma eficiente as trajetórias de veículos ou pessoas em caminhos muito grandes é uma tarefa praticamente impossível para uma única câmera.

A alternativa empregada para a observação de vias extensas é a adoção de câmeras ligadas por um sistema de rede de comunicação. Estas câmeras mesmo separadas por longas distâncias, podem trocar informações e prover o monitoramento mais eficiente de uma ampla região de interesse. Sistemas de câmeras que compartilham informações dotadas com a habilidade de movimento para se posicionar de maneira adequada na observação



Figura 1.2: Caminhos, vias e faixas de tráfego formam direções preferenciais para a trajetória. Nota-se nas fotografias que pessoas e veículos tendem em seguir a trajetória definida pelas vias.

da trajetória de alvos¹ é uma alternativa atraente e desafiante, caracterizando-se como o objeto de estudo desta tese.

O problema de observar a trajetória de alvos com câmeras dotadas de movimento de translação e rotação planar é denominado aqui como *Observação Dinâmica Cooperativa (ODC)*. Este problema insere-se na área de *Visão Computacional Ativa*, subárea da Visão Computacional que trata da modelagem de estratégias de controle para a percepção[Bajcsy (1988)]. Um sistema de visão ativa é capaz de interagir com o ambiente por meio da modificação da direção da visada da câmera, em vez de observar passivamente, operando em tempo real sobre uma seqüência de imagens[Bajcsy (1988); Y. Aloimonos e Bandopadhyay (1988)]. Nesse contexto, o problema de pesquisa pode ser assim definido:

Dada uma região de monitoramento \mathcal{S} , um conjunto \mathcal{A} de alvos e um conjunto \mathcal{C} de câmeras dotadas da habilidade de movimentação planar, o Problema de Observação Dinâmica Cooperativa (ODC) consiste na obtenção de informações visuais que viabilizem a postura das câmeras $C_i \in \mathcal{C}$ no ambiente a partir das informações obtidas por uma das câmeras e compartilhadas pelo grupo, garantindo a observação visual dos alvos $a_j \in \mathcal{A}$ e de suas trajetórias descritas no espaço de monitoramento.

■

Especificamente, deseja-se obter informações visuais por meio da identificação de movimento dos alvos e a partir do compartilhamento da informação da trajetória observada

¹Nesta Tese, assume-se o termo *alvo* como sinônimo de objeto de interesse.

no sistema. As outras câmeras devem se posicionar para observar os objetos de interesse que se movimentam seguindo as trajetórias previamente identificadas. A motivação para o desenvolvimento deste trabalho e os objetivos a serem alcançados estão descritos na seção seguinte.

1.1 Motivação

A principal motivação para este trabalho é a demanda por sistemas de vigilância e proteção ambiental que podem ser atendidas pelas tecnologias de visão computacional. Além das tradicionais tarefas de vigilância, nota-se a potencial aplicação dos sistemas de visão computacional ativa em tarefas de sensoriamento ambiental, coleta de dados e observação de grandes regiões geográficas.

O Projeto SIVAM [SIVAM (2005)], por exemplo, é uma experiência recente no Brasil que apresenta a viabilidade potencial para o emprego de sistemas de monitoramento visual distribuído, provendo o auxílio ao controle, vigilância, segurança e aquisição de dados na Região Amazônica.

O objetivo do SIVAM é fornecer a infra-estrutura de meios operacionais para a coleta, processamento e difusão de dados de interesse para as organizações integrantes do Sistema de Proteção da Amazônia – SIPAM,² comporta, fundamentalmente, um sistema de aquisição de dados formado por uma rede dos mais variados tipos de sensores, incluindo sensoriamento remoto por satélite e por aeronaves, equipamentos de coleta de dados meteorológicos, hidrológicos e ambientais.

Sistemas com a estrutura do SIVAM apresentam, naturalmente, uma arquitetura formada por diferentes níveis de sensoriamento, favorecendo um esquema mais acurado de aquisição e integração de dados multi-escala. Esta estrutura pode adequadamente incluir sistemas de visão distribuída baseados, por exemplo, em grupos de robôs móveis cooperativos, provendo ações de vigilância e proteção ambiental com a tarefa de rastreamento e observação de alvos em ambientes desconhecidos, inóspitos ou de difícil acesso ao ser humano.

Um exemplo de aplicação da ODC pode ser idealizado a partir da Figura 1.3. A figura mostra parte do campus da UFMG onde estão destacadas algumas das suas principais vias de acesso e áreas de estacionamento. Um sistema de visão distribuída baseado nos conceitos de troca de informações entre as câmeras e da eventual capacidade de movimentação destas, viabilizaria ações como a obtenção das características de todos os veículos que acessam as

²Em geral, órgãos dos Governos Federal, Estaduais, Universidades e Forças Armadas.

áreas do campus e a possibilidade de verificação remota e controle dos veículos.

No sistema de segurança idealizado na Figura 1.3, os veículos ao ingressarem por alguma das entradas do campus, têm suas características visuais extraídas (cor predominante, forma, textura, dimensões, número da placa, entre outras). Ao mesmo tempo, são também realizadas inferências sobre a velocidade, direção e tempo de chegada do veículo a um dos estacionamentos ou saídas do campus. Estas informações são enviadas a todos os postos de observação (câmeras) do sistema que procuram, por sua vez, identificar e casar as informações recebidas com as informações que estão sendo obtidas em suas respectivas áreas de cobertura.

Eventualmente, a detecção de problemas, ocorrência de ações ilícitas ou falhas no sistema exemplificado podem gerar a emissão de sinais de controle, acionar alarmes ou solicitar a verificação de locais específicos por meio de agentes móveis. Neste último caso, robôs móveis dotados com câmeras seriam automaticamente deslocados para as possíveis regiões de ocorrência usando as informações providas pelo sistema.

Por outro lado, considerando-se os problemas atuais como crescimento da violência urbana e a ameaça do terrorismo internacional, aplicações dos conceitos desenvolvidos para a ODC podem ser estendidos a várias outras aplicações. Entre essas podem-se citar o apoio a operações policiais em aglomerados urbanos, repreensão a ilícitos, inspeção em áreas de conflito, identificação de movimento de passageiros em aeroportos, entre outras.

1.2 A Abordagem Desenvolvida

Para se observar os alvos dinamicamente, o sistema deve ser capaz de adquirir informações sobre estes e produzir, a cada instante de tempo, uma estimativa de medidas de interesse dos alvos sob observação. Dado que uma câmera do sistema tenha identificado uma trajetória, esta informa aos outros membros do grupo as características visuais e os parâmetros observados.

Com base nas informações recebidas, as outras câmeras do grupo devem procurar posições que viabilizem a aquisição de observações a respeito da trajetória e dos alvos anteriormente detectados. As observações devem ser então associadas com as informações recebidas, reconhecendo se o objeto visto em outras regiões corresponde ao mesmo que esteja sendo monitorado a partir da posição encontrada.

Na sua forma geral, esta abordagem se relaciona com as soluções propostas para dois tipos de problemas: *rastreamento visual de múltiplos objetos* e o *planejamento de sensores*.

Abordagens relacionadas com rastreamento de múltiplos objetos têm sido estudadas

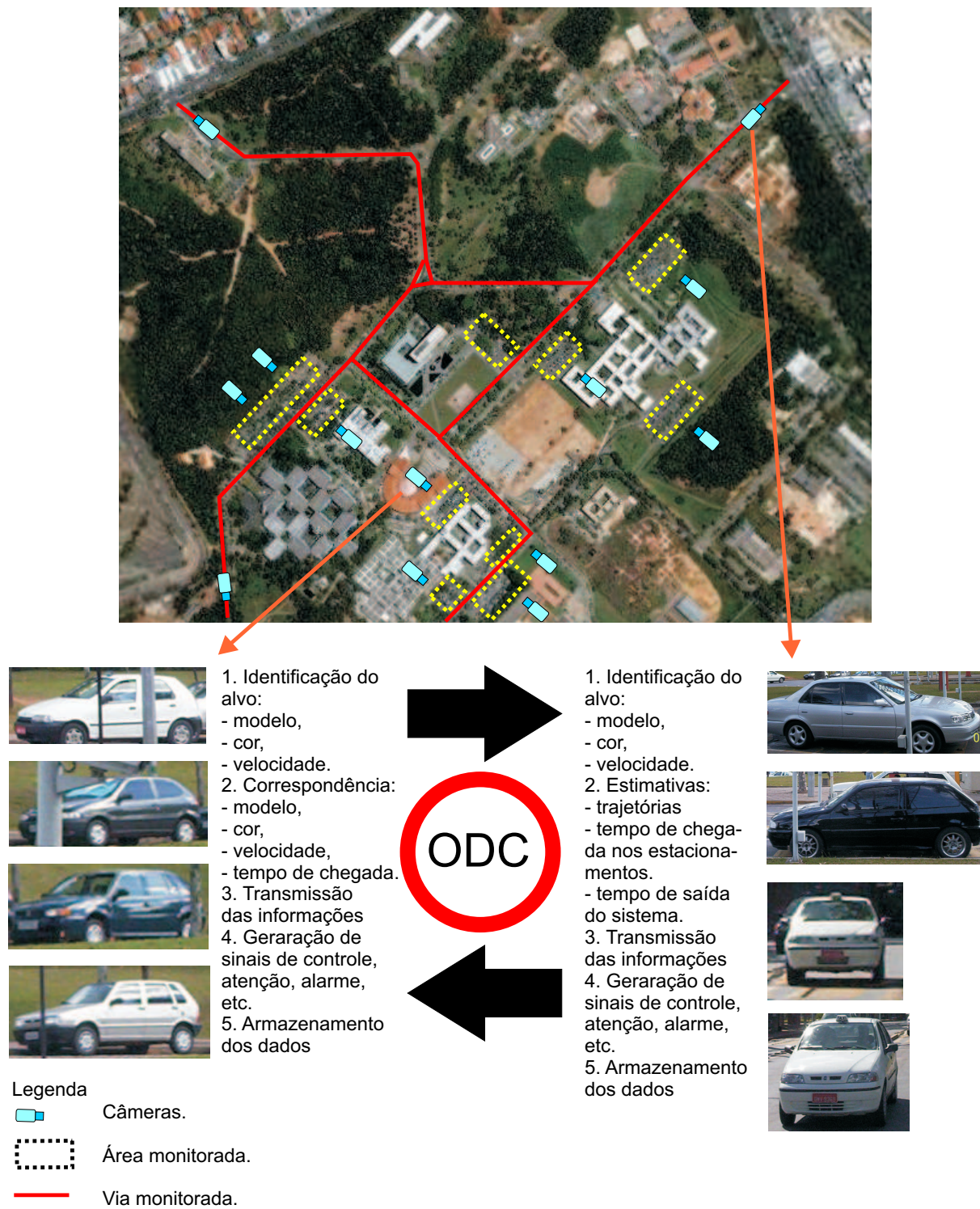


Figura 1.3: Sistema idealizado de câmeras de segurança nas vias de acesso do campus da Universidade Federal de Minas Gerais. As câmeras estão distribuídas pelas principais vias e nas áreas de estacionamento. O sistema extrai as características dos veículos e distribui por todas as câmeras. As câmeras podem correlacionar as características e identificar a chegada dos veículos nas áreas de estacionamento.

em diferentes contextos e aplicações. Tradicionalmente, a comunidade de processamento de sinais foi pioneira em tratar questões relacionadas ao rastreamento de alvos [Bar-Shalom (1978, 1990); Gutman e Velger (1990); Bar-Shalom (1992); Gordon et al. (1993); Lane et al. (1998); Hue (2002); Blackman (2004); Vercauteren et al. (2005)], particularmente devido aos problemas demandados pelos sistemas de RADAR estudados desde o período da Segunda Guerra Mundial [Rao e Dhawas (1995); Daeipour et al. (1997); Park e Lee; (1998)]. Ainda hoje é muito grande a contribuição dos grupos de pesquisa em processamento de sinais no desenvolvimento de filtros preditivos e técnicas de rastreamento para uso militar [Song et al. (1990); Yeddapanudi et al. (1995); Siouris et al. (1997); Penny (1999)].

Embora a contribuição dos resultados da área de processamento de sinais sejam importantes para a fundamentação da proposta desenvolvida neste trabalho, são os resultados obtidos em visão computacional que fornecem o direcionamento seguido para o desenvolvimento da solução proposta.

Sob o ponto de vista da visão computacional, um problema de rastreamento é essencialmente um problema de correspondência [Trucco e Verri (1998)]. Os sistemas de visão utilizam câmeras para rastrear alvos a partir de suas características visuais e estabelecer a correspondência entre entidades (objetos e características visuais) vistas na imagem corrente e aquelas vistas nas imagens anteriores [Murray e Basu (1994); Isard e Blake (1998); Saito et al. (1999); Tao et al. (1999); Javed et al. (2000, 2003); Chen et al. (2001); Ito e Sakane (2001); Jennings (2002); Forestie Micheloni (2003); Ukita e Matsuyama (2005)].

Mais recentemente, o problema de rastreamento de alvos tem atraído soluções em robótica móvel e em redes de sensores sem fio. Em robótica móvel, utiliza-se robôs móveis dotados com diferentes tipos de sensores para rastrear os alvos, procurando mantê-los sob contínua observação [Leonard et al. (1991); Balche Parker (2001); Vidal e Sastry (2002); Parker (2002); Jung e Sukhatme (2002a,b, 2004)]. O movimento dos robôs é realizado a partir de sinais de controle obtidos das informações sensoriais capturadas dos alvos [Han e Veloso (1997); LaValle et al. (1997); LaValle e Hutchinson (1998); Stroupe e Balch (2003)].

A abordagem convencional usada para o rastreamento em redes de sensores sem fio, caracteriza-se pela distribuição de pequenos nós sensores pelo ambiente, formando uma rede de comunicação sem fio. Esses nós, com pequena capacidade de processamento, detectam o movimento ou a presença de objetos dentro da sua área de cobertura e propagam essa informação para os outros nós da rede. A informação é propagada até o ponto de acesso da rede com maior capacidade de processamento e capaz de identificar e posicionar os objetos na área de cobertura dos outros sensores [Chen et al. (2004); Hsieh (2004); Kotecha et al. (2005); Ssu et al. (2005)].

A questão fundamental em relação ao problema de *planejamento de sensores* é deter-

minar a posição e a orientação (*pose* ou *postura*) dos sensores de tal forma que se obtenham melhores observações sobre os alvos. As principais abordagens usam informações providas por meio da geometria da cena, da topologia do ambiente e da área de cobertura dos sensores.

Sob o ponto de vista teórico, o problema de posicionar sensores tem sua origem na Geometria Computacional, no problema denominado de “Problema da Galeria de Arte” [ORourke (1983, 1987); Liaw et al. (1993)]. Em sua formulação mais simples, o problema da galeria de arte consiste na determinação da posição de sensores (ou guardas) em um ambiente poligonal (a galeria de arte) [Michael e Pinciu (2003)]. Porém, diferentes variações desse problema incluem polígonos com buracos, guardas que podem ver em cone ou com visão omnidirecional, galerias de arte tridimensionais e guardas com capacidade de mobilidade [Bose et al. (1997); Bjorling-Sachs (1998); Guibas et al. (1997, 1999); Tóth (2000); Hlin (2001); Simov et al. (2002); LaValle et al. (2002); Tóth (2002a,b, 2003); Sachs et al. (2004)].

A abordagem desta tese pode ser vista aqui como uma rede de comunicação descentralizada e totalmente conectada, formada por câmeras assentadas sobre robôs móveis que se comunicam entre si e agem em um esquema de cooperação³ na condução da tarefa de observar alvos e suas trajetórias. Propõe-se assim, uma arquitetura geral que está subdividida em três módulos principais:

- o módulo de rastreamento visual;
- o módulo de associação e reconhecimento de trajetórias;
- e o módulo de planejamento de posição.

Os três módulos interagem em um esquema de malha fechada. A partir da associação das informações visuais obtidas pelas câmeras, cada câmera planeja o seu posicionamento em relação a trajetória descrita pelos alvos e movem-se de tal forma que a observação do movimento dos objetos seja garantida a partir da posição planejada.

A abordagem aqui desenvolvida diferencia-se das demais apresentadas por estruturar a tarefa de observação com base na associação e compartilhamento da informação visual em uma estratégia de planejamento de posição, gerando um esquema de auto-organização para a observação provida por meio do posicionamento das câmeras. Essa é a idéia principal que será explorada em profundidade no restante deste trabalho.

³No contexto deste trabalho, a cooperação está relacionada com a troca de informações e a organização das câmeras.

1.3 Contribuições

O trabalho aqui desenvolvido promove, fundamentalmente, contribuições na área de Visão Computacional Ativa. As principais contribuições nesta área, são:

- *Caracterização e formalização do problema de observação dinâmica cooperativa de alvos.*
- *Um método para o rastreamento de múltiplos alvos móveis em tempo real.*

É desenvolvida uma nova abordagem para o rastreamento visual de múltiplos objetos em tempo real. Cada alvo detectado é rastreado por um filtro de partículas separado. O método utiliza a informação de cor do movimento do alvo para a associação entre as medidas observadas e as trajetórias. Estas mesmas informações são utilizadas para redistribuir uma parcela das partículas nas regiões de maior probabilidade no espaço de estado. A complexidade do método cresce linearmente em relação ao número de partículas.

- *Desenvolvimento de uma plataforma sensorial dinâmica e flexível o suficiente para a mudança de tipo de sensoriamento e integração de outros tipos de sensores.*

Diferente da maioria das técnicas usuais de rastreamento, nas quais os sensores permanecem estáticos em posições previamente escolhidas, a plataforma sensorial aqui projetada utiliza a capacidade de mobilidade do grupo de robôs móveis para o posicionamento dinâmico das câmeras no ambiente.

- *Captura dinâmica de uma vasta área e entendimento da cena.*

As câmeras são distribuídas sobre uma determinada área para a captura de diferentes vistas simultâneas. Este esquema propicia o entendimento de situações dinâmicas e dos eventos que ocorrem em uma cena muito mais ampla que a obtida por uma única vista de um único sensor.

- *Um arcabouço computacional para o rastreamento por meio de múltiplas câmeras sem a necessidade sobreposição dos campos de visão*

O rastreamento visual em sistemas multi-câmeras convencionais geralmente utiliza as informações sobre os objetos visualizados nas áreas de sobreposição dos campos visão das câmeras. Na abordagem desenvolvida, relaxa-se a necessidade de sobreposição dos campos de visualização e utiliza-se a informação compartilhada pelo grupo para reconhecer e associar as informações sobre os objetos.

1.4 Organização do Texto

O texto está organizado em cinco capítulos, contando com este introdutório, e três apêndices que apresentam alguns conceitos relacionados com tema do trabalho. O conteúdo de cada capítulo está descrito na seqüência.

Capítulo 2 apresenta a revisão da bibliografia dos trabalhos relacionados com a observação dinâmica cooperativa. Em particular, enfoca-se o estado da arte dos sistemas de visão computacional distribuída e fornece-se uma visão geral sobre o rastreamento visual de múltiplos objetos e o posicionamento de sensores.

Capítulo 3 trata do desenvolvimento da metodologia. Nesse capítulo são descritos os algoritmos de rastreamento implementados com base nos métodos de Monte Carlo Seqüenciais (ou filtros de partículas) capazes de rastrear múltiplos alvos em tempo real. Também se descreve a estratégia adotada para a associação entre características visuais e trajetórias. A estratégia de planejamento de posição para as câmeras é desenvolvida em detalhes nesse capítulo.

Capítulo 4 mostra o planejamento, execução e a análise dos experimentos realizados com base na metodologia. Os resultados são validados por meio de um protocolo de experimentação que leva em consideração os aspectos teóricos e práticos.

Capítulo 5 apresenta as conclusões da tese, temas propostos para investigação e desenvolvimento futuro.

Capítulo 2

Trabalhos Relacionados

What we have learned from others, becomes our own by reflection.

Ralph Waldo Emerson (1803–1882), ensaísta e poeta norte-americano.

O objetivo deste capítulo é apresentar as abordagens e os trabalhos relacionados com problema da ODC. Devido a característica multidisciplinar desta tese, os trabalhos serão apresentados seguindo a sistematização proposta na Figura 2.1. Sob este enfoque, a Seção 2.1 apresenta os trabalhos relacionados com filtros de partículas e o rastreamento visual de múltiplos objetos. A Seção 2.2 destaca o rastreamento de alvos utilizando robôs, enfocando as soluções para o problema de rastreamento sob o enfoque da robótica móvel. Por fim, na Seção 2.3 são apresentadas as abordagens usuais para o planejamento de sensores, dando-se particular destaque aos problemas geométricos.

2.1 Rastreamento Visual de Múltiplos Alvos

O problema de rastreamento de múltiplos alvos tem sido estudado sob diferentes pontos-de-vista por diferentes grupos de pesquisa. Embora, os objetivos de estimar posição e características do movimento dos alvos sejam comuns, as abordagens diferenciam-se na maneira de tratar os dados e na forma de associação entre observações e trajetórias. [Bar-Shalom (1992)].

A literatura de rastreamento visual é muito vasta e diversificada em termos de abordagens. Uma análise exaustiva de todos os métodos e trabalhos realizados nessa área está fora do escopo deste trabalho. O enfoque aqui desenvolvido trata das soluções que rastreiam visualmente múltiplos alvos. Em particular, as abordagens Bayesianas que utilizam os métodos sequenciais de Monte Carlo, ou filtro de partículas.

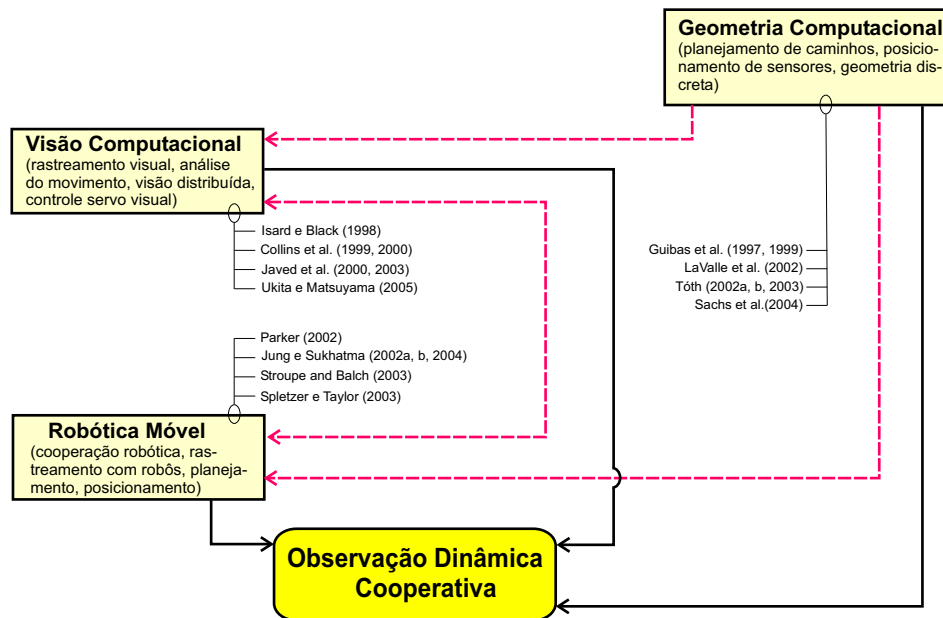


Figura 2.1: Principais áreas envolvidas no trabalho e algumas referências em destaque.

O trabalho pioneiro na aplicação do filtro de partículas em visão computacional foi apresentado por Isard e Blake (1998) como o algoritmo CONDENSATION. Nesse trabalho, os objetos são modelados como um conjunto de B -splines parametrizadas e então o método de Monte Carlo é usado para prever a posição das curvas durante o processo de rastreamento. O algoritmo CONDENSATION propaga as densidades de vários estados no espaço de busca, o que é muito atrativo em vários problemas de rastreamento, particularmente quando se procura rastrear múltiplos objetos. Porém, o algoritmo CONDENSATION básico, diferente da abordagem desenvolvida nesta tese, não foi projetado com o propósito de rastrear múltiplos alvos em aplicações de tempo real.

Em trabalho posterior, MacCormick e Blake (1999) aplicou o princípio de exclusão probabilística com o objetivo de incorporar ao modelo básico do CONDENSATION a habilidade de rastrear múltiplos objetos. Porém, o algoritmo é muito dependente do modelo de observação e os resultados apresentados não tratam um número considerável de objetos rastreados. Um problema perceptível nesse algoritmo é a sua incapacidade de rastrear um objeto que não tenha sido reconhecido e iniciado no no começo do processo de rastreamento. Para contornar este problema, um modelo de mistura no espaço métrico com tratamento probabilístico foi proposto por Toyama e Blake (2001).

Um tratamento probabilístico para a rastrear múltiplos objetos a partir do método *Condensation* foi desenvolvido por Meier e Ade (1999). Uma função de densidade de pro-

babilidade descrevendo o estado mais provável dos alvos é propagada no tempo por meio de um modelo dinâmico. As medidas obtidas pelo modelo de observação são incorporadas aos novos objetos. Durante o processo de reamostragem, em vez de se selecionar N amostras, como no algoritmo convencional, são selecionadas $N - M$ amostras. Estas novas amostras selecionadas são adicionadas a outras M amostras baseada nas medidas do modelo de observação. Este processo de reamostragem favorece a inclusão de novos objetos no processo de rastreamento. Um resultado importante deste trabalho, adaptado ao modelo de rastreamento desenvolvido nesta tese, é a representação da função de densidade de probabilidade por meio de um número fixo a amostras. Esta abordagem garante um tempo de execução constante em cada iteração do algoritmo.

O método denominado de rastreador Bayesiano de múltiplos *blobs* (*Bayesian multiple-blob tracker – BraMBLe*) foi proposto por Cormicke Isard (2001). Os autores adaptaram verossimilhanças como modelos de observação para o fundo e para o plano frontal de uma cena e aplicaram um filtro de partículas para rastrear múltiplos objetos. O método trabalha com múltiplos objetos ordenados pela profundidade devido a um modelo de estados de espaço tridimensional. O grande potencial desse método é a incorporação de estatísticas para a separação entre o plano frontal e os objetos, viabilizando a identificação sem ambigüidade entre os blobs. O maior problema é devido ao tempo de treinamento necessário para os diferentes casos.

Zia Khan e Dellaert (1999) descreve um filtro de partículas para o rastreamento de múltiplos objetos que leva em consideração a influência devido a proximidade ou comportamento de outros objetos. Os autores mostram nesse artigo um modelo de movimento baseado em um campo aleatório de Markov. Baseados na premissa de que dois corpos não ocupam ao mesmo tempo o mesmo lugar no espaço e que os alvos tendem a evitar colisões. As interações entre objetos são modeladas em uma aplicação para o rastreamento do comportamento social de insetos. Porém, diferente da abordagem desenvolvida nesta tese, o rastreador conjunto de partículas apresenta complexidade exponencial em relação ao número de alvos rastreados.

Mais recentemente, Patrick Pérez e Blake (2004) introduziu um mecanismo de amostragem por importância genérico para a fusão de dados de cor com som estéreo. O objetivo do trabalho são aplicações em teleconferência ou para a fusão com informações sobre o movimento em sistemas de segurança com câmeras fixas. O trabalho mostra como essas informações podem ser modeladas em uma função de verossimilhança apropriada e como informações intermitentes (o som e o movimento) são melhor arranjadas para a geração dessas funções de verossimilhança. O uso da fusão de informações de cor e movimento também é explorado no desenvolvimento do algoritmo de rastreamento desta tese.

Porém, de forma diferenciada, as informações são fundidas em tempo real em uma função de densidade de probabilidade usada para propagar hipóteses sobre o posicionamento dos múltiplos alvos que se movimentam sobre a região de observação.

Wu e Huang (2004) apresentou um método de rastreamento baseado na co-inferência entre modelos gráficos. O problema de rastreamento é tratado como um problema de estimação a partir da inferência sobre o estados dos alvos em uma seqüência de vídeo. A análise variacional estruturada dos modelos gera diferentes modalidades de gráficos que são sujeitas a um processo de co-inferência. Baseado na técnica de reamostragem por importância (descrita com detalhes no Apêndice A) o método de Monte Carlo é então aplicado para estimar o estado dos alvos. Os resultados obtidos são avaliados por meio de estudos de simulações.

Nesta tese implementou-se um rastreador consistindo de múltiplos filtros de partículas independentes, capazes de rastrear múltiplos objetos em tempo real com qualidade de rastreamento similar aos apresentados e com baixo custo computacional. O rastreador utiliza a cor dos alvos associada ao movimento destes em uma seqüência de quadros. Estas duas informações são fundidas e utilizadas para direcionar a propagação das partículas, conveniente redistribuídas e propagadas em regiões do espaço de estado de mais alta probabilidade.

Uma segunda inovação desenvolvida neste trabalho, ainda no contexto do rastreamento visual, consiste em um maneira diferenciada de iniciar o estado dos alvos. Aqui, os blobs são identificados e associados às partículas no processo de associação de dados. Assim, blobs que possuem valores de probabilidade crescente são automaticamente instanciados em objetos a serem rastreados. Por outro lado, blobs que apresentam baixo valor de probabilidade são excluídos do processo de rastreamento, considerados perdidos ou saídos do campo de visão da câmera. Este processo resultou em um rastreador de múltiplos blobs com performance igual ou superior aos métodos usuais.

A seção seguinte apresenta alguns trabalhos onde serão consideradas múltiplas câmeras com o objetivo de monitorar regiões, rastrear alvos e obter informações da cena.

2.1.1 Sistemas com Múltiplas Câmeras

Atualmente existe uma grande variedade de aplicações nas quais um conjunto de câmeras interligadas é distribuído em um ambiente objetivando a visualização simultânea de diversas regiões de uma área de monitoramento. Em geral, essas aplicações são voltadas para segurança em bancos, parques, grandes estacionamentos, portos, aeroportos e estações de trem ou metro. O objetivo principal desses sistemas é adequar maior eficiência ao sistema

de segurança por meio da combinação das técnicas de visão computacional com tecnologia de redes de computadores em aplicações de rastreamento de pessoas e veículos [Collins et al. (1999)].

Em geral, os sistemas de múltiplas câmeras são projetados em circuito fechado de TV (CFTV) e têm apresentado resultados surpreendentes na identificação de suspeitos e na detecção de atividades ilícitas. Porém, a grande maioria ainda deixa por conta do operador humano o processo de controle e tomada de decisão [Ellis (2002)].

No entanto, o avanço da capacidade de computação aliado ao desenvolvimento de algoritmos de visão computacional mais eficientes, viabilizam soluções para a aquisição automática de informações em um ambiente de câmeras distribuídas. Woo e Capson (2000) descreve um sistema distribuído de câmeras para o rastreamento de características no espaço tridimensional. Os autores projetaram uma unidade de câmera de baixo custo que incluem uma pequena câmera CCD e dois mecanismos servo-controlados para rotações nos ângulos de *pan* e *tilt*¹. A Figura 2.2 mostra o modelo esquemático do sistema de câmeras distribuídas. No detalhe, vê-se a unidade de câmera pan/tilt. Um procedimento de triangulação a partir de duas ou mais câmeras calibradas é utilizado para calcular o posicionamento do alvo por meio da posição das câmeras e dos valores dos ângulos de tilt e pan. Os autores apresentam resultados experimentais que mostram a viabilidade do sistema e a vantagem do uso de dados redundantes obtidos pelas câmeras.

O trabalho de Kogut e Trivedi (2001) trata da infraestrutura necessária para a utilização de redes de câmeras em áreas metropolitanas. Uma vez que estas redes fornecem uma grande quantidade de informação visual, provida, na maioria dos casos, por um pequeno número de câmeras comuns, o trabalho mostra um algoritmo que executa a fusão dos dados visuais, procurando manter a identidade de veículos, ou de grupos de veículos, que atravessam as áreas monitoradas por essas câmeras. O algoritmo usa a combinação da cor com a organização espacial dos veículos, objetivando minimizar a ocorrência de falsos positivos e a obtenção de informações sobre o tráfego. Seguindo o mesmo princípio de associação entre características, Dockstadere Tekalp (2001) aborda os problemas de oclusão em sistemas de rastreamento de múltiplos alvos por meio da fusão da informação obtidas por múltiplas vistas de uma mesma cena, utilizando uma rede de crença Bayesiana.

Durante os últimos anos, têm-se notado o crescimento considerável do interesse em sistemas de segurança por meio de câmeras de vídeo. Este interesse dá-se, principalmente, por causa dos problemas relacionados à segurança em aeroportos e estações ferroviárias

¹ *Tilt* corresponde movimento efetuado com a câmera verticalmente, geralmente em velocidade lenta, de cima para baixo ou vice-versa. *Pan*(de panorâmica), por sua vez, é o movimento efetuado com a câmera horizontalmente, geralmente em velocidade lenta, de um lado para outro.

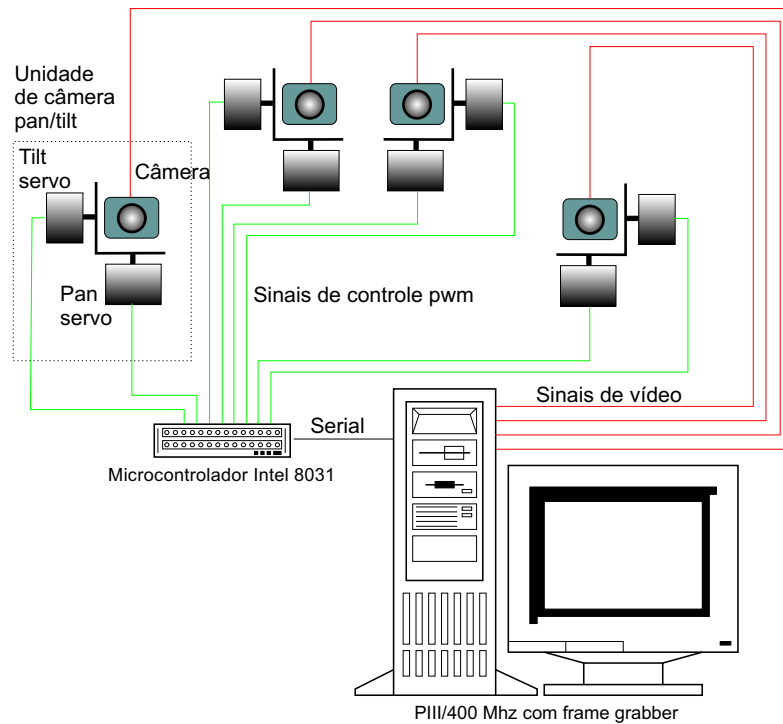


Figura 2.2: Esquema do sistema de câmeras pan/tilt distribuídas [Woo e Capson (2000)]. As câmeras possuem movimento de rotação em relação ao eixo vertical e capacidade de aproximação visual por meio das lentes zoom. Sinais de vídeo e de controle são coordenados por um microcontrolador associado a uma CPU.

devido aos recentes ações do terrorismo internacional[McKenna (2003)].

Boult et al. (2001) oferece uma visão geral dos sistemas automáticos de segurança com vídeo sob o enfoque histórico/evolutivo desses sistemas e destaca suas aplicações no contexto militar. O foco é dado sobre a importância da utilização desses sistemas em situações críticas de monitoramento de áreas de floresta, onde ocorre a infiltração de alvos convenientemente camuflados. Situações dessa natureza requerem apurada capacidade de sensoriamento e maior campo de visão das câmeras, o que levou os autores ao desenvolvimento de um sistema de segurança baseado em câmeras omnidirecionais. O principal problema abordado é a identificação de um alvo camuflado em meio a vegetação, problema este que pode levar os sistemas convencionais de detecção de alvos a um elevado nível de falsos positivos, uma vez que, em cenas naturais, o movimento e a mudança de luminosidade produzem cenas extremamente dinâmicas. O artigo discute, fundamentalmente, o estado da arte da detecção de alvos, incluindo os algoritmos de análise e adaptação de fundo. Destaca-se ainda o sistema LOTS (*Lehigh Omnidirectional Tracking System*) e seus componentes, que incluem a modelagem adaptativa de múltiplos fundos, agrupamento de componentes quase-conexos (uma nova abordagem para o agrupamento de componentes espaço-temporais), análise de subtração de fundo e a avaliação experimental de todo o sistema.

Ellis (2002) apresenta a descrição completa de um sistema de segurança em vídeo com múltiplas vistas. Nesse trabalho os autores também descrevem os algoritmos usados para detectar e rastrear objetos (baixa densidade de pedestres, ciclistas ou automóveis) que se movem em ambiente externo e são detectados por meio de uma rede de câmeras de um sistema de segurança. O sistema foi projetado para se adaptar às diferentes condições de iluminação e ao tratamento do movimento de objetos espúrios (vegetação) provenientes de cenas em ambientes externos. O sistema utiliza, quando possível, as múltiplas vistas de um mesmo objeto para tratar os problemas de oclusão e aumentar a robustez do rastreamento. A correspondência entre os objetos nas áreas de sobreposição das câmeras é analisada por meio da geometria projetiva. O objetivo geral do sistema apresentado é capturar a cena por meio de um processo de aprendizado dependente da informação, construído modelos da cena usando as observações obtidas a partir da rede de câmeras de vídeo.

Berriss et al. (2003) mostra a aplicação dos algoritmos de processamento digital de imagens e vídeo no campo de sistemas de segurança por meio de vídeo digital. O objetivo é identificar os algoritmos de processamento de imagens que são mais indicados para a análise rápida e eficiente, em tempo real, de grandes bases de dados de vídeo. Os algoritmos investigados foram projetados para extrair descritores do conteúdo das imagens, tais como cor, forma, textura e movimento dos objetos. Os autores propõem ainda uma abordagem

que utiliza a cor como descritor para a recuperação da informação por conteúdo aplicado na identificação de uma determinada pessoa em uma base de dados de seqüências de vídeo. O trabalho restringe-se à análise dos métodos aplicáveis à recuperação de imagens por conteúdo, tratando pouco dos casos de rastreamento e aplicações em tempo real.

Boult (2003) faz uma revisão histórica de alguns sistemas de vigilância visual automática, desde a primeira geração desses sistemas, no início dos anos de 1990, até o corrente estado da arte. Nesse trabalho, Boult discute as limitações, tanto de hardware quanto de software, dos primeiros sistemas de segurança visual, destacando as soluções obtidas para o aumento da performance, vigentes até os dias atuais, nos sistemas de segurança comerciais. A abordagem adotada objetiva identificar as limitações relacionadas com a visão ativa sob regras geo-espaciais. Os sistemas de visão ativa destacados nesse trabalho, utilizam dados controlados por computador obtidos por meio de câmeras *Pan/Tilt/Zoom* (ou unidades PTZ - câmeras dotadas com movimentos de pan e tilt e com lentes zoom), os quais são combinados em algoritmos de detecção e rastreamento visual. Embora os conceitos apresentados no trabalho não apresentem nenhuma novidade, a inclusão de regras derivadas das informações geo-espaciais associadas aos dados multi-sensoriais fornecem conceitos fundamentais para o projeto de redes de sensores sem fio.

As soluções tecnológicas que propiciam avanços ao estado da arte dos sistemas de segurança em vídeo são analisadas por Black (2004). Nesse trabalho, os autores tratam especificamente dos problemas relacionados com os sistemas de segurança em vídeo em ambientes dinâmicos e complexos dos aeroportos. Entre as várias soluções propostas, destaca-se a perseguição automática por vídeo (*video pursuit*), que integra a segurança em vídeo com estratégias militares por meio de um programa do tipo "comando e controle". O trabalho aborda ainda a integração entre vídeo e outros sistemas de comunicação. O projeto das aplicações mostradas no trabalho, tem como princípio o uso de tecnologias que favoreçam o desenvolvimento de uma plataforma sólida para sistemas de segurança, objetivando a expansão e integração do sistema sem a necessidade de expor o sistema a situações críticas. Porém, toda a operacionalização do sistema e parte do controle ainda está presente nas mãos do operador humano.

A segurança doméstica controlada por meio de uma rede de câmeras é o tema abordado por Cucchiara et al. (2005). Problemas de propósito geral em segurança doméstica e as técnicas usuais em visão computacional são discutidas e implementadas como uma solução para a detecção e rastreamento de objetos capaz de tratar sombras, oclusão e postura de pessoas. As câmeras são convenientemente distribuídas e instaladas em cada sala da casa, com o objetivo de monitorar a movimentação das pessoas. O rastreamento é baseado em uma abordagem probabilística na qual a aparência e probabilidade de oclusão

são calculadas a partir de uma câmera corrente e transmitida para uma câmera seguinte para o posicionamento, evitando a eventual oclusão. Esta aplicação se contextualiza no área emergente denominada de *domatics* (*domus*, do latim, significando lar e *informatics* - informática, em inglês). Em particular, o trabalho mostra como a tecnologia de segurança por meio de câmeras pode ser convenientemente aplicada em uma residência com grau de autonomia suficiente a custos convenientes e elevado nível de confiança.

Fora do contexto dos sistemas de segurança, encontram-se ainda aplicações dos sistemas de múltiplas câmeras em problemas que envolvem além do rastreamento, a reconstrução tridimensional. O trabalho de Jennings (2002) usa pares de imagens estéreo para o rastreamento dos dedos da mão humana em espaço tridimensional. Um par de câmeras estéreo também é utilizado em Harville (2002) para o rastreamento de pessoas. Um plano de visualização é montado baseado nas estatísticas obtidas pelo movimento tridimensional das pessoas. Porém, sistemas dessa natureza partem de premissas que viabilizam a obtenção direta da informação tridimensional, como a utilização da visão estéreo, tornando necessário a utilização de técnicas de calibração complexas e o conhecimento *a priori* das áreas de sobreposição dos campos de visão das câmeras.

Ishiguro (1997) foi pioneiro em propor um Sistema de Visão Distribuída (DVS – *Distributed Vision System*) que funciona como uma infraestrutura de informação perceptual para a navegação de robôs móveis em um ambiente estruturado. O sistema é formado por Agentes Visuais (VA – *Visual Agent*) que são, na realidade, câmera fixas dotadas de algum poder de computação convenientemente distribuídas pelo ambiente e interligadas por um sistema convencional de rede de computadores. A infraestrutura de câmeras do sistema é denominada de Infraestrutura de Informação Perceptual (PI² – *Perceptual Information Infrastructure*) e seu objetivo é manter uma representação consistente da dinâmica do mundo e o envio de informações para a navegação dos robôs móveis.

A utilização dos conceitos definidos por Ishiguro, encontraram aplicações nos trabalhos desenvolvidos por Sogo et al. (1999). O sistema proposto em Sogo et al. (2001) objetiva a realização da navegação robótica apoiada pelo DVS em duas fases: a fase de ensino e a fase de navegação. Na fase de ensino os VAs observam e memorizam o caminho realizado por um robô controlado por um operador humano. Na fase de navegação, os VAs comunicam-se entre si para obter a informação visual importante para a navegação, e então os robôs são guiados pelos dos caminhos memorizados em cada câmera.

Aplicações de múltiplas câmeras também são observadas para o acompanhamento de jogos, em particular em jogos de futebol [Kang et al. (2003)] e futebol americano. Saito et al. (1999), por exemplo, utiliza trinta câmeras para cobrir uma partida de *Super Bowl*, viabilizando a reconstrução do movimento dos jogadores a partir de qualquer ponto de

visualização e a utilização de quarenta e nove câmeras em um quarto tridimensional (*the 3D room*).

O rastreamento de jogadores de futebol por uma rede de câmeras foi tratado recentemente por Júnior e de Oliveira Anido (2004). Rastrear jogadores em um campo de futebol é um problema desafiante devido, principalmente, ao rápido movimento dos objetos e à ocorrência freqüente de oclusões. O uso de mais de uma câmera para cobrir uma região muito grande, como um campo de futebol, aumenta consideravelmente a acurácia do rastreamento, porém, leva a uma grande quantidade de dados que devem ser processados e convenientemente associados. As câmeras são posicionadas objetivando maximizar a cobertura do campo de futebol e rastrear objetos que se tornam pequenos no plano de projeção das câmeras (entre dez a quarenta pixels de altura). O trabalho desenvolvido trata dos casos onde o processo de rastreamento é aplicado em diferentes câmeras que cobrem partes específicas da região de cobertura. Seis câmeras digitais são posicionadas nos pontos mais altos do estádio e as imagens que são processadas praticamente em tempo real por 66 microcomputadores. Cada jogador é rastreado por uma máquina; as restantes servem para fazer o processamento e detecção dos jogadores nas imagens. O sistema identifica os jogadores por meio de padrões como a cor da camisa ou do calção.

A maioria dos métodos para rastreamento de objetos por meio de múltiplas câmeras anteriormente apresentados, necessitam para o rastreamento a calibração prévia das câmeras e a sobreposição dos campos de visão. Em alguns casos (por exemplo, Saito et al. (1999); Kang et al. (2003); Júnior e de Oliveira Anido (2004)), informações topológicas e geométricas também são necessárias. Métodos mais atuais relaxam a necessidade de calibração, mas mantêm a sobreposição de campos de visão para determinar as correspondências. O interesse pelo rastreamento em sistemas de câmeras sem sobreposição de campos de visão (*handoff camera*) tem crescido consideravelmente nos últimos tempos. Uma característica importante na abordagem adotada nesta tese para a ODC é a capacidade do sistema associar observações à trajetórias em campos de visão disjuntos. Alguns trabalhos que relaxam a necessidade de sobreposição são descritos em seguida.

Huang e Russel (1997) desenvolveram um sistema para rastrear veículos em uma rodovia utilizando câmeras separadas. Os autores descrevem um conjunto de padrões que viabilizam a identificação dos veículos por meio das observações obtidas. É introduzido um critério de identificação que seleciona quais eventos correspondem a identificação dos objetos observados. A partir desta seleção, o sistema está apto para calcular a probabilidade de dois objetos serem o mesmo, dado o conjunto de observações sobre vários objetos. Porém, informações adicionais para o cálculo das probabilidades de transição, como os pontos de entrada e saída dos veículos, devem ser informados ao sistema.

Kettner e Zabih (1999) apresentam uma abordagem Bayesiana para rastrear objetos em rede de câmeras sem sobreposição de campos de visão. O objetivo é reconstruir a trajetória de todos os objetos que são temporariamente visíveis. A solução ótima é um conjunto de trajetórias com maior probabilidade para um dado conjunto de observações. Uma solução aproximada para o máximo *a posteriori* é dado por meio de programação linear. Porém, informação adicional também se faz necessária para o funcionamento efetivo do sistema. O sistema requer que um conjunto de trajetórias possíveis e um conjunto de probabilidades de transição e tempos sejam dados como entrada. Sistemas de segurança em tráfego funcionam bem com câmeras sem sobreposição devido ao fluxo contínuo de veículos em caminhos bem definidos que se estendem por longas distâncias, o que os torna impraticável monitorá-las inteiramente. Na abordagem desenvolvida na tese, os dados são obtidos automaticamente para se descobrir as relações entre alvos e trajetórias, sem a necessidade de intervenção do usuário.

Javed et al. (2000, 2003) apresenta uma solução geral para o problema das câmeras disjuntas e das probabilidades de ocorrência dos eventos. O sistema estima automaticamente a trajetória onde os pedestres devem aparecer em uma câmera secundária, tão logo eles deixem a área de cobertura de uma câmera de referência. São utilizados as relações entre os contornos do campo de visão da câmera para se estabelecer a correspondência entre a observação de um mesmo objeto em múltiplas câmeras. O método assume um conjunto de correspondências iniciais desses parâmetros e utiliza dados fornecidos pelo usuário, acumulando a evidência de padrões no decorrer do tempo. Os padrões são usados para se classificarem seqüências como também para individualizar instâncias dentro de uma mesma seqüência de quadros. Em Javed et al. (2003) os autores usam *Parzen windows* para estimar a probabilidade do objeto entrar no campo de uma certa câmera em um certo tempo, sendo dados a localização, tempo e a velocidade de saída do objeto dos campos das outras câmeras. Embora a abordagem seja dada como geral, esta carece das informações providas por usuários e da aprendizagem do sistema, executada por meio do movimento repetitivo de um alvo através de todos caminhos possíveis no espaço de monitoramento.

No trabalho de Makris et al. (2004), o sistema aprende automaticamente a topologia das câmeras e as zonas de entrada e saída de uma rede de câmeras não calibradas, porém com regiões de sobreposição e sem sobreposição conhecidas. O artigo investiga como o aprendizado não supervisionado de um modelo de atividades para uma rede de câmeras de segurança pode ser desenvolvido a partir de grande conjunto de observações. Essa abordagem favorece o desenvolvimento de um algoritmo que busca encontrar as ligações entre a associação dos campos de visão das câmeras com uma atividade. O algoritmo de aprendizado opera explorando estatisticamente a consistência sobre os dados observados.

Um modelo derivado é usado para se determinar automaticamente a estrutura de uma rede de câmeras e fornecer os meios para o rastreamento de alvos através dos campos de visão disjuntos do sistema.

Recentemente, Bowdene Kaew-Tra-Kul-Pong (2005) desenvolveram uma abordagem que utiliza a consistência estatística dos dados para o aprendizado automático dos períodos de visualização dos objetos. O processo atua sobre um longo período de tempo e requer a calibração geométrica explícita das câmeras. A abordagem viabiliza a predição de onde os objetos deverão aparecer e por quanto tempo. Os períodos de reaparição são usados para modelar as trajetórias físicas dos objetos entre regiões de visualização disjuntas. Porém, o sistema requer um período de treinamento para adequação dos parâmetros estatísticos. Os autores não descrevem como o treinamento é realizado nem uma análise conclusiva sobre os aspectos do aprendizado é desenvolvida.

O aprendizado é uma forma atrativa de calibrar as câmeras da rede, porém esta abordagem sofre de uma série de problemas para sistemas de múltiplas câmeras reais. Abordagens que utilizam aprendizado requerem explicitamente uma fase de treinamento onde no qual a calibração das câmeras é efetuada. Porém, se alguma das câmeras for movida (de maneira intencional ou não), todo o sistema deve ser novamente recalibrado. Outro problema notório é a incapacidade dos sistemas com aprendizado em tratar seqüências de vídeo pré-gravadas, sem informação explícita do treinamento.

Nesta tese faz-se uso de uma metodologia para se associarem trajetórias à alvos em regiões com cobertura disjuntas. A abordagem desenvolvida não utiliza nenhuma forma de aprendizado e relaxa a necessidade de informação adicional fornecida pelo controlador humano. Na abordagem desenvolvida no trabalho, as trajetórias são determinadas no espaço da imagem. O rastreamento é baseado na probabilidade de associação entre as cores e o movimento dos blobs em quadros sucessivos da imagem. O posicionamento das câmeras é obtido por meio do planejamento da movimentação dos robôs em uma estrutura comum a todo o grupo de agentes.

A cor é usada como descritor para rastreamento de objetos na abordagem desenvolvida para a ODC. Porém, o tratamento é dado em tempo real por meio da criação do histograma de cor dos objetos. As informações do histograma são convenientemente associadas as informações do movimento do alvo, tornando o sistema capaz de manter o rastreamento através de uma rede de câmeras cooperativas com campos de visão disjuntos. As abordagens cooperativas para sistemas de múltiplas câmeras são apresentadas em seguida.

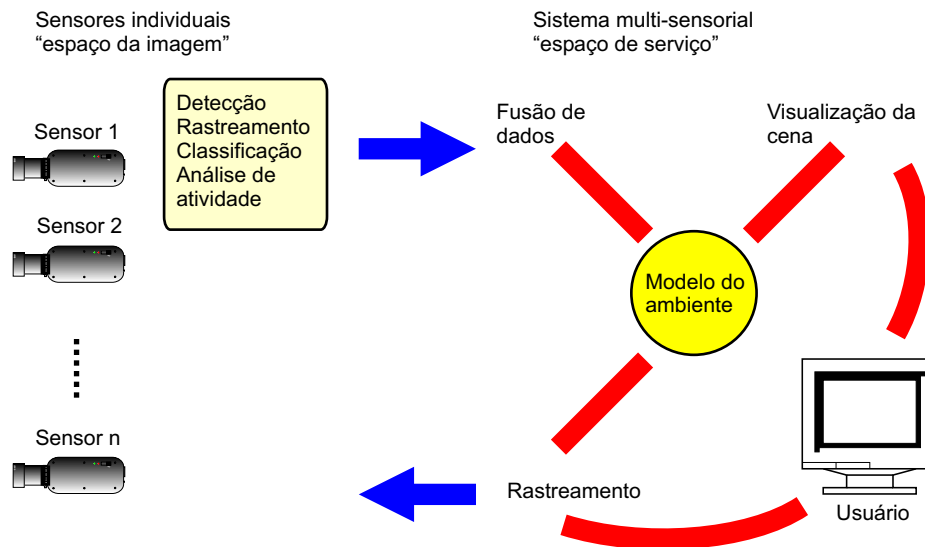


Figura 2.3: Esquema do modelo lógico das tarefas do sistema de segurança de Collins et al. (2000). O sistema é dividido em duas estruturas lógicas denominadas de espaço da imagem e espaço de serviço. No espaço da imagem, tarefas como a detecção do movimento, rastreamento e classificação são realizadas. No espaço de serviço realizam-se tarefas como a fusão de dados sensoriais e adequação das informações a um modelo do ambiente.

2.1.2 Sistemas Multi-Câmeras Cooperativos

Collins et al. (2000) desenvolveram um sistema de segurança com múltiplas câmeras que torna possível a um único operador humano monitorar atividade de pessoas e veículos em um ambiente desordenado, usando uma rede distribuída de sensores de vídeo ativos e cooperativos. No trabalho também se desenvolve um conjunto de algoritmos para a detecção e rastreamento de pessoas e veículos. O rastreamento utiliza um esquema de cooperação entre sensores que objetiva determinar a posição tridimensional dos alvos em relação a um modelo geoespacial do local. Esta informação é apresentada ao operador humano que controla as atividades do sistema por meio de uma interface gráfica do sistema. O esquema lógico do sistema pode ser visto na Figura 2.3. O objetivo geral do sistema é prover automaticamente, coletar e disseminar a informação de segurança para os agentes tomadores de decisão. O sistema é testado em um grande estacionamento, provendo constante vigilância de toda a área monitorada em tempo real. Embora o sistema mantenha o rastreamento de múltiplos alvos, parte do controle ainda é direcionada pelo operador humano e a informação compartilhada pelas câmeras não é utilizada para atualização do posicionamento ou dos parâmetros óticos da câmera.

Micheloni et al. (2005) monta uma rede de câmeras cooperativas para a segurança

por meio de visão em grandes parques de estacionamento para veículos. Esta rede utiliza sub-redes capazes de gerenciar câmeras estáticas e câmeras ativas em uma estrutura hierárquica. O sistema é capaz de rastrear múltiplos alvos simultaneamente em tempo real através das áreas monitoradas. As posições dos diferentes objetos, calculadas a partir das diferentes câmeras, são associadas considerando o fator dinâmico do objeto sob observação em cada leitura do sensor. Os movimentos próximos são obtidos por meio de tarefas direcionadas ao sistema de câmeras ativas. A cooperação é realizada por meio de um sistema de comunicação *multicast* estruturado para transmitir dados para as redes e sub-redes. A informação sobre os objetos a serem rastreados é obtida por meio do sistema de câmeras estáticas, então é usado o sistema de câmeras ativas para operar o reposicionamento. O sistema de câmeras ativas compensa as mudanças de fundo devido a movimentação das câmeras, detecta os objetos móveis e automaticamente rastreia os objetos de interesse.

O rastreamento de objetos utilizando robôs móveis será abordado em maiores detalhes na seção seguinte.

2.2 Rastreamento com Robôs

Diferente das abordagens para o rastreamento vistas anteriormente, o rastreamento com robôs utiliza um conjunto de diferentes sensores para a identificação dos alvos e usa as informações obtidas para atuar no ambiente. Suzuki e Yamashita (1992) foi um dos pioneiros em tratar o problema de rastreamento de um alvo móvel com um robô móvel em ambiente interior com obstáculos poligonais. O trabalho trata da busca e rastreamento realizada com um único robô por um alvo que se desloca e forma um rastro de luz. Todas as condições necessárias e suficientes para a realização da busca são previamente fornecidas, restringindo-as a situações bem determinadas como a não existência de reentrâncias ou saídas do ambiente poligonal. A busca é condicionada a regiões estruturadas, sendo posteriormente estendido para a utilização de vários agentes “observadores” no trabalho de Yamashita et al. (1997).

O rastreamento de um alvo móvel utilizando um robô considerando aspectos geométricos do ambiente, foi desenvolvido por LaValle et al. (1997) e LaValle e Hutchinson (1998). Nesses trabalhos, os autores primeiramente consideraram o caso na qual o movimento do alvo é observável. O algoritmo quebra o intervalo de tempo em $K + 1$ pontos de tempo distintos para os quais a posição do alvo é conhecida e cria um espaço gerado pelo produto Cartesiano do espaço livre com o conjunto dos $k + 1$ inteiros, de 1 até $K + 1$. Para cada k de 1 até $K + 1$, um algoritmo de varredura é utilizado para mapear os pontos no

espaço visível a partir da posição do alvo. Os valores de um funcional de perda para cada ponto no polígono de visibilidade também é pré-computado nesse passo. Então, a cada passo, uma função de custo é calculada para representar o custo acumulado de um ponto de partida em uma configuração inicial e escolher as ações ótimas a cada passo. A partir dessa função de custo determinam-se as ações de movimento que devem ser tomadas pelo observador para se obter uma próxima configuração.

No trabalho de LaValle e Hutchinson (1998), os autores preocuparam-se em abordar a minimização do tempo de fuga de um alvo da área de cobertura dos sensores do robô. Por meio da maximização da velocidade do alvo, pode-se calcular o tempo mínimo que o alvo levará para alcançar a aresta mais próxima na área visível do observador. Essa abordagem considera o problema de rastreamento com o conhecimento prévio das posições do alvo. Os autores não discutem as possíveis extensões do método para inclusão de mais alvos ou de mais robôs observadores.

Técnicas de rastreamento possuem aplicações típicas no futebol de robôs. Nesse contexto, Os robôs não possuem nenhum conhecimento sobre os objetos (bola, obstáculos e os outros robôs), que podem estar, eventualmente, sob oclusão de outros objetos ou obstáculos contidos no ambiente. Uma maneira de se sobrepor essa limitação é fazer com que um agente mantenha um histórico das possíveis localizações dos objetos no ambiente em uma grade de ocupação² As células da grade representam as possíveis localizações dos objetos. Porém, se o ambiente é dinâmico, as células de uma grade de ocupação refletem somente instâncias onde os dados foram registrados. O problema é agravado quando o ambiente é considerado grande, devido, principalmente, ao aumento da complexidade computacional de espaço para a representação e armazenamento da grade de ocupação.

Na abordagem de Han e Veloso (1997), os robôs jogadores não incorporam toda a capacidade de processamento. Em particular, o processamento visual é centralizado em uma CPU que conecta vários clientes de software individuais para o controle dos robôs. O sistema de visão requer que o campo visual cubra todo campo de futebol para calcular o posicionamento dos robôs jogadores em tempo real. Os robôs são identificados pela cor e o rastreamento destes é baseado no Filtro de Kalman Estendido (FKE).

O emprego de múltiplos robôs no ambiente viabiliza a troca de informação e fusão multisensorial. Essa última abordagem é especialmente útil em ambiente dinâmico com restrição de visibilidade. O trabalho de Dietl et al. (2001) mostra métodos para o rastreamento de objetos obtidos pelo sensoriamento de um grupo de robôs. O método é baseado no Filtro de Kalman e localização de Markov. O método é aplicado no contexto do futebol

²*Occupancy grid*, abordagem desenvolvida para a navegação de robôs móveis em Elfes (1990).

de robôs na liga de participantes de tamanho médio.

Schulz et al. (2001) propõem um método para rastrear múltiplos alvos com um robô móvel. A técnica utilizada consiste em uma abordagem baseada em amostragens variantes dos filtros de associação de dados com probabilidade conjunta para o rastreamento de características individuais dos objetos e a solução do problema de correspondência entre as características detectadas com o resultado dos filtros. A abordagem é limitada devido ao custo computacional elevado ao grande número de estados utilizados no filtro de partículas. O modelo assume também a visão global e conhecimento de todo ambiente.

A observação dinâmica de múltiplos alvos caracteriza o trabalho desenvolvido por Stroupe e Balch (2003). Nesse trabalho se apresenta uma aplicação do algoritmo VBORT, um algoritmo distribuído que objetiva maximizar a movimentação do grupo de robôs em tarefas de observação. O algoritmo escolhe o movimento apropriado para cada robô do grupo baseado em um valor ótimo apropriadamente escolhido. O valor é escolhido baseado na redução da estimativa da incerteza sobre todas as observações simultâneas do grupo. O algoritmo é aplicado no caso dinâmico para rastrear e manter sob observação múltiplos alvos móveis.

Jung e Sukhatme (2004) propõem uma abordagem generalizada baseada em região para o rastreamento de múltiplos alvos, podendo ser aplicada tanto em ambiente estruturado quanto em ambiente não estruturado. O rastreamento é visto como um problema de alocação de recursos. Os robôs constroem individualmente regiões virtuais a partir das últimas informações sobre o rastreamento dos alvos realizada por outros robôs do grupo. Cada robô estima a região com maior urgência de ser visitada, indicando, assim, as regiões com presença de alvos e navegam para essa região. Um sistema de rastreamento multi-robô desenvolvido torna possível a distribuição dos robôs em resposta ao movimento dos alvos. A arquitetura do sistema é desenvolvida baseada no modelo de mensagem pública/inscreve (ou jornal/agenda), que aloca o robô mais capacitado para a tarefa de rastreamento.

A abordagens cooperativas são soluções ainda pouco exploradas no contexto de rastreamento de alvos com robôs. Sob o aspecto cooperativo, grupos de robôs devem ser coordenados para rastrear alvos que se deslocam em um ambiente.

O trabalho Matsuyama (1998) introduz robôs móveis em um ambiente de visão distribuída. O sistema fundamenta-se na implementação de um protocolo cooperativo de rastreamento de objetos. A idéia central desse protocolo é o conceito de *agência*. Uma agência é um conjunto de agentes visuais (VA - *Visual Agent*) que visualizam o objeto a ser rastreado e mantém um histórico do rastreamento. Um VA sai da agência se ele não está apto para visualizar o objeto rastreado. Um agente visual novo está apto para a agência assim que um objeto rastreado entre em seu campo de visão. Se nenhum dos

VAs visualizar um alvo, então não existe a agência. Esse sistema possui câmeras montadas sobre um suporte especial com rotações e garantindo um ponto de visada fixa. O sistema rastreia um carro de brinquedo controlado por rádio em ambiente interno, não sendo, portanto, um sistema formado por agentes móveis verdadeiramente autônomos.

Uma abordagem posterior, porém ainda fortemente inspirada nos trabalhos de Matsuyama e Ishiguro, é apresentada por Menegatti e Pagello (2001) e por Menegatti et al. (2003). Essa abordagem utiliza um grupo de robôs móveis dotados com câmeras, sendo estes os próprios VAs. Já que o ponto de visada dos agentes muda continuamente, as mudanças na imagem não são somente mudanças que ocorrem na cena capturada, mas mudanças que ocorrem devido a mudança de posição dos robôs. Assim, essa nova abordagem passa a incluir a incerteza na posição do alvo, uma vez que a posição dos VA não pode ser mais determinada de forma exata e devido aos erros no cálculo da distância relativa entre os VA e o alvo.

Parker (2002) apresenta uma abordagem descentralizada para resolver o problema denominado de "observação cooperativa por múltiplos robôs de múltiplos alvos" (do inglês, *CMOMMT – Cooperative Multi-Robot Observation of Multiple Moving Targets*). A solução para o problema é proposta por meio da combinação do controle de baixo nível com um esquema de controle de alto nível dos robôs. O controle de baixo nível é descrito em termos de campos de força que emanam dos robôs e dos alvos. O controle de alto nível é baseado no ambiente ALLIANCE [Parker (1996)], que trata o controle cooperativo, a tolerância a falhas e viabiliza o ajuste das ações de baixo nível aos membros do grupo. O grande problema deste tipo de metodologia é a falta de provas formais de estabilidade e garantias de que a tarefa será completada. O trabalho deixa de mostrar os níveis de coordenação e de controle dos comportamentos.

O trabalho de Sukhatme et al. (2001) desenvolve dois estudos de casos como exemplos de coordenação e controle de grupos heterogêneos de robôs. O primeiro estudo de caso trata do transporte marsupial, o segundo de um processo de localização cooperativa. O primeiro estudo de caso objetiva demonstrar a capacidade de embarque e desembarque de um sistema multi-robô. O experimento é realizado fazendo com que o helicóptero robô AVATAR (*Autonomous Vehicle Tracking and Reconnaissance*) carregue pequenos robôs a bordo. O AVATAR é usado para detectar e rastrear um intruso em terra. Quando o intruso desaparece do campo de visão do helicóptero, o AVATAR aterrissa próximo a posição onde o intruso foi visto do ar pela última vez. Um pequeno veículo rádio controlado é desembarcado do AVATAR, o qual localiza o intruso na área de busca. O segundo estudo de caso, Sukhatme et al. (2001) trata do processo de localização cooperativa. Embora os estudos de caso demonstrem a viabilidade da cooperação entre robôs, o controle dos robôs

não é totalmente autônomo. O posicionamento do helicóptero sobre o alvo é realizado pelo operador humano. Percebe-se ainda que se relaxadas as restrições impostas ao problema, compromete-se a precisão da abordagem. Alguns aspectos importantes como o posicionamento e a correção dos erros oriundos dos sensores de posicionamento e do GPS não são tratados no trabalho.

Timothy W. McLain e Kelsey (2002) descreve uma abordagem para colaboração entre UAVs e UGVs objetivando a interceptação de um alvo móvel. Robôs aéreos, voando em altura constante, seguem o alvo em terra e os robôs terrestres procuram imitar o comportamento dinâmico dos robôs aéreos. O objetivo do grupo é a interceptação simultânea de alvos. As trajetórias candidatas para cada robô são escolhidas em um grafo de Voronoi derivado de um conjunto de pontos marcados no espaço de navegação. A abordagem, desenvolvida ainda no contexto de simulações, precede de um modelo do ambiente e do conhecimento prévio do posicionamento dos alvos.

2.3 Planejamento de Sensores

O planejamento de sensores tem motivado estudos e proposto soluções em áreas como robótica industrial, robótica móvel, sistemas de controle servo visuais, inspeção visual, síntese automática de objetos, entre outras. Tarabanis et al. (1995) aborda o planejamento de sensores sob o enfoque da visão computacional, oferecendo uma visão geral por meio de uma taxonomia para o problema de posicionamento de câmeras.

O problema conhecido como a "Próxima Melhor Visada" (*Next Best View Problem – NBV*) objetiva encontrar as melhores posições para varredura dos sensores sobre um dado objeto e se obter todas as informações geométricas referente às superfícies que o formam. Kemmotsu e Kanade (1994) mostra uma abordagem para o NBV para determinar a postura (posição e orientação) de um objeto poliédrico por meio de sensores de profundidade. O problema é abordado por meio da determinação do melhor posicionamento dos sensores no espaço de trabalho.

Mais recentemente, Chene Li (2004) apresenta um abordagem para o NBV visando o posicionamento de sensores para manipuladores robóticos. Nesse método, o sensor deve ser movido de uma posição a outra sobre o objeto de interesse. Durante esse movimento, toda característica visual de interesse é observada. A tarefa envolve a determinação do posicionamento ótimo do sensor e o menor caminho tridimensional entre esses pontos de observação. Durante o processo de planejamento de sensor, as características visuais dos objetos são reamostradas como informações pontuais e associadas aos vetores normais das

superfícies observadas. Um grafo de posicionamento ótimo de sensores é obtido por meio de algoritmo genético com base em um critério min-max de avaliação.

O posicionamento estático de sensores tem sido empregado em problemas relativos às redes de sensores sem fio [Chen et al. (2004); Hsieh (2004); Kotecha et al. (2005); Ssu et al. (2005)]. Chiue Lin (2004); Lin e Chiu (2005) apresentam um algoritmo para o posicionamento de sensores em uma rede de sensores sem fio com restrições de limitação de custos e cobertura completa do ambiente de monitoramento. O algoritmo tem como cenário uma grade regular onde cada ponto da grade comporta exclusivamente um sensor. O posicionamento ótimo dos sensores é alcançado quando, para uma dada configuração de sensores posicionados na grade, as posições de alvos de interesse sejam obtidas a partir de qualquer ponto da grade do campo de sensores. Em abordagem similar, Wong et al. (2004) propõe uma técnica de auto-organização para aprimoramento da cobertura de uma rede de sensores sem fio.

Sob enfoque mais teórico, o posicionamento de sensores tem sido abordado por meio do problema da galeria de arte e suas variações [ORourke (1983, 1987); Liaw et al. (1993)]. Em suas diferentes caracterizações, o problema da galeria de arte inclui polígonos com buracos, guardas que podem ver em cone ou com visão omnidirecional, galerias de arte tridimensionais e guardas com capacidade de mobilidade [Bose et al. (1997); Bjorling-Sachs (1998); Guibas et al. (1997, 1999); Tóth (2000); Hlin (2001); Simov et al. (2002); LaValle et al. (2002); Tóth (2002a,b, 2003); Sachs et al. (2004)].

Guibas et al. (1999) trata do problema da busca por um alvo cuja a posição é desconhecida e se move arbitrariamente rápido, fazendo uso de um grupo de robôs móveis. O algoritmo desenvolvido decompõe o espaço identificando curvas críticas onde ocorrem mudanças de visibilidade. Essa decomposição determina áreas no espaço livre (áreas descontaminadas) onde o fugitivo pode ser encontrado. Um valor binário é assegurado as arestas de cada célula da decomposição para indicar se a área já fora descontaminada (o valor “0” indica que a área está limpa e o valor “1” indica que o fugitivo pode ser encontrado naquela área). Quando uma solução for encontrada, todos os valores são trocados para “0”.

O conjunto de células construído pelo processo de decomposição é usado para formar um grafo com os vértices representando as células e as arestas ligando células adjacentes. Duas células estão conectadas no grafo se e somente se elas possuem em comum uma mesma aresta no espaço livre. Cada célula corresponde ao conjunto de condições de visibilidade. Quando o robô se move de um vértice do grafo para outro, as condições de visibilidade mudam, isto é, novas informações de visibilidade são inseridas e as informações anteriores são removidas. O grafo de células é utilizado para construir um grafo no espaço de

informações, o qual pode ser percorrido para se determinar uma solução.

A construção de grafos a partir da decomposição espacial do ambiente é uma solução interessante que foi adotada para o planejamento de posicionamento das câmeras para a ODC. Porém, de maneira diferenciada, na abordagem desenvolvida aqui impõe-se restrições de visibilidade a construção do grafo, procurando estabelecer o posicionamento dos robôs dentro do espaço de percepção alcançável das câmeras, agrupando em uma única estrutura o posicionamento eficiente da câmera à habilidade de navegação do robô.

Spletzer e Taylor (2003) objetiva controlar um grupo de robôs de tal maneira que a configuração do grupo otimize a estimativa das medidas derivadas das observações dos robôs. Nesse trabalho se desenvolve uma abordagem teórica para o problema do posicionamento de sensores. Os autores mostram também uma estratégia computacional prática para o posicionamento dos robôs, baseada em filtro de partículas. Estas idéias são aplicadas a tarefas de rastreamento de alvos por meio de simulações e por experimentos com robôs reais. Os resultados obtidos utilizam robôs com câmeras omnidirecionais, com visão panorâmica ampla sem impor nenhuma restrição à área de cobertura dos robôs. Embora o método não requiera nenhum conhecimento *a priori* do ambiente, utiliza as técnicas convencionais de planejamento de movimento, como a possibilidade da ocorrência de mínimos locais.

Navarro-Serment et al. (2004) descrevem um método para a observação de alvos por um grupo de robôs móveis equipados com câmeras. Busca-se posicionar as câmeras de tal maneira que estas cubram da melhor maneira possível os alvos. O problema consiste em posicionar e manobrar as câmeras por meio da avaliação da informação da posição corrente que pode manter a melhor visão frontal do alvo. A idéia geral é distribuir essa tarefa por vários robôs, uma vez que o consumo de energia seria excessivo para um único robô manter a constante visualização do alvo. Os autores usam uma função de avaliação baseada na posição do robô e na estimação da posição do alvo para encontrar a melhor posição da câmera. Embora o método desenvolvido apresente uma abordagem interessante para o posicionamento de câmeras montadas em robôs móveis, os resultados são preliminares e baseados em simulações. A função de posicionamento não é avaliada na presença de obstáculos e o método não incorpora funções mecanismos de navegação ou desvio de obstáculos.

A revisão bibliográfica apresentada neste capítulo oferece uma visão geral do estado da arte nas principais áreas correlacionadas com o problema da ODC. Muito das idéias aqui apresentadas serão adequadas à metodologia proposta nos capítulos restantes. Em seguida, o próximo capítulo irá explorar em maior profundidade a ODC, fundamentando o problema e oferecendo uma descrição geral do arcabouço desenvolvido.

Capítulo 3

Metodologia

It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories instead of theories to suit facts.

*Sherlock Holmes, em "The Whole Art of Detection",
por Sir Arthur Conan Doyle (1859–1930), médico e escritor inglês.*

Neste capítulo desenvolve-se a metodologia proposta. Aqui serão detalhados os métodos para o rastreamento visual, para a associação e reconhecimento entre observações e trajetórias e as estratégias de planejamento de posição. Primeiramente, o problema será formalizado na Seção 3.1. Uma descrição da arquitetura funcional do rastreador é mostrado na Seção 4.2, onde cada componente será discutido separadamente. Em seguida, na Seção 3.3.2, será descrito o procedimento para a associação de observações e trajetórias entre câmeras com campo de visão disjuntas, para garantir o posicionamento correto das câmeras. Por fim, na Seção 3.3.3 será detalhado o esquema de planejamento de posição desenvolvido.

3.1 Formalização do Problema

Considere um sistema constituído de n câmeras calibradas C_1, C_2, \dots, C_n , dotadas da habilidade de movimento de translação e rotação planar que monitoram uma região $S \subset \mathbb{R}^2$. Considere também a existência de m alvos a_1, a_2, \dots, a_m que se movimentam em uma direção sobre o espaço S . Considere, ainda,

$$T_{ij} = \{(x_1, y_1, h_1), (x_2, y_2, h_2), \dots, (x_k, y_k, h_k)\},$$

a trajetória de um alvo a_j rastreada por uma câmera C_i do sistema, onde (x_t, y_t) são as coordenadas que descrevem a trajetória do alvo em relação a um sistema de referência global $\{L\}$ e h_t é a informação de cor, ambas tomadas em instantes de tempo $t = 1, \dots, k$. Supondo-se que o conjunto de pontos (x_t, y_t) seja aproximado por um modelo linear $y = ax + b$, representado pelos parâmetros $\bar{T}_{ij} = (a, b)$, e que \bar{h} represente o valor médio observado da cor do alvo, então conjunto de parâmetros composto por \bar{T}_{ij} e \bar{h} forma uma observação

$$k_{ij} = \{\bar{T}_{ij}, \bar{h}_{ij}\} \in \mathcal{K},$$

onde \mathcal{K} é o conjunto de observações obtidas pelas câmeras do sistema.

Seja C_l outra câmera do sistema suficientemente afastada de C_i que compartilhe a informação k_{ij} por meio de um sistema de comunicação. O problema consiste em estimar k_{ij} e determinar, por meio dessa estimativa, a postura de C_l em S de tal maneira que as observações k_{lj} estimadas por C_l mantenham uma relação de correspondência entre as observações k_{ij} , obtidas por C_i . A postura de C_l deve ser determinada respeitando-se os critérios de visibilidade impostos pelas restrições da câmera e do ambiente por meio de uma função de observação $V_c(P, \theta)$ sobre um ponto P e uma orientação θ . O problema pode ser assim formalizado:

Dados,

$S \subset \mathbb{R}^2$, uma área de monitoramento;

$C_i = \{C_1, \dots, C_n\}, i = 1, \dots, n$, um grupo de câmeras;

$A_j = \{a_1, \dots, a_m\}, j = 1, \dots, m$, um grupo de alvos;

$T_{ij} = \{(x_1, y_1, h_1), \dots, (x_k, y_k, h_k)\}$, o resultado do rastreamento de uma câmera C_i sobre um alvo a_j ;

$V_c(P_l, \theta_l) : \mathbb{R}^3 \rightarrow \mathbb{R}$, uma função de observação que associa a posição P_l e a orientação θ_l de uma câmera C_l do sistema a um valor real.

Estimar,

$k_{ij} \in \mathcal{K}$, a observação do alvo a_j pela câmera C_i e

$k_{lj} \in \mathcal{K}$, a observação do alvo a_j pela câmera C_l ,

e determinar

$$P_l(x_l, y_l), \theta_l,$$

tal que $V(P_l, \theta_l)$ seja mínimo e

se tenha $k_{lj} \sim k_{ij}$, para qualquer alvo $a_j \in A$ percebido pelas câmeras.

O símbolo \sim indica uma relação de similaridade entre os padrões contidos nos campos de visão das câmeras. As câmeras estão calibradas e referenciadas em relação a um sistema de referência global $\{L\}$. Isso significa que toda informação visual obtida ou transmitida pelas câmeras através do sistema são mapeadas em relação ao sistema de referência $\{L\}$.

O aspecto cooperativo do problema é devido a troca assíncrona de informações entre os robôs por meio de um sistema de comunicação e a conseqüente ação dos agentes a partir das informações compartilhadas. Um esquema de planejamento deve posicionar as câmeras dentro de seu espaço de percepção, objetivando garantir a observação dos alvos.

3.2 Suposições

Para melhor delimitar o escopo do problema, sem modificar sua natureza, assumem-se as seguintes suposições:

1. Em relação as câmeras:

- As câmeras são calibradas, seus parâmetros intrínsecos previamente conhecidos, ou seja, para cada câmera C_i , conhece-se previamente a distância focal f_i , o centro de projeção $Cp(x_u, y_v)$ e a dimensão $u \times v$ do plano da imagem.
- Todas as câmeras conhecem sua posição em relação a um sistema de referência inercial $\{L\}$. Em outras palavras, são conhecidas *a priori* as matrizes de rotação \mathbf{R}_i e os vetores de translação \mathbf{t}_i de cada sistema de referência da câmera $\{C_i\}$ em relação ao sistema de referência global $\{L\}$.
- Todos os componentes do sistema estão contidos na área de cobertura do sistema de comunicação. Seja r o raio de cobertura do sistema de comunicação definido por meio do disco de centro $(x_r, y_r) \in \mathbb{R}^2$ que define a região de cobertura de comunicação R . Então, para toda câmera $C_i \in S$, onde S é a região de monitoramento, sempre teremos $C_i \in R$, ou seja, $S \subset R$.

2. Em relação aos alvos:

- Os alvos movem-se de forma suave. Explora-se suavidade a redundância do movimento que pessoas ou veículos tendem a seguir em uma determinada via ou caminho, como em ruas, calçadas ou corredores. As trajetórias são aproximadas por um modelo linear.
- Os alvos não são evasivos ou maliciosos, ou seja, os alvos não objetivam fugir, se esconder das câmeras, nem possuem "inteligência" para burlar, de alguma maneira, o sistema.
- Os alvos são identificados de forma única. Marcadores ou rótulos visualmente identificáveis, como a cor, a geometria ou a forma do objeto estão associados como rótulos em cada alvo. Precisamente, os alvos são unicamente identificados por meio do valor da matiz de cor h .

3.3 A Metodologia

O princípio da metodologia aqui proposta consiste na subdivisão das atividades inerentes à tarefa de observação dinâmica cooperativa, divididas em três módulos principais:

- Rastreamento visual: onde se identifica, faz-se a extração das características visuais e predição do movimento dos alvos. O módulo de rastreamento visual recebe como entrada uma seqüência de imagens e fornece como saída um conjunto de informações visuais obtidas pelo processamento de imagens e rastreamento dos alvos.
- Associação de trajetórias: onde formam-se as trajetórias, identificam-se os alvos e associam-se as observações obtidas pelo sistema. O módulo de associação de trajetórias recebe como entrada o conjunto de informações visuais obtidas pelo rastreamento visual e fornece como saída modelos de trajetórias para o módulo de planejamento de posição.
- Planejamento de posição: responsável pela escolha da movimentação e posicionamento das câmeras do sistema. Este módulo recebe como entrada os modelos de trajetórias e fornece como saída as posições para observação das trajetórias.

A Figura 3.1 apresenta a arquitetura funcional da metodologia. Nos retângulos destacam-se os módulos principais e as setas indicam o fluxo de informação entre os

componentes do arcabouço metodológico. O esquema de cooperação entre os módulos de associação de dados e planejamento de posição está destacado por meio de uma linha tracejada, indicando a troca de informações entre esses dois módulos propiciam a organização do sistema para a observação da trajetória dos alvos. Os módulos que compõem a metodologia estão apresentados em detalhes nas seções seguintes.

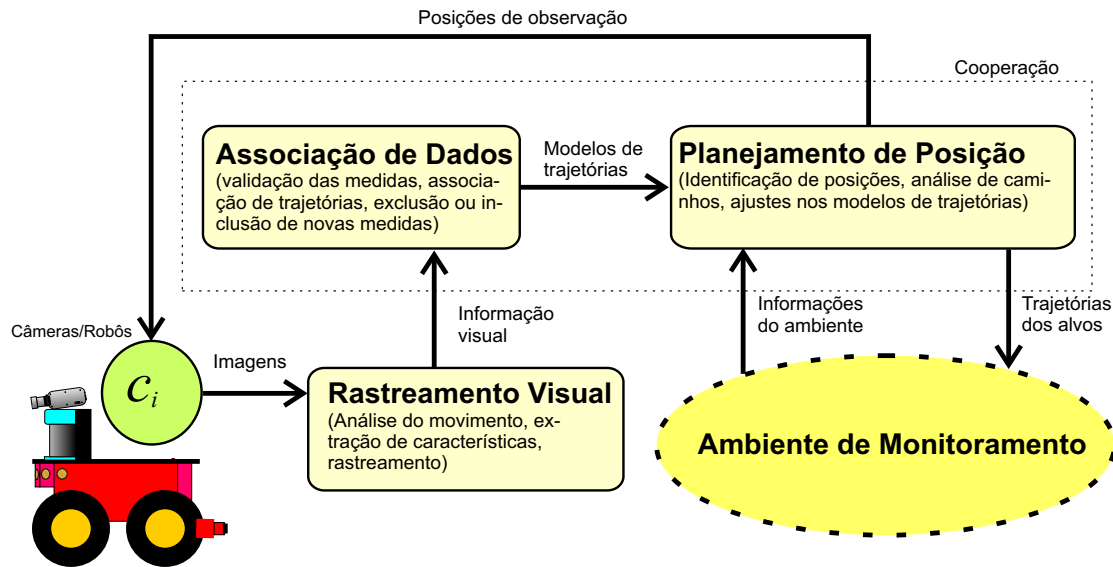


Figura 3.1: Arquitetura funcional do sistema de observação dinâmica. As caixas representam os principais módulos. As setas indicam o fluxo de informações entre os módulos do sistema e a área de monitoramento. As câmeras C_i estão representadas pelo círculo posicionado na parte inferior a esquerda da figura.

3.3.1 O Rastreamento Visual

A arquitetura do rastreador visual é dividida em três módulos principais, cada um projetado para a execução de um conjunto de tarefas específicas. Os módulos do rastreador, são:

1. O módulo de aquisição e processamento de imagens;
2. O módulo de rastreamento ou filtro de partículas;
3. O módulo de associação de dados.

Cada módulo possui seus limites e interfaces bem definidas, viabilizando a implementação e testes de diferentes procedimentos de maneira independente. Essa abordagem no projeto da arquitetura funcional do sistema de rastreamento, propicia a flexibilidade no

desenvolvimento do sistema, tornando possível a modificação em um módulo sem alterar significativamente o sistema como um todo.

A Figura 3.1 ilustra a arquitetura proposta e as interações entre seus principais módulos. Central a todo o processo está um filtro de partículas descentralizado, capaz de rastrear múltiplos alvos simultaneamente e compartilhar seus dados para o planejamento e identificação dos alvos e trajetórias.

3.3.1.1 Módulo de Processamento de Imagens

Para haver a ODC, o sistema deve ser suficientemente hábil para identificar o movimento dos alvos e rastreá-los. Quando o sistema inicia o processo de rastreamento de um alvo, o módulo de processamento de imagens é utilizado para obter informações sobre a cena e sobre os objetos. Nesse módulo é que se identificam os alvos, se obtém o modelo de fundo, as características visuais e do movimento para o rastreamento. A filtragem das imagens para a atenuação do ruído também é realizada nessa etapa.

O diagrama de blocos da Figura 3.2 mostra os módulos funcionais do sistema de rastreamento e seus principais componentes. Cada módulo será tratado em particular nas subseções seguintes.

A identificação do movimento é realizado por meio da técnica de subtração de fundo. Algumas adaptações foram aqui implementadas visando a aquisição das imagens e o processamento em tempo real. Estas adaptações estão descritas em seguida.

Assumindo-se que a câmera/robô esteja parada e que não existem movimentos significativos na cena, o modelo de fundo é obtido utilizando os n primeiros quadros (*frames*). Para cada quadro obtido aplica-se um filtro passa baixa com a finalidade de reduzir o ruído na imagem e a variabilidade da cor do objeto, preparando o sistema para a obtenção do modelo de fundo.

Para um dado quadro $\mathbf{I}_{f \times q}$ com intensidade de pixel $I_f(i, j), i = 1, \dots, p$ e $j = 1, \dots, q$, o modelo de fundo \mathbf{B} pode ser obtido por meio da seguinte relação recursiva [Chen et al. (2001)].

Para cada quadro $\mathbf{I}_f, f = 2, \dots, n$, tem-se:

$$B_1(i, j) = I_1(i, j), \quad (3.1)$$

$$B_f(i, j) = B_{f-1}(i, j) + \frac{I_f(i, j) - I_{f-1}(i, j)}{2}, \quad (3.2)$$

onde f denota o quadro da imagem e \mathbf{B}_f é o modelo de fundo.

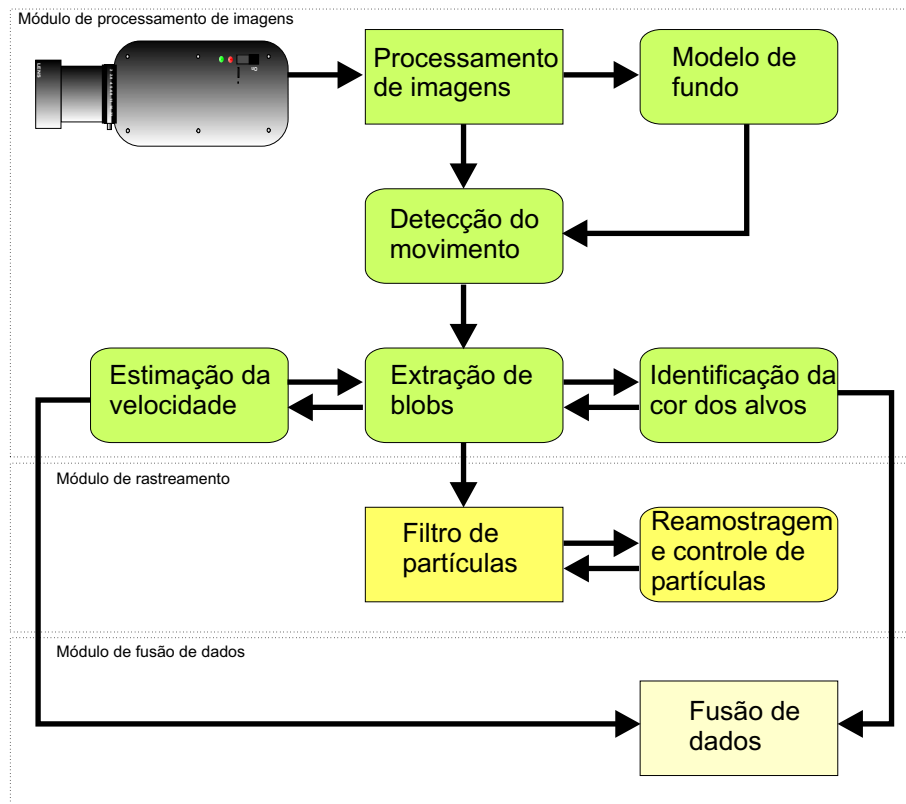


Figura 3.2: Diagrama de blocos do sistema de rastreamento visual. O sistema é decomposto funcionalmente em três módulos principais: processamento de imagens, rastreamento e fusão de dados.

Uma vez determinado o modelo de fundo, cada novo quadro pode ser subtraído deste para se determinar a movimentação dos alvos. Assumindo a premissa que a câmera é fixa e que se pode manter o mesmo modelo de fundo, o procedimento de detecção de movimento consiste em se comparar cada pixel da imagem $I_f(i, j)$ com os pixels do modelo de fundo $B(i, j)$. Para se verificar se um pixel da imagem corresponde a um pixel de um objeto em movimento, basta verificar se

$$|I_f(i, j) - B(i, j)| > \tau. \quad (3.3)$$

O limiar τ é um valor que pode ser ajustado para a obtenção de maiores ou menores regiões de movimentação na seqüência de imagens.

Assim, para a identificação dos alvos, consideram-se as regiões que apresentam um determinado número fixo de pixels. Grupos de pixels que formam regiões que atendam a este critério são identificados como alvos e armazenados em uma lista de alvos Q mantida durante todo o processo de rastreamento do alvo. Regiões formadas por uma quantidade

de pixels menor que a determinada, não são levadas em consideração e nenhum processo de rastreamento é associada a esta.

A Lista Q tem uma grande importância no rastreamento. Cada novo alvo válido é incluído na lista e as informações extraídas sobre ele são atualizadas até que o alvo saia do campo de visão da câmera. A lista Q mantém a identificação e todo o histórico sobre o rastreamento de cada alvo, servindo como a base de dados para a obtenção das observações efetivadas sobre os alvos.

3.3.1.2 Extração de Características

As características rastreadas pelo sistema são a cor do alvo e os parâmetros de movimento. Assim, simultaneamente à identificação do alvo, faz-se necessária a obtenção da informação de cor e dos parâmetros de movimento associado a este.

A cor do alvo é obtida por meio do cálculo do histograma de cor (descrito em detalhes no Apêndice B). O histograma descreve a distribuição das cores no objeto, independente de sua forma e orientação, definindo, assim, uma representação da cor para o alvo. A informação sobre a cor do alvo é obtida pelo valor da matiz de cor, fornecida por meio do modelo de cor HSI. Com o histograma de cor montado, pode-se utilizar qualquer métrica de similaridade de cor, que também estão mostradas no Apêndice B, ou se obter diretamente a média a partir do histograma.

Outra característica visual obtida da seqüência de imagens são as inerentes ao movimento. Para a isso, se utiliza um modelo de atualização do fundo \mathbf{B} . A última atualização de fundo \mathbf{B}_{k-1} e o último quadro \mathbf{I}_{k-1} são espacialmente alinhados em relação ao quadro corrente \mathbf{I}_k . Considera-se que para a obtenção deste último quadro que a câmera esteja estática e que os objetos na cena não exerçam nenhum movimento significativo.

Seja $\mathbf{P} = [X, Y, Z]^T$ o vetor de características do alvo, dado pelo centróide das regiões de movimento (ou *blobs*) em relação a um sistema de referência estacionário $\{L\}$, e seja $\mathbf{p} = [x, y]^T$ a projeção de \mathbf{P} no plano de imagem da câmera, então os vetores de deslocamento nas direções x e y podem ser obtidos diretamente no processo de subtração de fundo, sendo tomados como a distância percorrida pelo centróide do *blob* em dois quadros consecutivos.

Assumindo para simplificação da notação que a distância focal f é unitária (porém obtida com precisão razoável no processo de calibração) e que o modelo de projeção da câmera é perspectiva, sabe-se que a velocidade do vetor de características está associada a velocidade do objeto por meio de

$$\dot{\mathbf{x}} = \left(x \frac{v_z}{Z} - \frac{v_x}{Z} \right) + [xy\omega_x - (1 + x^2)\omega_y + y\omega_z], \quad (3.4)$$

$$\dot{\mathbf{y}} = \left(y \frac{v_z}{Z} - \frac{v_y}{Z} \right) + [(1 + y^2)\omega_x - xy\omega_y - x\omega_z]. \quad (3.5)$$

As Equações 3.4 e 3.5 representam o fluxo óptico do vetor de características \mathbf{p} .

3.3.1.3 Módulo de Rastreamento

A estimação dos estados é realizada por um filtro de partículas descentralizado. Como observado nos capítulos anteriores, um filtro de partículas é uma ferramenta probabilística ideal para a estimação recursiva do vetor de estados quando não se dispõe de informações suficientes sobre o modelo de movimento do objeto.

Assim, objetivando estabelecer um relacionamento entre a carga computacional e a eficiência da estimação em tempo real, algumas adaptações foram incluídas no algoritmo do filtro de partículas convencional, tornando a abordagem desenvolvida rápida e computacionalmente eficiente para o rastreamento de múltiplos alvos em tempo real. Estas adaptações serão descritas com mais detalhes a seguir.

Pode-se rastrear múltiplos alvos com um único filtro de partículas, um filtro conjunto, por meio da concatenação da configuração do sistema em um único vetor de estados. Infelizmente, para aplicações em tempo real um filtro conjunto torna-se impraticável devido ao crescimento exponencial da complexidade na dimensão do espaço de estados.

O filtro de partículas requer dois modelos para o processo de estimação: o modelo de ação, ou de movimento, e o modelo de medidas, ou de sensor. Um objeto a ser rastreado é caracterizado como um vetor de estados $\mathbf{x} \in \mathbf{X}$. Como este estado é geralmente desconhecido, o conhecimento sobre este último é representado por meio de uma distribuição de probabilidade $p(\mathbf{x})$.

Na abordagem multi-alvo desenvolvida, considerou-se um vetor de estados para cada alvo independentemente. Assim, em um instante de tempo t , para m alvos identificados, o estado do sistema pode ser descrito por

$$\mathbf{x}_t = (\mathbf{x}_t^{(1)}, \mathbf{x}_t^{(2)}, \dots, \mathbf{x}_t^{(m)}), \quad (3.6)$$

onde,

$$\mathbf{x}_t^{(i)} = [x, y, Vx, Vy, h]^\top, \quad (3.7)$$

na qual o par ordenado (x, y) representa a posição estimada de um alvo individual no plano da imagem, Vx e Vy é a velocidade linear nas direções x e y , respectivamente, no plano da imagem e h é a informação de cor.

Assim, o modelo dinâmico para a representação dos estados de cada alvo em um

instante de tempo t pode ser descrito como:

$$\begin{aligned}
 x_t^{(i)} &= x_{t-1}^{(i)} + Vx_{t-1}^{(i)}\Delta t + u_x, \\
 y_t^{(i)} &= y_{t-1}^{(i)} + Vy_{t-1}^{(i)}\Delta t + u_y, \\
 Vx_t^{(i)} &= Vx_{t-1}^{(i)} + u_{Vx}, \\
 Vy_t^{(i)} &= Vy_{t-1}^{(i)} + u_{Vy}, \\
 h_t^{(i)} &= h_{t-1}^{(i)} + u_c,
 \end{aligned} \tag{3.8}$$

onde u_x e u_y são os ruídos adicionados as posições, u_{Vx} and u_{Vy} são os ruídos adicionadas às velocidades e u_c é o ruído adicionado à informação de cor.

Assume-se ainda que um único observador, uma câmera localizada em uma posição fixa e conhecida, captura uma seqüência de imagens \mathbf{I} de uma região de interesse S , e que em uma etapa de tempo t , o rastreador recebe m vetores de observação dados por:

$$\mathbf{y}_t = [\mathbf{y}_t^{(1)}, \mathbf{y}_t^{(2)}, \dots, \mathbf{y}_t^{(m)}]^\top, \tag{3.9}$$

com o valor do número de observações m variando no tempo.

O modelo de observação fornece as coordenadas (x, y) dos centros dos *blobs* no espaço da imagem associada a informação de cor h . As medidas são assumidas como independentes entre si e podem ser originadas tanto dos alvos verdadeiramente observados como de observações espúrias. Os vetores de ruído $\mathbf{u}_t^{(i)}$ para a transição de estado e $\mathbf{v}_t^{(i)}$ para as observações são supostos como brancos e independentes no tempo e no espaço.

O método desenvolvido segue os passos elementares do filtro de partículas:

1. Iniciação de alvos.
2. Propagação.
3. Observação e fusão.
4. Seleção

A iniciação dos alvos começa a partir da obtenção do conjunto de partículas para cada alvo detectado

$$\{\mathbf{x}_{t-1}^{(i,j)}, w_{t-1}^{(i,j)}\}. \tag{3.10}$$

O peso amostrado w_{t-1} representa a verossimilhança de uma determinada amostra pertencer a uma localização verdadeira do alvo. Esta verossimilhança é calculada por meio da comparação entre os dados obtidos no módulo de processamento de imagens com os dados obtidos *a priori* de posição e cor.

A inicialização dos alvos é realizada no instante de detecção do movimento, assim que se estabelece a formação do *blob* no espaço da imagem. Assim que é detectado um *blob*, o conjunto de partículas é iniciado com N amostras distribuídas sobre todos os possíveis estados iniciais, identificados a partir da formação da área dos *blobs* no espaço da imagem e espalhadas sobre estas regiões com uma dispersão inicial relativa a $2A_b$, sendo A_b a área inicial do *blob*. Cada amostra recebe um peso inicial de $1/N$.

As informações de cor e movimento no espaço da imagem são usadas para se obter a verossimilhança das partículas amostradas e identificar regiões de provável movimento no espaço de estados. A idéia básica é a verificação da probabilidade de associação entre a cor e o movimento dos alvos em quadros sucessivos da seqüência de imagens. O processo se repete a cada iteração do método.

Ao se agruparem os pixels na formação do *blob*, obtém-se sucessivas medidas da cor e de movimento. Estas medidas são posteriormente associadas às novas observações na seqüência de imagens, servindo como um indicador de iniciação ou de saída do alvo de cena.

No caso da iniciação do alvo, o número de partículas aumenta consideravelmente devido ao aumento da verossimilhança entre cores e movimento. A cada novo quadro regiões com maior verossimilhança são obtidas, gerando um número de partículas maior sobre a região de movimento estimada. O alvo é então incluído na lista de alvos válido Q e seu estado passa a ser rastreado.

A verificação de saída de cena de um alvo, ou sua finalização, é realizada de forma similar ao processo de iniciação. Ao sair do espaço da imagem, o alvo deixa de fazer associação de cor e movimento nos quadros sucessivos, gerando um número de partículas cada vez mais baixo. Logo, se um pequeno número de partículas for associada ao vetor de observação de um *blob* já iniciado, então o alvo está ocluído ou saiu de cena.

Esta técnica de iniciação das partículas evita a perda de alvos durante os períodos de movimento e viabiliza a investigação das regiões de movimento sob a abordagem Bayesiana.

A etapa de propagação ocorre por meio da aplicação do modelo de transição de estados do tipo

$$\mathbf{x}_t = f(\mathbf{x}_{t-1}, \mathbf{w}_{t-1}), \quad (3.11)$$

onde \mathbf{w}_{t-1} é um vetor de ruído adicionado, com distribuição conhecida e que representa as incertezas do sistema. A função $f(\mathbf{x}_{t-1}, \mathbf{w}_{t-1})$ é descrita por meio do modelo dinâmico

dado pela Equação 3.8 sobre os estados correntes do sistema.

O processo de observação compreende a estimativa do estado de cada alvo e a ponderação de cada elemento de \mathbf{x}_t em termos das medidas de cor e movimento observadas. A dinâmica da distribuição de probabilidade é descrita pelo mapeamento de $p(\mathbf{x}_{t-1})$ para $p(\mathbf{x}_t)$ por meio de

$$p(\mathbf{x}) = \int p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1})d\mathbf{x}, \quad (3.12)$$

onde $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ representa a função de densidade de probabilidade do processo. A função de densidade $p(\mathbf{x}_t)$ é modelada como um processo Markoviano, o qual depende somente da distribuição anterior $p(\mathbf{x}_{t-1})$.

A estimativa do estado dos alvos é obtida a partir de uma aproximação da Equação 3.12. Esta aproximação é dada por meio de

$$\tilde{\mathbf{x}}_{k+1}^{(i)} \leftarrow \frac{1}{p} \left(\sum_{j=1}^p \mathbf{x}_{k+1}^{(j)} \right). \quad (3.13)$$

Um cuidado constante a cada iteração do algoritmo é o controle da geração de partículas. O objetivo deste controle é evitar a degeneração do conjunto de partículas. Este problema ocorre quando poucas partículas apresentam peso considerável em relação ao peso das outras partículas do conjunto.

Objetivando evitar a degeneração das partículas e o processamento em tempo real, desenvolveu-se um procedimento adaptativo para a reamostragem. Este procedimento caracteriza-se pelo processo de seleção do filtro de partículas convencional. O procedimento é executado quando os pesos das partículas estão a baixo de um determinado limiar.

Para se reamostrar o conjunto de partículas, emprega-se o procedimento de ordenação e obtém-se, a partir deste conjunto ordenado, um percentual fixo das k maiores verossimilhanças. As partículas obtidas são adicionadas na formação de um novo conjunto de partículas geradas aleatoriamente, criando novas hipóteses no espaço de estado dos alvos.

A etapa de ponderação e a conseqüente atualização dos pesos do filtro de partículas é executada por meio de

$$\mathbf{w}_{k+1}^{(i)} \leftarrow p(\mathbf{y}_t|\mathbf{x}_{(k+1)}), \quad (3.14)$$

onde $p(\mathbf{y}_t|\mathbf{x}_{(k+1)})$ é a fdp obtida pelo processo de fusão entre a cor e a distância percorrida pelo alvo, que será descrito em seguida.

Se a quantidade de alvos observados for maior que um, então faz-se necessário a

utilização de um método de fusão de dados para a formação de trajetórias. O objetivo é garantir que a cada observação forme corretamente uma trajetória por meio da associação das características de cor e movimento do alvo.

A idéia geral é comparar informação de um modelo de cor e de movimento com modelos similares extraídos da imagem como candidatos a alvos. Quanto menor for a discrepância entre o modelo de referência e o modelo do candidato, maior será a probabilidade de localização do alvo na região da imagem onde se observa o eventual alvo.

A associação de dados é efetivada por meio de uma abordagem probabilística. Emprega-se uma função exponencial parametrizada por meio da associação da distância entre as cores e o deslocamento do alvo. A aplicação da função exponencial garante a fusão entre as informações de cor e de movimento e a flexibilidade para o ajuste conveniente dos parâmetros, como será mostrado em seguida.

Seja

$$d_c^{(j,l)}(\mathbf{h}_t^{(j)}, \mathbf{H}_t^{(l)}), \quad (3.15)$$

a métrica de cor, definida como:

Definição 3.1 (Distância entre cores) – A distância do espaço de cor é obtida por meio do histograma de cor do alvo e o valor de cada hipótese de cor, caracterizada por meio da diferença absoluta entre as hipóteses de cor $H_t^{(l)}$ para cada partícula l e as observações do histograma $h_t^{(j)}$,

$$d_c^{(j,l)}(h_t^{(j)}, H_t^{(l)}) = |h_t^{(j)} - H_t^{(l)}|. \quad (3.16)$$

■

Seja ainda,

$$d_{pos}^{(j,l)}(\mathbf{Y}_t^{(j)}, \mathbf{y}_t^{(l)}), \quad (3.17)$$

a distância entre cada hipótese de posição e o vetor de observação no intervalo de tempo t , definida como

Definição 3.2 (Deslocamento do Alvo) – A distância percorrida pelo alvo, ou o deslocamento do alvo, em relação ao conjunto de observações é dada por:

$$d_{pos}(y_t^{(j)}, Y_t^{(l)}) = \|(\mathbf{y}_t^{(j)} - Y_t^{(l)})^2\|, \quad (3.18)$$

onde $\mathbf{y}_t^{(j)}$ é a posição observada j no instante de tempo t e $Y_t^{(l)}$ é o vetor de observação com a hipótese de posição de cada partícula l .

Então, a probabilidade de cor pode ser obtida por meio de

$$P_c(\mathbf{y}_t^c | \mathbf{x}_t) = e^{-\frac{d_c}{k_1}}, \quad (3.19)$$

onde k_1 é um parâmetro da fdp.

De maneira análoga, a probabilidade de movimentação em relação a distância percorrida pelo alvo pode ser encontrada como

$$P_{pos}(\mathbf{y}_t^{pos} | \mathbf{x}_t) = e^{-\frac{d_{pos}}{k_2}}, \quad (3.20)$$

onde k_2 é um parâmetro da fdp.

A fusão entre as informações de cor $d_c^{(j,l)}$ e de movimento $d_{pos}^{(j,l)}$ do alvo é obtida por meio da seguinte proposição.

Proposição 3.1 *Seja $P_c(\mathbf{y}_t^c | \mathbf{x}_t) = e^{-\frac{d_c}{k_1}}$ a probabilidade de uma cor e $P_{pos}(\mathbf{y}_t^{pos} | \mathbf{x}_t) = e^{-\frac{d_{pos}}{k_2}}$ a probabilidade de movimentação de um alvo. Então, a fusão entre a informação de cor e a informação de movimentação do alvo pode ser obtida por meio de*

$$p^{(j,l)}(\mathbf{y}_t | \mathbf{x}_t) = P_c(\mathbf{y}_t^c | \mathbf{x}_t) \cdot P_{pos}(\mathbf{y}_t^{pos} | \mathbf{x}_t) = e^{-\frac{k_2 d_c + k_1 d_{pos}}{k_1 k_2}}, \quad (3.21)$$

onde os fatores k_1 e k_2 são os parâmetros da fdp exponencial.

Prova. Seja $P_c(\mathbf{y}_t^c | \mathbf{x}_t) = e^{-\frac{d_c}{k_1}}$ e $P_{pos}(\mathbf{y}_t^{pos} | \mathbf{x}_t) = e^{-\frac{d_{pos}}{k_2}}$ duas fdps com medidas independentes e identicamente distribuídas. Então, pode-se escrever

$$\begin{aligned} p^{(j,l)}(\mathbf{y}_t | \mathbf{x}_t) &= P_c(\mathbf{y}_t^c | \mathbf{x}_t) \cdot P_{pos}(\mathbf{y}_t^{pos} | \mathbf{x}_t), \\ &= e^{-d_c/k_1} e^{d_{pos}/k_2}, \\ &= e^{-\frac{k_2 d_c + k_1 d_{pos}}{k_1 k_2}}. \end{aligned}$$

Os parâmetros k_1 e k_2 podem ser convenientemente utilizados e empregados de forma eficiente no processo de fusão. Quando se desejar, por exemplo, dar maior representatividade à informação de cor, utiliza-se $k_1 \gg k_2$.

O Algoritmo 1 sumariza o processo de rastreamento de múltiplos alvos utilizando os filtros de partículas.

Algoritmo 1 MULTITARGETSMCM

- 1: **para** cada alvo detectado $\mathbf{x}^{(i)}, i = 1, 2, \dots$ **faça**
 - 2: com $\{x_{t-1}^{(i,j)}, w_{t-1}^{(i,j)}\}, j = 1, \dots, n_p$, partículas obtidas no passo anterior;
 - 3: **para** $k = t - 1 \dots, final - time$ **faça**
 - 4: $\mathbf{x}_{k+1}^{(i)} \leftarrow \text{SORT}(\mathbf{w}_k^{(i,j)}), j = 1, \dots, p$ {reamostra os p primeiros maiores valores}
 - 5: $\tilde{\mathbf{x}}_{k+1}^{(i)} \leftarrow \frac{1}{p} \left(\sum_{j=1}^p \mathbf{x}_{k+1}^{(i)} \right)$;
 - 6: $\mathbf{x}_{k+1}^{(i)} \leftarrow f(\tilde{\mathbf{x}}_{k+1}^{(i)})$; {Propagação}
 - 7: $\mathbf{w}_{k+1}^{(i)} \leftarrow \text{FUSE}(\mathbf{x}_{k+1}^{(i)}, \mathbf{y}_{k+1})$; {Observação e atualização dos pesos}
 - 8: **fim para**
 - 9: **fim para**
-

Algoritmo 2 FUSE

- 1: **para** cada vetor de observação $\mathbf{y}^{(l)}, l = 1, 2, \dots, n_{obs}$ **faça**
 - 2: $d_c^{(j,l)}(h_{k+1}^{(j)}, H_{k+1}^{(l)}) = |h_t^{(j)} - H_{k+1}^{(l)}|, j = 1, \dots, n_{obs}, l = 1, \dots, n_p$; {Encontra a distância de cor. Os valores n_{obs} e n_p são o número de observações e o número de partículas, respectivamente.}
 - 3: $P_c(\mathbf{y}_{k+1}^c | \mathbf{x}_{k+1}) = e^{\frac{d_c}{k_1}}$; {Encontra a verossimilhança de cor.}
 - 4: $d_{pos}^{(j,l)}(y_t^{(j)}, Y_t^{(l)}) = \|(\mathbf{y}_t^{(j)} - Y_t^{(l)})\|, j = 1, \dots, n_{obs}, l = 1, \dots, n_p$; {Encontra a distância.}
 - 5: $P_{pos}^{(j,l)}(\mathbf{y}_t^{pos} | \mathbf{x}_{k+1}) = e^{\frac{d_{pos}}{k_2}}$ {Encontra a verossimilhança da posição.}
 - 6: $p^{(j,l)}(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}) = e^{\frac{k_2 d_c + K_1 d_{pos}}{k_1 + k_2}}, \dots, n_{obs}, l = 1, \dots, n_p$; {Fusão.}
 - 7: **fim para**
-

O algoritmo 2 sumariza como o processo de associação entre a cor e o movimento é executado. O método pode ser convenientemente empregado em tempo real. O custo computacional envolvido está diretamente relacionado com o número de alvos e o número de partículas, como formalizado na proposição seguinte:

Proposição 3.2 *O custo computacional do método de rastreamento de múltiplos alvos e da ordem $O(mn)$, onde m é o número de objetos e n é o número de partículas distribuídas pelo filtro.*

Prova. O algoritmo principal analisa individualmente cada objeto rastreado. O custo computacional para m objetos é $O(m)$. Por sua vez, o processo de fusão opera sobre cada cada partícula distribuída sobre cada um dos m objetos. Assim, para um processo com n partículas, o custo computacional é $O(mn)$. ■

O problema de associação entre observações e trajetórias obtidas por câmeras diferentes ocorre quando a observação obtida por uma câmera deve ser confrontada com as observações e trajetórias obtidas por outra câmera do sistema. Como os campos de visão das câmeras não estão sobrepostos, então o sistema não pode usar os métodos convencionais de correspondência entre características. Porém, o sistema deve verificar se a informação recebida corresponde a alguma das observações medidas. Este problema é tratado em detalhes na seção seguinte.

3.3.2 Associação de Trajetórias em Campos de Visão Disjuntas

Uma necessidade da ODC é determinar se o alvo que foi observado e rastreado por uma câmera C_i será o mesmo observado e rastreado por outra câmera C_j que se posicionou dinamicamente para cobrir a trajetória estimada por meio das observações propagadas no sistema pela câmera C_i .

A ODC caracteriza-se como sendo um sistema de câmeras distribuídas cujo objetivo é posicionar as câmeras a partir da observação inicial do movimento de alvos no ambiente de monitoração. Uma suposição comum nos sistemas distribuídos de visão é que os campos de visão de sobreponham. A sobreposição dos campos de visão é importante para o sistema registrar os alvos em relação ao conjunto de seqüências de imagens capturadas por diferentes câmeras durante o processo de rastreamento.

Porém, a sobreposição sofre severas restrições, especialmente quando o a região de monitoramento é muito grande ou as trajetórias sofrem algum tipo de oclusão. Objetivando

aferir o rastreamento e a cobertura correta dos alvos, desenvolveu-se uma abordagem para se manter o rastreamento dos alvos em uma rede de câmeras com campos de visão disjuntos, sem sobreposição.

3.3.2.1 Associação Entre Observações e Trajetórias

Para solucionar o problema de associação, assume-se, sem perda de generalidade, que as câmeras $C_i(x_i, y_i, z_i)$ estão fixas e posicionadas em locais conhecidos em relação a um sistema inercial de referência $\{L\}$. Assume-se também o aspecto cooperativo do sistema, ou seja, que as câmeras podem se intercomunicar e trocar informações por meio de um sistema de comunicação *multicast*. Assume-se ainda que o sistema de comunicação é perfeito e que todos os robôs/câmeras estão na área de cobertura do sistema de comunicação.

Como pode ser percebido, o problema se relaciona em encontrar quais medidas observadas por uma câmera podem ser associadas às quais características de múltiplos alvos observados por outra câmera do sistema, sem que estas não compartilhem informação visual provida pela sobreposição de campos de visão. A Figura 3.3 mostra o diagrama funcional do método desenvolvido.

O procedimento de associação de observações e trajetórias é composto de duas etapas principais:

1. mapeamento,
2. separação e ajuste de trajetórias.

O sistema recebe um conjunto de dados brutos, obtidos a partir da posição dos alvos no espaço da imagem. Assim, torna-se evidente que estes dados sejam convenientemente ajustados e separados em suas possíveis trilhas e mapeados em relação a um sistema de referência comum.

O mapeamento consiste na determinação da localização dos alvos observados em relação a sistema de referência global $\{L\}$. A etapa de separação das trajetórias trata as informações obtidas pelo filtro de partículas segmentando a informação de movimento e de cor dos alvos. Por fim, a etapa de associação entre observações e trajetórias certifica se alguma observação informada ao sistema corresponde a alguma trajetória segmentada. Cada etapa está discutida em detalhes nas seções seguintes.

3.3.2.2 Mapeamento

O processo de mapeamento envolve a determinação das coordenadas do alvo em relação a um sistema de referência $\{L\}$, comum a todos os membros do sistema. O mapeamento

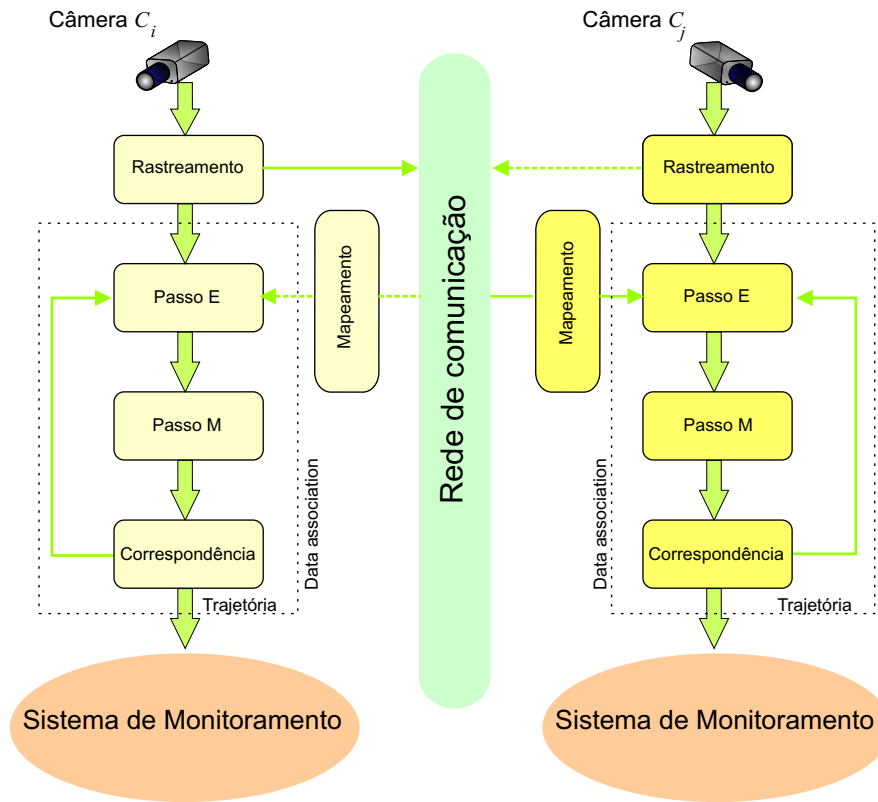


Figura 3.3: Diagrama funcional do subsistema de associação de trajetórias. O sistema recebe de uma câmera C_i informações sobre o movimento e da trajetórias de alvos que devem ser associados a movimentos e trajetórias de alvos vistos por outra câmera C_j do sistema.

usa o conhecimento dos parâmetros de calibração das câmeras e da postura do robô para mapear as observações em relação a $\{L\}$.

A geometria provida por meio do processo de formação da imagem é expressa matematicamente como um processo de concatenação de várias transformações de coordenadas e de projeção que leva ao mapeamento de um ponto de características observado no espaço da imagem a uma coordenada no espaço tridimensional.

Assumindo-se que as coordenadas do vetor posição $\mathbf{X} = [x_l, y_l, z_l]^T \in \mathbb{R}^3$, estejam referenciadas em relação ao sistema global $\{L\}$ e que uma câmera C_i visualiza um ponto a partir da posição $\mathbf{P}_c = -\mathbf{R}\mathbf{T}$, onde \mathbf{R} é a matriz de rotação e \mathbf{T} é o vetor de translação da câmera, e ainda que a câmera possui seu sistema de referência local $\{C\}$, então as coordenadas do vetor de características são mapeadas por meio de

$$\mathbf{x} = \mathbf{R}^T \mathbf{X} + \mathbf{T}. \quad (3.22)$$

A Figura 3.4 mostra um ponto de características \mathbf{X} mapeado para o plano de imagem $[x_i, y_i]^\top$.

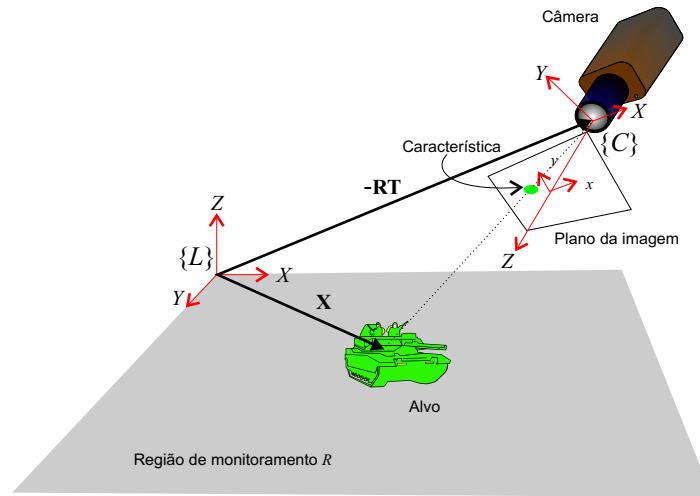


Figura 3.4: O processo de mapeamento. Um ponto com características no plano de imagem está associado a um alvo no sistema referência global.

Como a câmera é modelada por meio da transformação perspectiva, um ponto no plano da imagem está associado a um ponto do alvo por meio de

$$[x_i, y_i]^\top = [x_c, y_c]^\top / z_c. \quad (3.23)$$

Assumindo que os parâmetros extrínsecos e intrínsecos da câmera sejam conhecidos por meio de calibração, e que dispomos de informação sobre a geometria do alvo, o processo de mapeamento fornece as coordenadas de \mathbf{X} em relação ao sistema de referência $\{L\}$.

3.3.2.3 Separação e Ajuste de Dados

O procedimento de separação e ajuste de dados objetiva determinar as trajetórias que melhor se aproximam de um conjunto dados brutos referentes a posição dos alvos observados.

Foi mostrado anteriormente que quando se rastreiam múltiplos alvos o problema de associação de dados apresenta-se como um problema combinatorial. O problema é agravado quando se tenta combinar informações visuais obtidas de câmeras sem sobreposição de seus campos de visão. As observações apresentam-se como um modelo de mistura. O sistema deve, então, procurar segmentar essas informações e ajustá-las convenientemente as possíveis trajetórias.

Em outras palavras, o sistema deve encontrar para um conjunto de parâmetros e observações k_{ij} , obtidos por uma câmera C_i sobre um alvo a_j , qual é a possível trajetória do

alvo observada em outra câmera C_l do sistema. A idéia geral é agrupar itens (observações) em classes (trajetórias), por meio de propriedades representativas (características – cor e movimento do alvo).

Para resolver este problema, desenvolveu-se uma adaptação do algoritmo EM. Em nossa abordagem, o algoritmo EM executa seus dois passos convencionais objetivando segmentar os dados e identificar os parâmetros de um modelo que descrevam a trajetória dos alvos. Ou seja, o passo-E (expectativa) tenta resolver, de maneira probabilística, a associação entre características e trajetórias do conjunto de dados. O passo-M (maximização) re-estima os parâmetros por meio de uma otimização quadrática.

O processo de associação de trajetórias recebe como entrada um conjunto finito de parâmetros observados k_{ij} , enviado por uma câmera C_i , e outro conjunto de parâmetros k_{lh} , observado por outra câmera C_l , e fica alternando entre os passos E e M até que a convergência seja alcançada.

Assumindo-se que as trajetórias dos alvos descritas por meio de um conjunto de pontos $\{(x, y)\}$ são representadas por segmentos de retas descritos como

$$y = ax + b, \quad (3.24)$$

onde a é o coeficiente angular e b o coeficiente linear da reta descrita em relação ao sistema de referência $\{L\}$, então os parâmetros a e b do modelo são encontrados pela aplicação dos passos principais do algoritmo EM.

Para um conjunto de pares $l\{(a_j, b_j)\}$ de parâmetros de segmentos de reta, o algoritmo determina, no passo-E, para cada ponto de entrada (x_i, y_i) um resíduo entre cada ponto e um modelo da reta. O resíduo é uma medida que reflete a distância entre o valor observado do ponto e o valor predito pelo modelo, sendo definido para os propósitos deste trabalho como:

Definição 3.3 (Resíduo) – *O resíduo é a diferença entre o valor estimado pelo modelo e o valor observado pelo sistema. No caso em que a trajetória é modelada por uma reta, o resíduo é dado por meio de*

$$d_j^i = a_j x_i + b_j - y_i, \quad (3.25)$$

onde a_j e b_j são os parâmetros das retas e (x_i, y_i) são os pontos observados da trajetória.

Ao valor do resíduo, associa-se, a cada ponto, a variação na informação de cor, representada pela diferença entre os valores da matiz

$$h_j^i = h_i - h_{i-1}, \quad (3.26)$$

onde h_i é o valor da informação de cor (matiz) do alvo observada no quadro corrente e h_{i-1} é o valor da matiz observada no quadro anterior. Assume-se que $h_{i-1} = 0$ para $i = 1$. O valor final do resíduo fica definido pela associação entre as diferenças de movimento obtidas em cada ponto acrescidas da variação da informação de cor, sendo dado por

$$r_j^i = d_j^i + h_j^i. \quad (3.27)$$

Em seguida, encontra-se o peso de cada ponto em relação a cada modelo de trajetória. Esse peso representa a probabilidade do ponto i pertencer a reta ou trajetória j . Os resíduos são então convertidos em pesos por meio da Equação 3.28, derivada diretamente da regra de Bayes,

$$w_j^i = \frac{e^{-\left(\frac{r_j^i}{\sigma}\right)^2}}{\sum_{j=1}^n e^{-\left(\frac{r_j^i}{\sigma}\right)^2}}, \quad (3.28)$$

onde o parâmetro σ corresponde ao nível de ruído esperado nos dados e n é o número de alvos observados no campo de visão da câmera.

A entrada para o passo-E (maximização) é o conjunto de pesos. Cada segmento de reta j é aproximado ao conjunto de pontos por meio de um processo de minimização quadrática, obtido por

$$\begin{pmatrix} \sum_i w_j^i x_i^2 & \sum_i w_j^i x_i \\ \sum_i w_j^i x_i & \sum_i w_j^i \end{pmatrix} \begin{pmatrix} a_j \\ b_j \end{pmatrix} = \begin{pmatrix} \sum_i w_j^i x_i y_i \\ \sum_i w_j^i y_i \end{pmatrix}. \quad (3.29)$$

Pontos que pertencem a trajetória j possuem maior peso no processo de aproximação.

O número de trajetórias é atualizado dinamicamente com base em estatísticas calculadas. Se um ponto não se aproxima de nenhuma trajetória já encontrada, então um novo modelo de trajetória é introduzido para esse ponto.

A verificação da aproximação entre dado bruto observado e trajetória é executado pelo módulo de correspondência mostrado na Figura 3.3. Considera-se que as trajetórias estão suficientemente próximas aos pontos observados por meio da métrica

$$E_t = \frac{1}{n} \sum_i |\min(r_i^j)|. \quad (3.30)$$

Como o algoritmo é executado iterativamente, tenta-se, a cada passo, aproximar os dados às trajetórias. Assim, se a diferença $E_t - E_{t+1} \leq \tau\%$, onde τ representa um valor percentual fixo para o erro relativo, então o novo dado pertence a trajetória l ; caso contrário,

um novo modelo é iniciado para o dado corrente. O desenvolvimento da abordagem para o posicionamento das câmeras é mostrado na seção seguinte.

3.3.3 Planejamento do Posicionamento das Câmeras

O objetivo do planejamento é posicionar apropriadamente as câmeras para que estas atendam a um conjunto de critérios que viabilize a observação das trajetórias e dos alvos. Para se executar esta tarefa de maneira eficiente, serão definidas uma função de observação e uma estrutura topológica, pelo meio das quais se planejará o posicionamento.

Então, seja $C(x_c, y_c) \in \mathbb{R}^2$ as coordenadas da posição da câmera no plano de referência e θ_c a sua orientação, ambos tomados em relação a um sistema de referência ortogonal $\{L\}$. O ponto C e o ângulo θ_c especificam o ponto de visada da câmera (veja Figura 3.5). Por conveniência, a *direção de visada* é caracterizada por um vetor unitário \mathbf{v} com direção θ_c

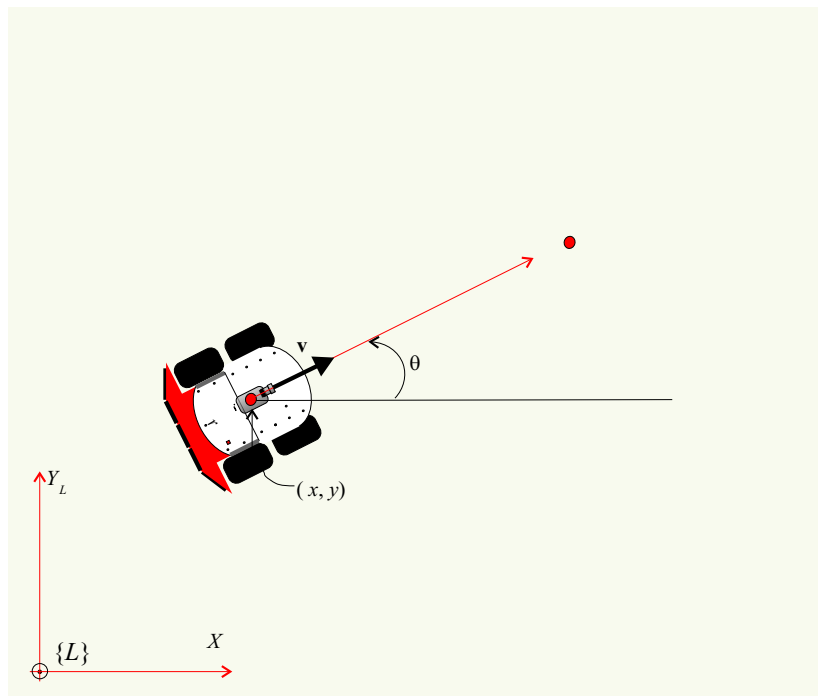


Figura 3.5: As coordenadas $C(x, y)$ e o ângulo θ definem o ponto de visada da câmera.

O ponto de visada está posicionado sobre o centro de projeção da câmera que coincide com o centro do sistema de referência do robô. Isto facilita a operacionalização do método, uma vez que as rotações e translações executadas sobre o robô serão as mesmas executadas sobre a câmera e vice-versa.

Para observar efetivamente trajetórias de alvos, um ponto de visada deve atender

um conjunto de requisitos impostos pelas características ópticas e geométricas da câmera e pelo ambiente. Nesse sentido, visando atender aos propósitos da ODC, um ponto de visada está sujeito as seguintes restrições:

Restrição sobre o ângulo de visada.

O ângulo de visada caracteriza-se como a rotação aplicada à câmera em torno do ponto de referência para torná-la ortogonal em relação à trajetória, sendo definido por meio de:

Definição 3.4 (Ângulo de visada) – *Seja \mathbf{v}_p o vetor definido entre a posição da câmera C_i e o ponto de referência P sobre a trajetória T , tal que $\|\mathbf{v}_p\|$ é a distância entre a câmera e o ponto de referência. O ângulo de visada α é o ângulo formado entre o vetor \mathbf{v}_p e o vetor normal à trajetória \mathbf{n} no ponto de referência P .*

■

A Figura 3.6 mostra o ângulo de visada para uma câmera posicionada à distância $\|\mathbf{v}_p\|$ trajetória T .

O ângulo de visada está sujeito a seguinte restrição:

$$R1 : \langle \mathbf{n}, \mathbf{v} \rangle \leq 0. \quad (3.31)$$

Essa restrição indica que o ângulo de visada não pode ser maior que 90° , condicionando o posicionamento da câmera mais ortogonal possível em relação a trajetória e evitando o efeito de encurtamento causado pela projeção perspectiva. A Figura 3.7 mostra o encurtamento devido a posição de câmera em relação ao alvo.

A restrição R1 objetiva condicionar a capacidade de observação do alvo. Para tal, a câmera deve ser posicionada na direção ortogonal à trajetória, viabilizando a visualização do alvo em sua maior dimensão e evitando o efeito do encurtamento.

Restrição sobre a abertura do ângulo de visada.

Posicionamento próximos a 90° não são possíveis e levam a obtenção de observações ruins. A idéia é manter o a abertura do ângulo de visada entre 0° e um valor máximo α_{max} . Assim, a abertura do ângulo de visada deve ser tal que:

$$R2 : \alpha = \pi - \arccos \left(\frac{\langle \mathbf{n}, \mathbf{v} \rangle}{\|\mathbf{n}\| \|\mathbf{v}\|} \right) < \alpha_{max}. \quad (3.32)$$

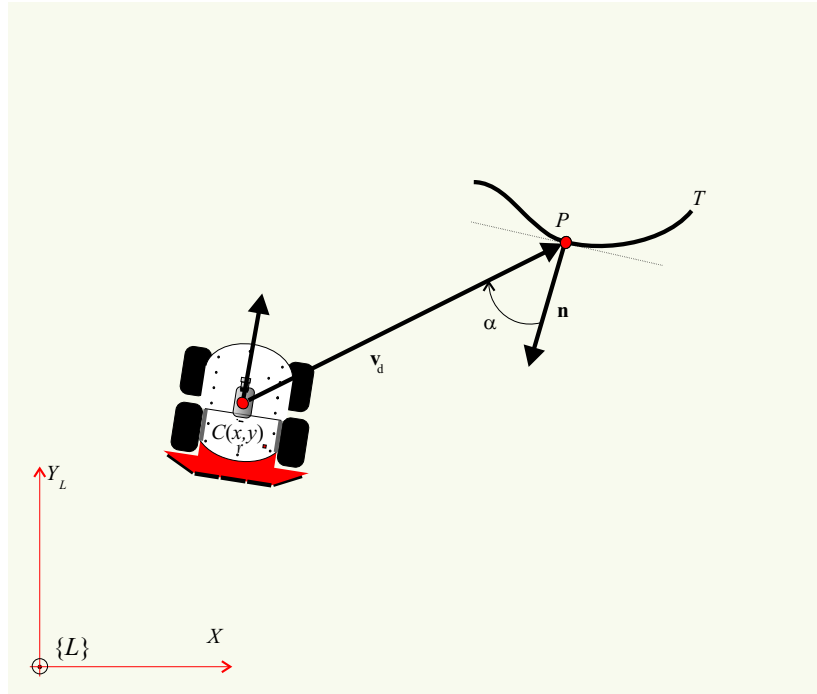


Figura 3.6: O vetor \mathbf{v}_p é formado entre a posição do robô e o ponto de referência. O ângulo de visada é formado por \mathbf{v}_p e \mathbf{n} .

Restrição à visibilidade.

Seja $e_j, j = 1 \dots, n$ o conjunto de todas as arestas que formam um espaço poligonal. Seja l_{cp} um segmento de reta que liga a posição da câmera C a um ponto de referência P_t sobre a trajetória. Então,

$$R3 : P_t \text{ é visível a partir de } C \text{ se } l_{cp} \cap \left(\bigcup_{j=1}^n e_j \right) = \emptyset. \quad (3.33)$$

A expressão 3.33 indica que para um ponto P_t sobre uma trajetória ser visível a partir da posição C , a reta l_{cp} que une o ponto C ao ponto P_t não deve se interceptar com nenhuma aresta dos polígonos que compõem o ambiente.

Restrição à rotação para o rastreamento.

Seja ϕ o ângulo de rotação da câmera sobre seu eixo vertical necessário para apontá-la em direção ao ponto de referência P sobre a trajetória. O ângulo ϕ é denominado de *ângulo de rastreamento*, sendo definido como:

Definição 3.5 (Ângulo de rastreamento) – O ângulo de rastreamento ϕ é o ângulo

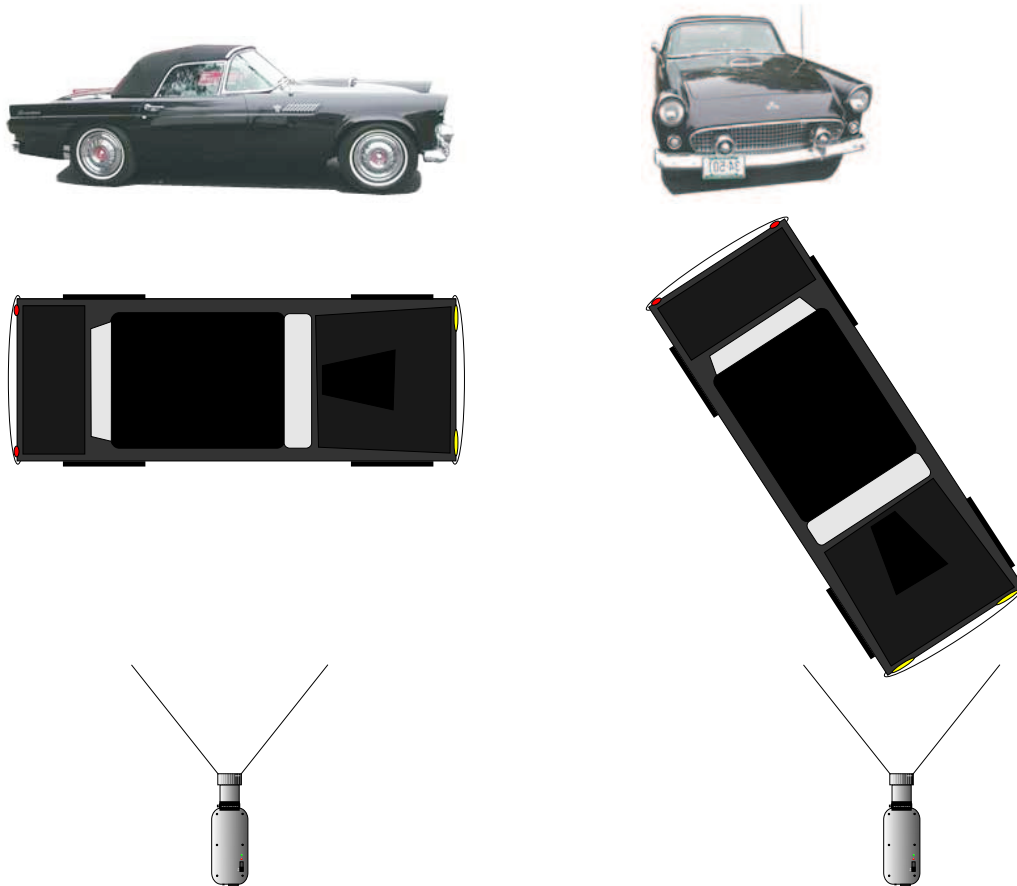


Figura 3.7: O encurtamento do alvo no espaço da imagem. Dependendo da posição da câmera, o alvo pode ser encurtado em uma de suas dimensões devido a transformação perspectiva.

formado entre o vetor de visada \mathbf{v} e o vetor \mathbf{v}_p , que une a posição da câmera ao ponto de referência P sobre a trajetória (ver Figura 3.8).

■

A restrição sobre o ângulo de rastreamento objetiva condicionar o vetor direção em relação ao ponto de referência. Isto pode ser formulado como:

$$R4 : \langle \mathbf{v}, \mathbf{v}_p \rangle = 0. \quad (3.34)$$

Restrição à cobertura.

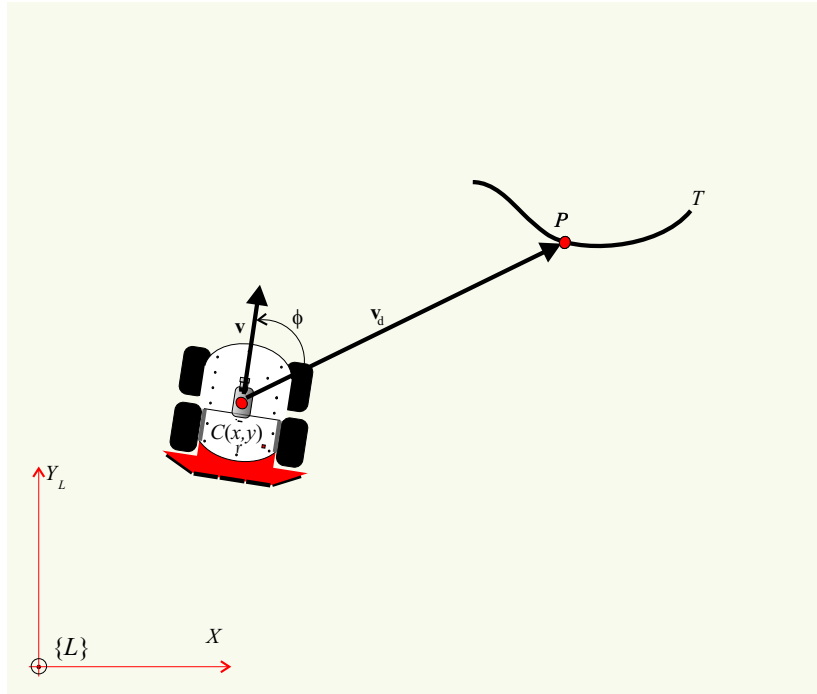


Figura 3.8: Parâmetros que influenciam a capacidade de observação: distância entre a trajetória e a câmera e os ângulos de visada α e de rastreamento ϕ .

A restrição de cobertura impõe que todo ponto P de referência esteja contido no FOV (campo de visão) da câmera. Seja \mathbf{v}_r cujo o módulo define o tamanho do raio de cobertura do FOV e γ sua abertura. Então,

$$R5 : \langle \mathbf{v} \mathbf{v}_r \rangle - \|\mathbf{v}\| \|\mathbf{v}_r\| \cos\left(\frac{\gamma}{2}\right) \geq 0. \quad (3.35)$$

Restrição à profundidade de campo.

A câmera deve estar afastada o suficiente para conter em seu campo de visão todo o espaço descrito na trajetória do alvo. Porém, as câmeras são modeladas por meio da projeção perspectiva com campo de visão fixo; o tamanho da região a ser coberta diminui, conforme a distância entre câmera e a trajetória aumenta.

A profundidade de campo da câmera está limitada entre as distâncias d_{min} e d_{max} , que podem ser estimadas por meio da geometria mostrada na Figura 3.9.

O valor de d_{min} é obtido diretamente na Figura 3.9, sendo dado por

$$d_{min} = H \tan \left[\frac{\pi}{2} - 2\gamma - \arctan \left(\frac{H-h}{d_{max}} \right) \right], \quad (3.36)$$

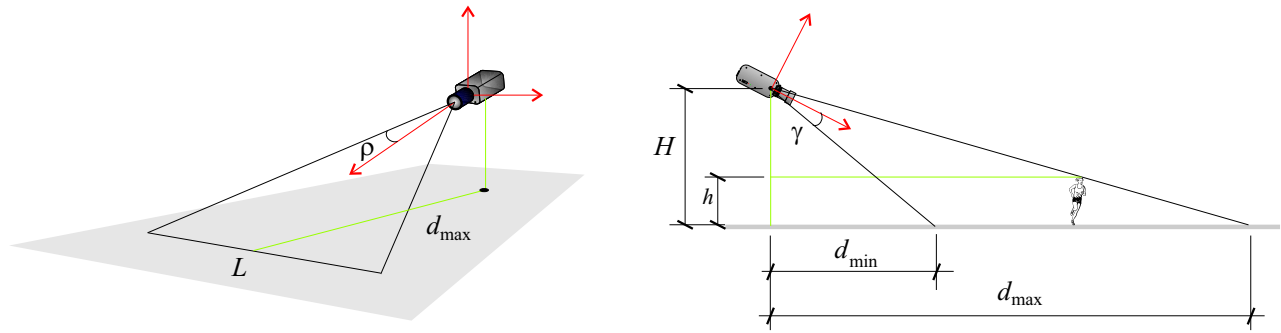


Figura 3.9: Cobertura da câmera. A altura da câmera e tamanho do ângulo de visada determinam a área de cobertura da câmera.

onde H é a altura da câmera em relação ao plano de referência, h é a altura do alvo e γ é o ângulo de abertura vertical. De maneira similar, o valor de d_{max} é obtido por meio de:

$$d_{max} = \sqrt{\left(\frac{L \times f}{ccd_u}\right)^2 - (H - h)^2}, \quad (3.37)$$

onde ccd_u é a maior dimensão do CCD, f a distância focal e ρ o ângulo de abertura vertical.

A restrição sobre a profundidade de campo fica melhor definida fazendo uso do vetor \mathbf{v}_p , sendo dada por meio de

$$R6 : d_{min} \leq \|\mathbf{v}_p\| d_{max}. \quad (3.38)$$

As técnicas convencionais de planejamento de posição utilizam uma função sobre as restrições impostas e aplica-se sobre esta alguma técnica de otimização [Chene Li (2004)]. De maneira diferenciada, a abordagem aqui desenvolvida utiliza um grafo sobre o qual serão impostas as restrições e associado a cada nó e aresta um custo referente a capacidade de observação. Essa abordagem está detalhada nas seções em seguida.

3.3.3.1 Posicionamento da Câmera em Relação a um Ponto de Referência

O posicionamento ideal deve atender a capacidade de observação da câmera, satisfazendo as restrições $R1$, $R2$, $R3$, e $R5$. A Figura 3.10 mostra uma câmera inicialmente posicionada em C_i com orientação θ_i . A linha tracejada representa a aproximação da trajetória sobre a qual está marcado o ponto de referência P_t .

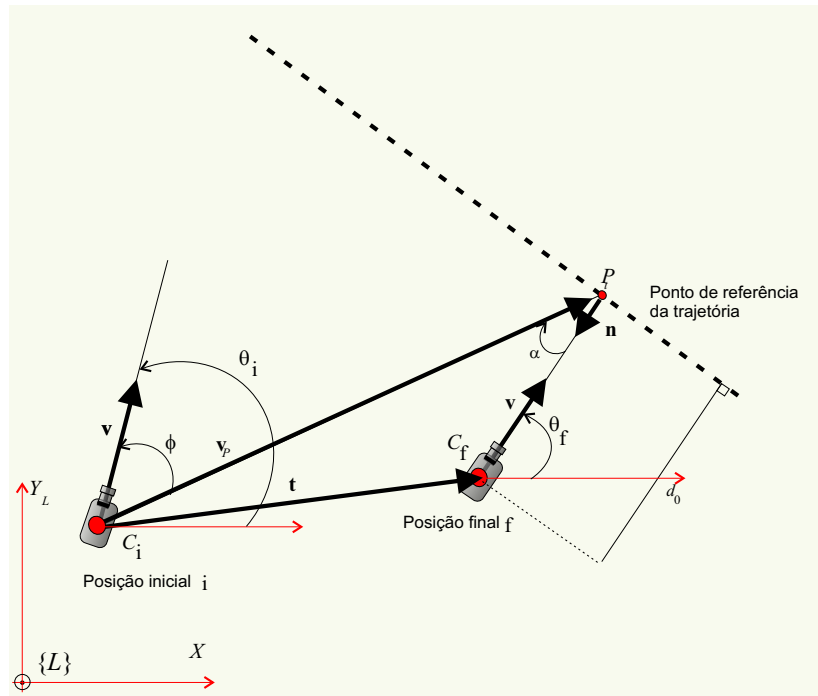


Figura 3.10: A distância entre a trajetória e a câmera e os ângulos de visada α e de rastreamento ϕ .

As restrições inerentes aos ângulos de rastreamento e de visada, assim como o de profundidade de campo, serão atendidas ao mover-se a câmera para a posição final C_f . Por meio da geometria mostrada na Figura 3.10, encontra-se o valor do ângulo de rastreamento ϕ , dado por

$$\phi = \theta_i - \theta_f = \arccos \left(\frac{\langle \mathbf{v}, \mathbf{v}_d \rangle}{\|\mathbf{v}\| \|\mathbf{v}_d\|} \right), \quad (3.39)$$

onde ϕ é o ângulo de rastreamento, θ_i é o ângulo de orientação inicial, θ_f é o ângulo de orientação final, \mathbf{v} é o vetor direção, \mathbf{v}_d é o vetor entre a posição inicial e o ponto de referência e $\langle \mathbf{v}, \mathbf{v}_d \rangle$ é o produto interno entre os vetores. A orientação final θ_f é obtida de maneira direta por,

$$\theta_f = \theta_i - \arccos \left(\frac{\langle \mathbf{v}, \mathbf{v}_d \rangle}{\|\mathbf{v}\| \|\mathbf{v}_d\|} \right). \quad (3.40)$$

De maneira semelhante, o ângulo de visada é obtido por meio de

$$\alpha = \arccos \left(\frac{\langle \mathbf{v}_d, -\mathbf{n} \rangle}{\|\mathbf{v}_d\| \|\mathbf{n}\|} \right), \quad (3.41)$$

onde α é o ângulo de visada, \mathbf{v} é o vetor direção, \mathbf{v}_d é o vetor entre a posição inicial e o ponto de referência, \mathbf{n} é o vetor normal à trajetória no ponto de referência e $\langle \mathbf{v}_d, \mathbf{n} \rangle$ é o produto interno entre os vetores.

O objetivo dessas equações é determinar as translações e rotações necessárias para o posicionamento das câmeras. Os ângulos de rastreamento e visada são determinados por meio das Equações 4.14 e 4.15, respectivamente. O vetor de translação \mathbf{t} também é obtido diretamente, sendo dado por

$$\mathbf{t} = \mathbf{v}_d - d_0 \mathbf{n}. \quad (3.42)$$

Esse conjunto de equações fornece a base para a construção de uma função capaz de avaliar a capacidade de observação de um ponto. Essa função será definida um pouco mais a frente, quando forem analisados os custos de observação sobre os nós e as arestas de uma estrutura topológica, que servirá como meio de percepção e agregação das restrições do ambiente com a função de observação. Na próxima seção será desenvolvida a estrutura topológica de apoio ao planejamento de posição.

3.3.3.2 Planejamento de Posição em Ambiente Poligonal

A principal idéia desenvolvida nesta seção é derivar, a partir do mapa geométrico do ambiente já conhecido *a priori*, uma estrutura topológica e perceptual que agregue parte das restrições impostas a visualização e torne possível o posicionamento da câmera em um ambiente com obstáculos.

Seja P um polígono simples, planar com n vértices e h polígonos simples contidos em P , representando obstáculos em P . As arestas externas de P representam os limites de um ambiente poligonal. Um conceito fundamental para o posicionamento das câmeras é o de ponto base:

Definição 3.6 (Ponto Base) – São considerados pontos de posicionamento base, ou pontos base, todos os vértices visíveis a partir de um vértice de referência v_r de P ou de seus h buracos e que estejam contidos na região formada por uma varredura rotacional Ω da câmera posicionada em v_r e limitada pelas arestas do polígono que se ligam em v_r .

■

A Figura 3.11 mostra o ponto mais escuro como o vértice de referência para a varredura e determinação dos pontos base. Todos os pontos visíveis a partir do vértice

de referência e dentro da região de cobertura obtida por uma rotação Ω da câmera, formam os pontos base para o vértice de referência v_r . Como exemplo, nota-se que o vértice 4 mostrado na figura não é um ponto base para o vértice de referência, não é visível por este último.

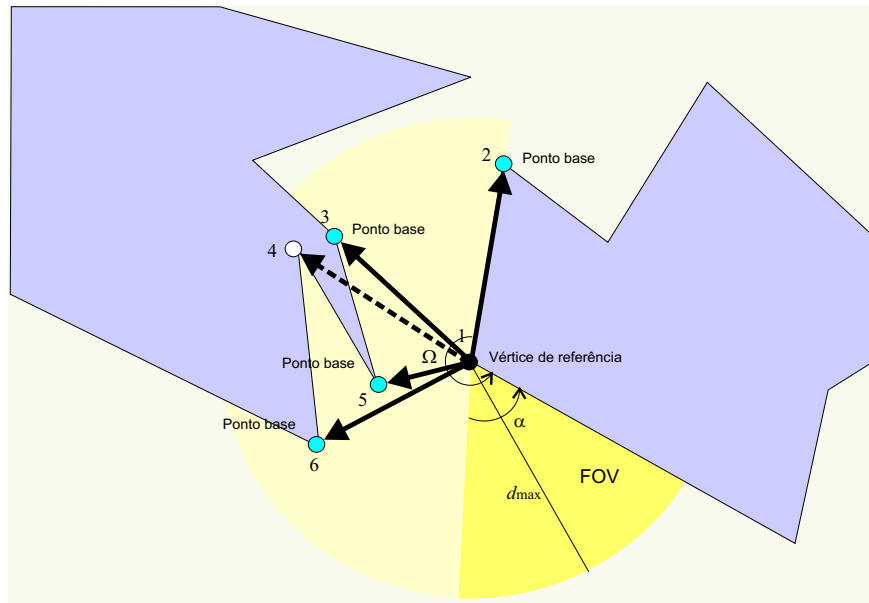


Figura 3.11: Determinação dos pontos base. São pontos base os vértices contidos na região de cobertura definida pela rotação da câmera e visíveis a partir de uma vértice de referência (ponto mais escuro).

Uma varredura rotacional Ω da câmera define uma região de cobertura, uma área circular no plano onde os pontos contidos nesta área podem estar visíveis.

Com os pontos base definidos, encontraram-se as posições a partir das quais as câmeras podem observar uma eventual trajetória contida na região de cobertura. Porém, falta ainda uma estrutura que viabilize a percepção do movimento por todo o sistema e flexibilize a movimentação da câmera (ou do robô) sobre o espaço perceptual a partir dos pontos base.

Para se construir uma estrutura topológica que agregue a viabilidade de navegação com o posicionamento na área de percepção das câmeras, serão utilizados os conceitos visibilidade e de grafo de visibilidade, apresentados no Apêndice C deste trabalho. Assim, seja $GVR(E, V)$ o grafo de vizinhança relativa, uma estrutura topológica cuja as características são dadas pela seguinte definição:

Definição 3.7 (Grafo de Vizinhança Relativa) – O GVR do conjunto X , denotado $GVR(X)$, é o grafo associado à matriz de adjacência M dada por:

$$\begin{aligned}
 M(i, j) &= 1 && \text{se} && d(p_i, p_j) \leq \max(d(p_j, p_k), d(p_i, p_k)), \\
 M(i, j) &= 0 && \text{caso contrário}
 \end{aligned}
 \tag{3.43}$$

onde $d(p_i, p_j)$ denota a distância Euclidiana entre os pontos p_i e p_j .

■

A Equação 3.43 diz que dois pontos p_i e p_j são ditos *vizinhos relativos* se não existir nenhum outro ponto p_k cuja distância em relação p_i e p_j seja inferior a $d(p_i, p_j)$. O $GVR(X)$ é o grafo $G(V, E)$ onde o conjunto de vértices V é formado por todos os pontos do conjunto X e o conjunto de arestas E é formado por todas as arestas e_{ij} cujo valor de $M(i, j)$ é igual a 1.

Como se pode observar, os n componentes conexos de $GVR(X)$ são agrupamentos que refletem uma visão de vizinhança mínima, ou seja, considerando m como sendo o número de componentes conexos obtidos de $GVR(X)$, pode-se esperar como número máximo de agrupamentos de pontos o valor $m < n$ e como número mínimo 2.

Com base na definição do $GVR(X)$ e restringindo os vínculos ao conjunto de pontos incluso no espaço perceptual das câmeras, define-se, então, o grafo de visibilidade restrita G_{vis} :

Definição 3.8 (Grafo de Visibilidade Restrita, G_{vis}) – O G_{vis} de um conjunto de pontos X é um grafo $G_{vis}(E, X)$ onde X é o conjunto de vértices selecionados como pontos base e E é o conjunto de arestas que ligam dois pontos base p_i e $p_j \in X$ de acordo com as seguintes regras:

1. $d(p_i, p_j) \leq d_{max}$, e
2. p_i é visível a partir de p_j ,

onde $d(p_i, p_j)$ é a distância Euclidiana entre os pontos p_i e p_j e d_{max} é o raio da região de cobertura formada por uma varredura Ω .

■

É importante perceber que o grafo $G_{vis}(E, X)$ agrega algumas propriedades relevantes, como:

Propriedade 1. Os componentes conexos de G_{vis} são agrupamentos que refletem uma visão de vizinhança mínima, ou seja, para um dado ponto base toda sua vizinhança está distante no máximo d_{max} deste.

Propriedade 2. Se uma rotação Ω com raio d_{max} cobrir mais da metade da área de monitoramento, então G_{vis} é o grafo de visibilidade.

Propriedade 3. O grafo de visibilidade restrita G_{vis} está contido no espaço livre de obstáculos, ou seja, suas arestas definem caminhos livres de obstáculos.

Propriedade 4 . Se G_{vis}^* for considerado como um grafo completo, aquele em que todos os vértices são conexos, então, existe pelo menos um caminho de custo mínimo em G_{vis}^* .

Essas propriedades darão suporte a utilização do G_{vis} como uma ferramenta útil na identificação de pontos base e no deslocamento da câmera em um ambiente com obstáculos. Um algoritmo para a construção do grafo de visibilidade restrita é desenvolvido na seção seguinte.

3.3.3.3 A Construção do Grafo de Visibilidade Restrita

O G_{vis} é construído para cada câmera independentemente. O grafo é calculado uma única vez, ou recalculado quando alguma mudança ocorre no ambiente, como a mudança de posição, retirada ou a inclusão de novos obstáculos, ou ainda quando são efetuadas alterações no espaço de sensoriamento ou nas características das câmeras. Assume-se que durante a execução de uma tarefa, as eventuais modificações causadas no ambiente são informadas as câmeras/robôs por meio do controle do sistema.

O cálculo do G_{vis} procede a partir das informações sobre a estrutura poligonal do ambiente. O ambiente é representado por meio de uma lista de polígonos P , constituída pelos limites do ambiente e pelos n polígonos que constituem os obstáculos. À lista P está associada a lista de arestas E com as m arestas de todos os polígonos e esta, por sua vez, está associada a lista de vértices dos polígonos V , com o total de m vértices.

Para cada polígono P_i da lista P , o algoritmo executa uma varredura de Ω_j graus em cada um dos seus $v_j \in V$ vértices. O ângulo de varredura Ω_j é o ângulo $\widehat{e_{j-1}e_j}$ formado pelas arestas que compartilham do vértice de referência v_j . A varredura define a região S_j coberta por Ω_j com raio d_{max} . O passo seguinte é a identificação dos pontos base.

Ao se executar a varredura angular, separam-se os polígonos que estejam cobertos por S_j , ou seja, encontram-se os polígonos P_k tal que $P_i \cap S_j \neq \emptyset$. A determinação dos

pontos base é executada em cada um dos P_k polígonos separados atendendo a dois critérios:

1. A visibilidade $vis(v_j, v_k)$, entre o vértice de referência v_j e os vértices do polígono em observação v_k (restrição R3).
2. O ponto escolhido está na área de varredura da câmera, ou seja, a distância de cobertura $d(v_j, v_k)$ é menor ou igual a d_{max} (restrição R5).

Assim, se o critério de visibilidade e da distância de cobertura for verificado, cria-se uma aresta unindo v_j ao ponto base encontrado em P_k . O processo se repete até que todos os polígonos tenham seus vértices verificados.

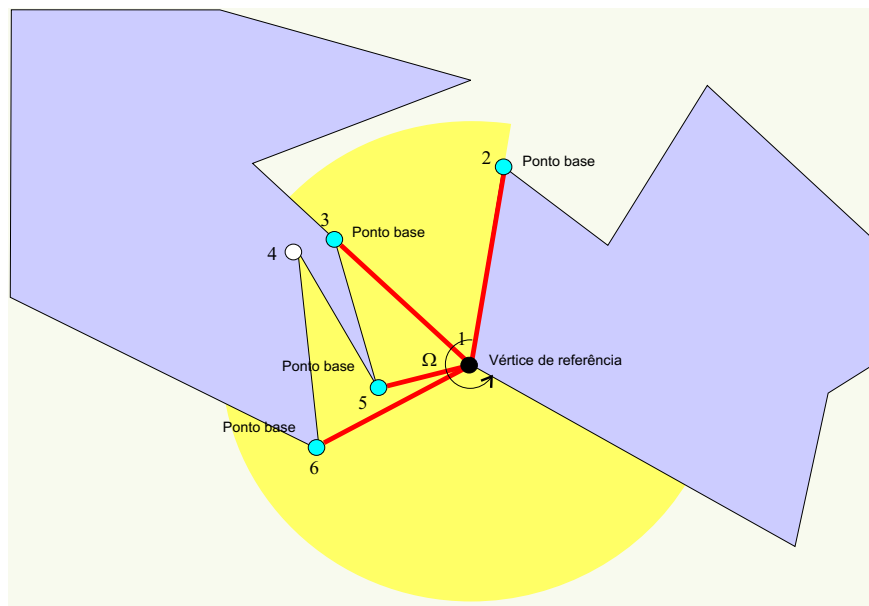


Figura 3.12: Formação das arestas do grafo de visibilidade restrita. As arestas são formadas por todo ponto base visível pelo vértice de referência.

A Figura 3.12 exemplifica a formação das arestas do grafo. Somente os pontos que atendem conjuntamente aos critérios de visibilidade e de pertinência a área de varredura formam efetivamente arestas do G_{vis} .

O Algoritmo 3 sintetiza os passos para a identificação dos pontos base e a formação do grafo G_{vis} . A formação do grafo é obtida por meio da estruturação de uma matriz de adjacência $M(j, k)$ que associa o vértice de referência v_j a um ponto base v_k .

A Figura 3.13 mostra a formação de um grafo de visibilidade restrita G_{vis} em um ambiente poligonal usando o Algoritmo 3. A formação da estrutura topológica é definida

Algoritmo 3 GVIS

```

1: para  $P_i, i = 1, \dots, n$  faça
2:   para  $v_j, j = 1, \dots, m$  faça
3:     se  $P_i \cap S_j \neq \emptyset$  então
4:        $P_k \leftarrow P_i$ 
5:       se  $vis(v_j, v_k) == TRUE$  então
6:         se  $d(v_j, v_k) \leq d_{max}$  então
7:            $M(j, k) \leftarrow 1$ 
8:            $E \leftarrow \overline{v_j v_k}$ 
9:            $V \leftarrow \{v_j, v_k\}$ 
10:        fim se
11:      fim se
12:    fim se
13:  fim para
14: fim para

```

por meio da inclusão das arestas entre os vértices visíveis a partir de um vértice tomado como base (o vértice mais escuro em cada uma das sub-figuras).

Existem ambientes em que a determinação dos pontos base é difícil ou até mesmo impossível. Ambientes sem vértices definidos em grandes extensões, como corredores muito grandes ou pistas de alta velocidade, são exemplos nos quais a determinação de pontos base não pode ser bem caracterizada. Para esses ambientes é conveniente a introdução de pontos base pelo operador do sistema, seguindo o procedimento descrito a seguir.

A inclusão de pontos base pode ser viabilizada de maneira simples e eficiente, apenas respeitando a distância d_{max} entre cada câmera e distribuído convenientemente sobre as arestas que formam as grandes extensões descobertas. Um critério eficiente é marcar os pontos de interseção da área de varredura com as arestas. O uso de uma representação em forma de grafo é muito flexível, permitindo a introdução de novos nós sem alterar a concepção original do modelo.

A Figura 3.14 exemplifica um caso de inclusão de pontos base. Os pontos mais claros na figura formam os pontos base determinados pelo critério da interseção entre as arestas dos polígonos e a área de varredura da câmera quando posicionada em cada novo ponto determinado.

Com base na estrutura topológica definida pelo grafo de visibilidade, será desenvolvida a seguir uma abordagem para o posicionamento das câmeras em ambiente poligonal.

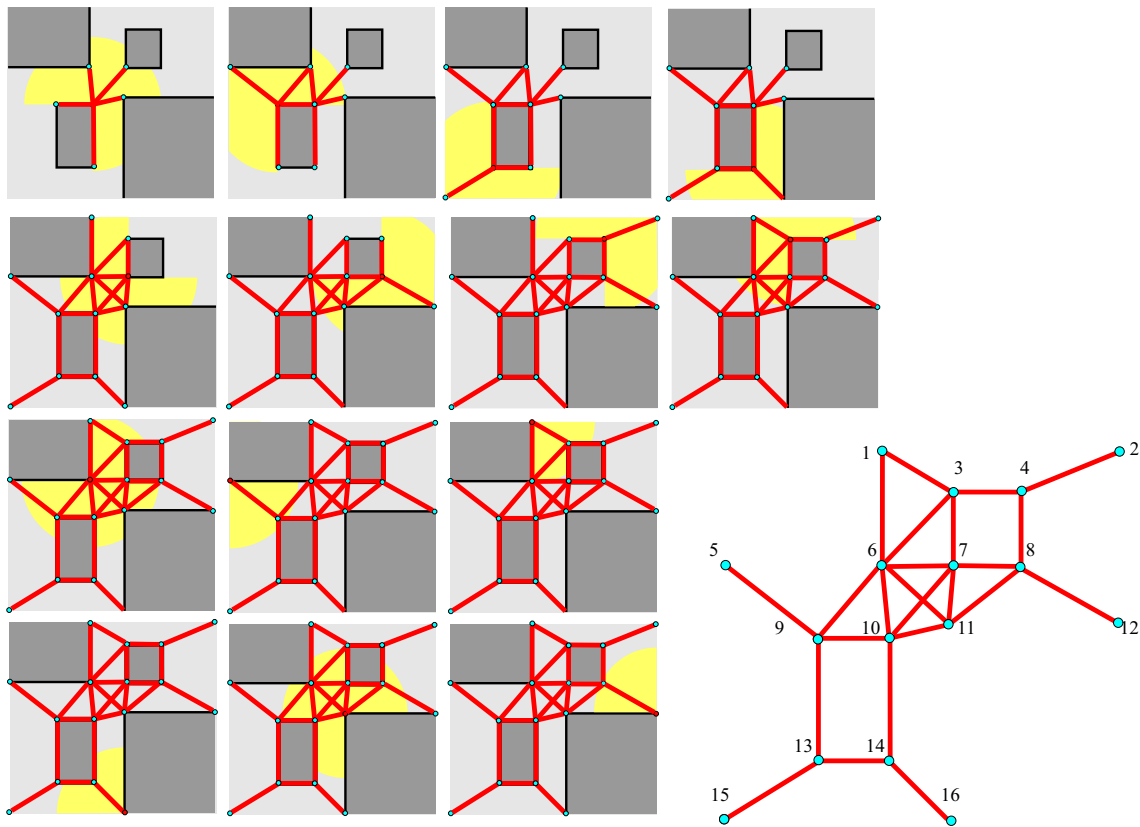


Figura 3.13: Formação de um grafo de visibilidade restrita $G_{vis}(E, V)$. A partir de um ponto de referência inicial, como mostrado no primeiro desenho no canto superior esquerdo da figura, o algoritmo determina os pontos base e inclui arestas entre estes e os pontos determinados. O processo se repete para todos os vértices de todos os polígonos.

3.3.3.4 Aplicando o G_{vis} para Posicionar as Câmeras

A finalidade do G_{vis} é a aplicação de uma ferramenta que amplie o espaço perceptivo e garanta a capacidade de observação às câmeras sobre todo o espaço de configuração do ambiente.

Em sua forma primária, o G_{vis} é um mapa topológico que associa todos os pontos visíveis pelas câmeras no espaço de configuração livre de obstáculos. Porém, para se ampliar o espaço perceptivo, é necessário que seja feita uma imersão do G_{vis} no espaço métrico, associando cada vértice a sua coordenada cartesiana e as arestas as distâncias e orientações entre os vértices em relação ao sistema de referência $\{L\}$.

Em robótica móvel, *mapas* são abstrações do ambiente essenciais para o controle e a navegação segura de um robô [Dudeke Jenkin (2000)]. Em um mapa topológico, posições específicas do ambiente são conectadas por meio de um grafo, com as arestas descrevendo

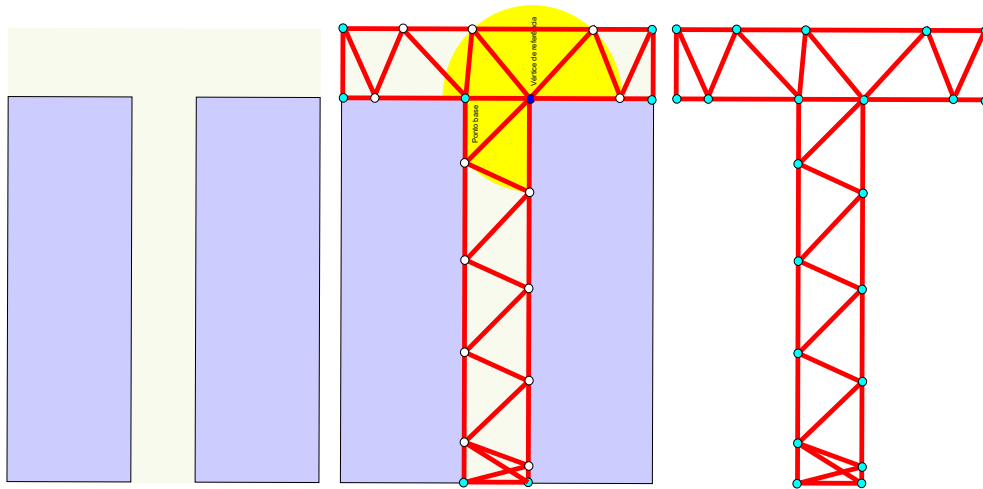


Figura 3.14: Determinação de pontos base em regiões com grande comprimento. Pontos podem ser incluídos para fornecer posições de visibilidade para as câmeras.

um caminho sobre o espaço livre. Estes mapas possuem uma descrição abstrata do ambiente e são, em geral, construídos por meio de dados sensoriais de baixo nível que determinam uma posição reconhecida no ambiente pelo robô.

Para a movimentação das câmeras assume-se que os robôs possuem a habilidade da navegação em um mapa topológico. Assume-se que o robô pode se localizar no ambiente em referência ao mapa, ou seja, o robô é capaz de determinar a sua posição mais provável no ambiente e avaliar sua movimentação no espaço a partir das informações contidas no mapa topológico.

O processo básico de navegação em um mapa topológico dá-se por meio de uma navegação exploratória a partir da identificação de lugares previstos com coordenadas pré-estabelecidas descritas no grafo. Nós não visitados pelo robô são referenciados como posições previstas, ou em aberto, no grafo. Se habilitado a mover-se fisicamente, o robô navega para a posição prevista sem encontrar nenhum obstáculo, tornando o lugar previsto em lugar confirmado, ou fechado. Caso o lugar previsto não seja alcançado ou identificado como lugar confirmado, esse nó é retirado do grafo e o mapa deve ser replanejado.

O Algoritmo BUSCAGRAFO mostrado a seguir, procura por um caminho, ou caminhos, para o robô posicionado em um vértice s do grafo e deve se deslocar para um vértice *objetivo*.

O algoritmo BUSCAGRAFO determina se existe um caminho de um vértice inicial s até um vértice ou vértices contidos no conjunto de vértices *objetivo*. O algoritmo funciona mantendo uma lista de vértices que já foram visitados (*Fechado*) e uma lista de vértices

Algoritmo 4 BUSCAGRAFO

```

1:  $Aberto \leftarrow \{s\}$ ;
2:  $Fechado \leftarrow \{\}$ ;
3:  $Achou \leftarrow falso$ ;
4: enquanto ( $Aberto \neq \emptyset$ ) .e. ( $\neg Achou$ ) faça
5:   Selecione um vértice  $n$  de  $Aberto$ ;
6:    $Aberto \leftarrow Aberto - \{n\}$ ;
7:    $Fechado \leftarrow Fechado \cup \{n\}$ ;
8:   se  $n \in objetivo$  então
9:      $Achou \leftarrow verdadeiro$ 
10:  senão
11:     $M$  é o conjunto de todos os vértices acessíveis a partir de  $n$  e que não estão em
         $Fechado$ 
12:     $Aberto \leftarrow Aberto \cup M$ ;
13:  fim se
14: fim enquanto

```

que devem ser visitados ($Aberto$) e que estejam direta ou indiretamente no caminho entre o vértice inicial e o objetivo. O algoritmo continua a busca pelo caminho até o objetivo seja encontrado ($Achou$ passa a ser verdadeiro) ou até o conjunto $Aberto$ torne-se vazio ($Achou$ é falso).

Alguns critérios ou restrições são adicionados a busca pela confirmação de posições de observações. A restrição mais comum é a de minimização de caminhos, ou seja, a posição confirmada deve atender as restrições de observação sobre os caminhos do grafo.

3.3.3.5 Custo de Observação sobre Vértices e Arestas

São definidos dois custos de observação no grafo de visibilidade restrita G_{vis} , ambos encontrados por meio da aplicação da função de observação sobre os nós e as arestas de G_{vis} . O custo de observação do nó i , $C_1(i)$ é decomposto em duas componentes: a primeira componente que relaciona o custo da distância que a câmera tem que se mover para colocar o alvo em sua cobertura, dado por

$$C_{cob}(i) = \|\mathbf{v}_{ip}\| - d_0, \quad (3.44)$$

onde $\|\mathbf{v}_p\|$ é a distância entre a posição corrente da câmera no nó i e o ponto de referência P sobre a trajetória e d_0 é a distância definida dentro da área de cobertura do sensor, atendendo a restrição R6.

A segunda componente está relacionada com os ângulos de rotação de rastreamento e

visada. O custo deve crescer de acordo com o aumento da amplitude dos ângulos de rotação. Isso sugere o uso dos cossenos dos ângulos de rotação, pois são obtidos de maneira direta por meio dos vetores de direção da câmera \mathbf{v} , do vetor \mathbf{v} e do vetor normal à trajetória \mathbf{n} , como já foi mostrado em seções anteriores. Esse custo é dado por

$$C_{ang}(i) = \frac{1}{\xi(\cos \phi_i \cos \alpha_i)}, \text{ para } 0 \leq \phi_i < \pi/2, \quad (3.45)$$

$$C_{ang}(i) = -\frac{1}{\xi(\cos \phi_i \cos \alpha_i)}, \text{ para } \pi/2 < \phi_i \leq \pi, \quad (3.46)$$

$$C_{ang}(i) = \frac{1}{\xi(\cos^2 \phi_i \cos(\pi - \phi_i) \cos \alpha_i)}, \text{ para } \pi \leq \phi_i < 3\pi/2, \quad (3.47)$$

$$C_{ang}(i) = \frac{1}{\xi(\cos^2(3\pi/2 - \phi_i) \cos(\pi - \phi_i) \cos \alpha_i)}, \text{ para } 3\pi/2 < \phi_i \leq 2\pi, \quad (3.48)$$

onde $\cos \phi_i$ e $\cos \alpha_i$ são respectivamente os cossenos dos ângulos de rastreamento e visada e ξ é um fator de escala empregado para se ampliar a importância das rotações na avaliação do custo.

Um problema freqüente no uso de funções trigonométricas é a ocorrência de pontos de singularidade. Nota-se que C_{vis} pontos de singularidade em $\cos \phi = \cos \alpha = 0$. Para contornar esse problema e aproveitando para satisfazer a restrição R2, o ângulo α fica condicionado a

$$\alpha = \pi - \arccos \left(\frac{\langle \mathbf{n}, \mathbf{v} \rangle}{\|\mathbf{n}\| \|\mathbf{v}\|} \right) < \alpha_{max}. \quad (3.49)$$

Assim, com os custos bem caracterizados, define-se a função de observação como a função que avalia os critérios de observação de um ponto de visada, escrita como

$$V_c = C_{cob} C_{ang}. \quad (3.50)$$

O cuidado ao se utilizar a função de observação é em relação à direção do ângulo ϕ , respeitando as Equações 3.45, 3.46, 3.47 e 3.48. Para $0 \leq \phi_i < \pi/2$, por exemplo, a função pode ser reescrita como

$$V_c = C_{cob} C_{ang} = (\|v_p\| - d_0) \left[\frac{1}{\xi(\cos \phi_i \cos \alpha_i)} \right]. \quad (3.51)$$

O custo de observação sobre um nó $C_1(i)$, nada mais é do que a aplicação da função de observação sobre um nó i de um grafo, sendo dada por meio de

$$C_1(i) = (\|v_p\| - d_0) \max\{C_{ang}(i)\}. \quad (3.52)$$

O custo $C_1(i)$ é obtido pela avaliação da função de observação sobre o nó de interesse i , considerando a distância $\|v_p\|$ entre o nó i e o ponto de referência P , o ângulo de visada α_i a direção e menor valor do $\cos \phi_{ij}, j = 1, \dots, n$, onde ϕ_{ij} são os ângulos de rastreamento formados pelas n arestas adjacentes ao nó i . Escolhe-se o menor valor do cosseno entre ângulos formados pelo nó e suas arestas incidentes $\cos \phi_{ij}, j = 1, \dots, n$, para que se tenha o maior valor de $C_{ang}(i)$ associado ao custo do nó. Figura 3.15 ilustra os parâmetros para o cálculo dos custos.

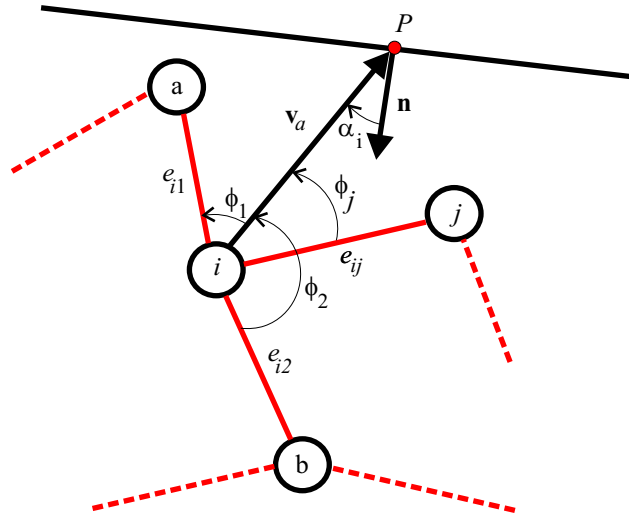


Figura 3.15: O custo da visualização está associado aos vértices do grafo e as suas arestas. O custos dos vértices leva em conta os as rotações dos ângulos de rastreamento ϕ_i e ao ângulo de visada α . Em relação as arestas, está associado o custo relativo às distâncias.

O custo de observação sobre as arestas C_2 , está associado ao custo C_1 e a distância entre os nós da aresta.

$$C_2(e_{ij}) = d(i, j) + |C_1(j) - C_1(i)|, \quad (3.53)$$

onde e_{ij} é a aresta entre os nós i e j e $C_1(i)$, $C_1(j)$ é o custo C_1 nos nós i e j , respectivamente.

A associação da distância $d(i, j)$ ao custo C_2 introduz o fator de custo relativo ao caminho. Com a determinação dos custos relativos à observação em cada nó e aresta, pode-se aplicar um algoritmo de caminho mínimo, determinando, assim, qual o melhor caminho para a câmera seguir atendendo os critérios de observação de uma trajetória. Para atender a essas restrições emprega-se, em geral, o algoritmo de Dijkstra, como mostrado em seguida.

O Algoritmo de Dijkstra (nome devido ao autor E.W. Dijkstra) é um dos algoritmos mais empregados para o cálculo do caminho de custo mínimo entre vértices de um grafo [Dudeke Jenkin (2000)]. Escolhido um vértice como raiz da busca, o algoritmo calcula o custo mínimo deste vértice para todos os demais vértices do grafo. Ele é bastante simples e com um bom nível de performance.

Algoritmo 5 DIJKSTRA

- 1: $T \leftarrow \{s\}$;
 - 2: **enquanto** $t \notin T$ **faça**
 - 3: Encontre a aresta $e \in G - T$ que para um dado vértice x possua a menor distância até s
 - 4: $T \leftarrow T + e$
 - 5: **fim enquanto**
-

O algoritmo de Dijkstra utiliza um conjunto T que armazena todos os vértices cujo caminho de custo mínimo a partir do vértice raiz s já tenham sido calculados. O algoritmo escolhe em cada etapa a aresta e de peso mínimo dentro do conjunto $G - T$, insere e no conjunto T e recalcula o custo de todos os vértices que são alcançados a partir de s .

3.3.3.6 O Planejamento de Posição

O método é executado de forma descentralizada. Cada câmera estima sua melhor posição em relação a um ponto de referência de observação e atualiza os custos de observação do grafo de visibilidade restrita. Com os valores do grafo atualizado, as câmeras/robôs deverão navegar e se posicionar em relação à trajetória estimada pelo sistema. O posicionamento deve, em síntese, obedecer os critérios de observabilidade: manter a maior parte possível da trajetória no campo de visão da câmera, evitando o encurtamento do alvo.

A idéia principal para a determinação dos pontos de observação sobre as trajetórias, é encontrar as interseções entre a trajetória estimada pelo sistema e as arestas do grafo G_{vis} . Os pontos determinados são incluídos como nós de referência para observação no grafo G_{vis} . A esses novos nós, são atribuídos custos de observação $C_1 = 0$.

As câmeras espalhadas pelo ambiente processam então a escolha por um ponto de referência, atendendo algum critério pré-definido. Um critério usual pode ser o da menor distância entre a posição atual da câmera, sobre algum nó do grafo, e o ponto de referência de observação.

Calculam-se então todos os custos de observação relativos aos nós e as arestas do G_{vis} em relação ao ponto de referência de observação escolhido. Com os custos do grafo

atualizado, procede-se então a determinação do caminho de custo de observação mínimo, aplicando o algoritmo de Dijkstra.

O Algoritmo 6 apresentado a seguir, resume o processo de planejamento de posição da câmera. São dados como entrada o ponto de referência sobre a trajetória p_t , já incluído em G_{vis} , e o nó corrente de posicionamento da câmera s .

Algoritmo 6 PLANNING

```

1:  $C_1(p_t) = 0$ ;
2: para  $i = i, \dots, n$  faça
3:    $v_i = C_1(i)$ 
4: fim para
5: para  $i = i, \dots, m$  faça
6:   para  $j = i, \dots, m$  faça
7:      $e_{ij} = C_2(i, j)$ ;
8:   fim para
9: fim para
10:  $Cm \leftarrow \text{DIJKSTRA}(G_{vis}, s, p_t)$ ;

```

A saída do Algoritmo 6 é o caminho de custo de observação mínima, que pode então ser usado no esquema de navegação e posicionamento da câmera.

3.3.4 Aplicação com o G_{vis}

Nesta seção desenvolve-se um exemplo simples do planejamento de posição usando o grafo de visibilidade restrita. Seja o grafo mostrado na Figura 3.16, o grafo de visibilidade restrita de um ambiente poligonal, por meio do qual escolheu-se o vértice 2 como o ponto de referência de visibilidade e o vértice 5 como a posição inicial da câmera.

Escolhendo-se os parâmetros $d_0 = 0$ e $\xi = 1$, aplica-se diretamente o custo C_1 sobre os vértices do grafo. Nos pontos de singularidade (em $\phi = \pi/2$ ou $\phi = 3\pi/2$) assume-se $\|\max C_{ang}\| = \|\mathbf{v}_{ip}\|$. Assim obtém-se,

$$C_1(1) = 0,000;$$

$$C_1(2) = 0,000;$$

$$C_1(3) = 4,000;$$

$$C_1(4) = 4,000;$$

$$C_1(5) = 36,00;$$

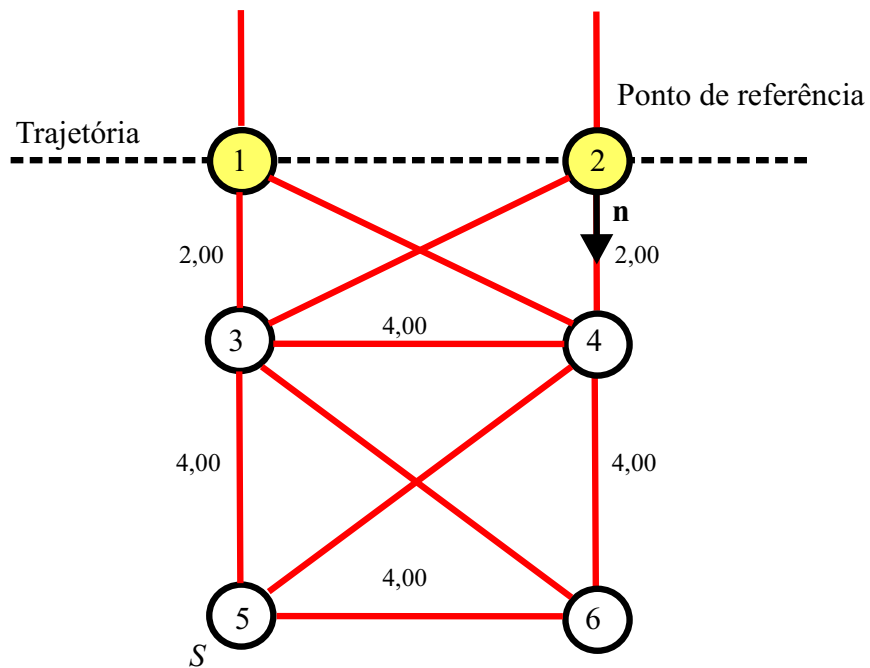


Figura 3.16: Exemplo de um grafo de visibilidade. Os vértices 1 e 2 (mais escuros) representam os pontos de observação sobre a trajetória.

$$C_1(6) = 36,00.$$

O custo sobre as arestas $C_2(e_{ij})$, é obtido por meio de

$$C_2(e_{13}) = 6,000;$$

$$C_2(e_{24}) = 6,000;$$

$$C_2(e_{23}) = 8,470;$$

$$C_2(e_{32}) = 8,470;$$

$$C_2(e_{34}) = 4,000;$$

$$C_2(e_{36}) = 39,750;$$

$$C_2(e_{35}) = 36,000;$$

$$C_2(e_{54}) = 39,750;$$

$$C_2(e_{56}) = 4,000;$$

$$C_2(e_{64}) = 36,000.$$

Com os custos das arestas calculados, encontra-se o melhor caminho de visualização, dados pelos nós $\{5, 3, 2\}$. Nota-se que a translação e as rotações são minimizadas através deste caminho.

No capítulo seguinte serão desenvolvidas as aplicações experimentais e avaliações que irão mostrar a aplicabilidade da metodologia proposta.

Capítulo 4

Aplicações Experimentais

São fúteis e cheias de erros as ciências que não nasceram da experimentação, mãe de todo o conhecimento.

Leonardo DaVinci (1452–1519), artista, arquiteto e inventor italiano.

A teoria é assassinada mais cedo ou mais tarde pela experiência.

Albert Einstein (1879–1955), físico alemão.

Projetar métodos de avaliação da performance para aplicações de visão computacional é uma tarefa desafiadora. Este capítulo trata especificamente do desenho de um protocolo para a experimentação e sua aplicação na avaliação em diversos aspectos da metodologia desenvolvida. O protocolo experimental leva em conta a acuidade do sistema em estimar a posição dos alvos e a performance na realização da tarefa de observação dinâmica.

4.1 Protocolo de Experimentação

Incertezas experimentais estão sempre associadas aos resultados obtidos pelos métodos de visão computacional. Partindo-se do princípio de que, em geral, é mais fácil planejar bem o experimentos do que tentar *a posteriori* tratar os dados e repetir milhares de vezes a experiência para minimizar o erro do resultado obtido, define-se nesta seção um *protocolo experimental*, cujo objetivo é avaliar a performance dos métodos desenvolvidos e a acurácia de seus resultados.

O principal objetivo das experimentações aqui desenvolvidas é a avaliação tecnológica e de cenário da metodologia. A *avaliação tecnológica* objetiva o entendimento do comportamento dos métodos e algoritmos desenvolvidos para a realização de tarefas específicas

. Os experimentos são realizados por meio de uma padronização ou especificação prévia dos dados de entrada. Os resultados são avaliados e a performance caracterizada por meio da comparação entre parâmetros previamente definidos e suas variações nos resultados experimentais obtidos [Thucker et al. (2002)].

Por sua vez, para a *avaliação de cenário* define-se uma aplicação particular dos algoritmos específicos, contidos em um protótipo do sistema e incluem-se outros componentes que podem ser, ou não, caracterizados [Thucker et al. (2002)]. Os dados de entrada são dados reais que podem ser controlados e reproduzidos no mundo real. Os resultados são avaliados em termos de métricas e estatísticas bem caracterizadas e definidas, que expressem graus de precisão e acurácia.

Um protocolo experimental especifica a quantidade a ser medida, a acurácia com que se quer medir e a caracterização das amostras e populações de seqüências de imagens nas quais os métodos serão aplicados. O protocolo experimental fornece, em síntese, um projeto experimental e um plano de análise de dados. Nas seções seguintes serão detalhados o projeto experimental e o plano de análise de dados.

4.1.1 Projeto de Avaliação Experimental

A idéia geral da avaliação experimental aqui desenvolvida é analisar sistematicamente o sistema. Sob este enfoque, primeiro se executa, de maneira independente, a avaliação tecnológica dos principais métodos desenvolvidos. Em seguida, busca-se a avaliação de cenário por meio de experimentos reais.

A Figura 4.1 mostra esquematicamente o protocolo de experimentação elaborada para avaliação da ODC. Assim, sob enfoque experimental, a observação dinâmica cooperativa é analisada em relação a três principais aspectos:

Ao erro obtido nas medidas observadas;

à detecção de padrões;

à performance da metodologia.

A análise das medidas consiste na verificação se o erro estatístico associado as medidas obtidas experimentalmente no processo de rastreamento são aceitáveis. Por "aceitáveis", compreende-se que o valor do erro médio obtido se encontra dentro de uma faixa definida previamente por nível de significância.

Com relação a precisão numérica dos valores processados, assume-se que dois dígitos decimais são suficientes. Uma vez que estamos trabalhando, na maioria dos casos, com

medidas em pixels. Uma precisão em centésimos de pixels garante uma representação numérica satisfatória para estas medidas.

A verificação se um determinado padrão de cor ou de movimento se associa efetivamente a um alvo ou a uma trajetória é o objetivo da análise da associação entre padrões. Essa associação pode ser medida em termos da quantidade identificada de falsos positivos e falsos negativos detectados pelo sistema em diversas seqüências de imagens. A performance da metodologia é analisada em termos de resultados obtidos por experimentos reais. As métricas utilizadas para avaliação dos métodos implementados estão descritas mais a frente.

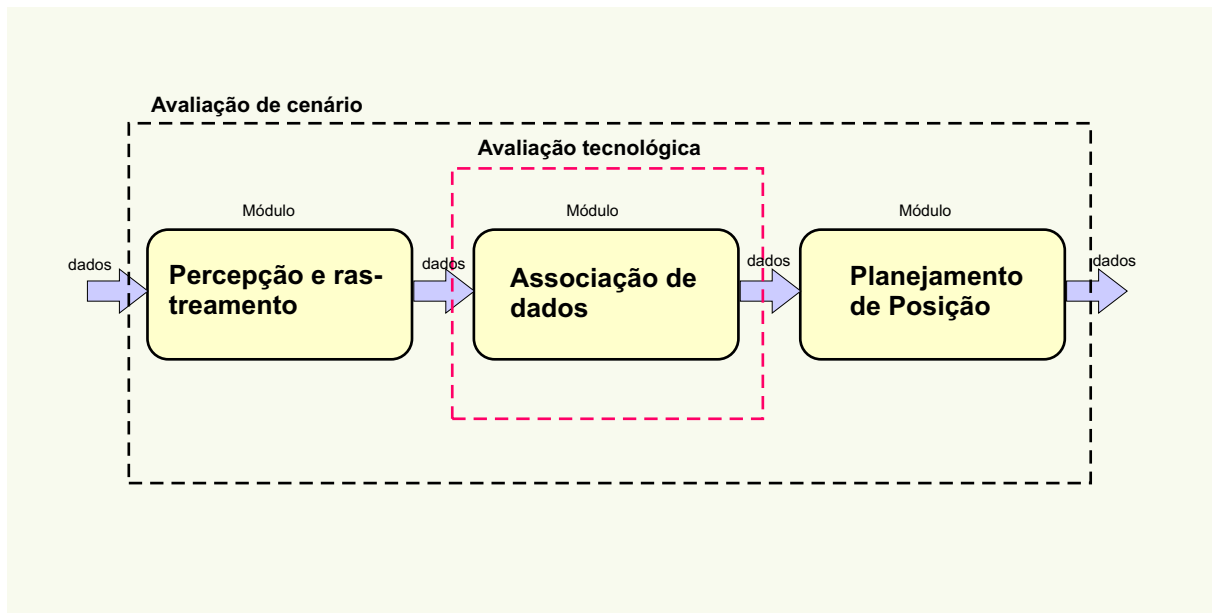


Figura 4.1: Metodologia de experimentação. A avaliação tecnológica é executada, separadamente, sobre cada um dos principais módulos do sistema. A avaliação de cenário usa um ou mais módulos em conjunto.

4.1.2 O Plano de Análise de Dados

A avaliação da acurácia e da performance da metodologia desenvolvida na tese segue os seguintes critérios de avaliação:

Acurácia do rastreamento – Este é o critério mais usual em sistemas de rastreamento.

Se o objetivo do rastreamento é estimar o estado do alvo a partir de medidas corrompidas pelo ruído, então filtra-se o ruído das medidas e observa-se quanto o estado medido se aproximou do valor do estado sem ruído [Bar-Shalom (1990)].

Uma maneira conveniente de se avaliar a acurácia do rastreador é comparar a trajetória estimada

$$T = \{(x_i, y_i)\}, i = 1 \dots, n, \quad (4.1)$$

tomadas em cada instante de tempo i , com uma trilha T_r marcada sobre uma base fixa. A trilha pode ser compreendida como uma trajetória não parametrizada pelo tempo; um caminho fixo determinado sobre uma região.

A métrica usual para medir o erro entre a trajetória e a trilha (ou entre duas trajetórias) é a distância média entre as posições tomadas a cada intervalo de tempo, dada por

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n d_i, i = 1, \dots, n, \quad (4.2)$$

onde d_i é a distância Euclidiana tomada entre as posições $(x1_i, y1_i)$ na trajetória T_1 e $(x2_i, y2_i)$ na trajetória T_2 ,

$$d_i = \sqrt{(x2_i - x1_i)^2 + (y2_i - y1_i)^2}. \quad (4.3)$$

A Figura 4.2 ilustra duas trilhas T_1 e T_2 que estão espacialmente separadas por n distâncias d_i . O distância média é avaliada pela aplicação direta da Equação 4.2.

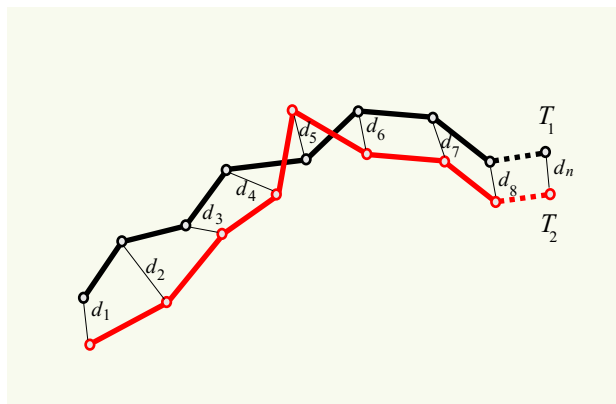


Figura 4.2: Cálculo da distância média entre as trajetórias T_1 e T_2 . As duas trajetórias estão separadas em pontos comuns tomados ao mesmo instante de tempo.

A acurácia do método de rastreamento também pode ser avaliada em relação aos erros referentes ao deslocamento horizontal Δ_x e vertical Δ_y para par de pontos tomados entre as trajetórias T_1 e T_2 , conforme pode ser observado na Figura 4.3. O modelo

ideal deve ter ambos deslocamento nulos ou muito próximos de zero. Os valores de dos deslocamentos são obtidos por meio de

$$\Delta_x i = |Xi_{T_2} - Xi_{T_1}|, \quad (4.4)$$

$$\Delta_y i = |Yi_{T_2} - Yi_{T_1}|, \quad (4.5)$$

onde (Xi_{T_1}, Yi_{T_1}) e (Xi_{T_2}, Yi_{T_2}) são as coordenadas dos pontos obtidos sobre as trajetórias T_1 e T_2 , respectivamente.

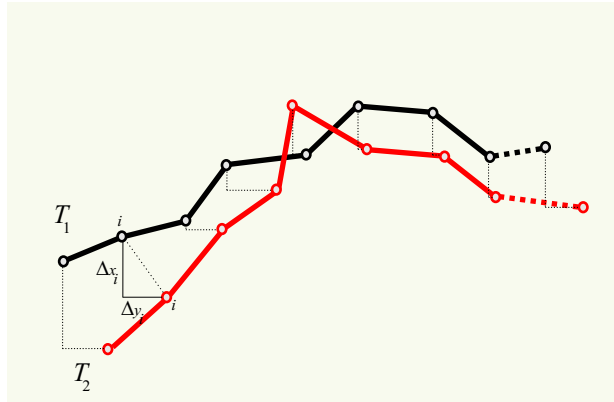


Figura 4.3: Cálculo do deslocamento horizontal e vertical entre as trajetórias T_1 e T_2 . Os pontos em comum são obtidos no mesmo instante de tempo.

Detecção de características – O sistema deve associar eficientemente características para a escolha e validação das trajetórias descritas pelos alvos. A escolha das características é o passo mais importante no sentido de se obterem os melhores resultados da avaliação da associação de dados.

Se o objetivo do método é identificar características visuais em diferentes quadros, então algumas métricas associadas a detecção de padrões devem ser definidas. Padrões são descritos em termos de seus vetores de características, que podem ser convenientemente avaliados por meio de uma métrica de similaridade.

Uma métrica de similaridade usual é a cosseno entre dois vetores. Quanto mais similar forem os dois vetores mais próximo do zero será o valor do cosseno entre eles. Esta relação pode ser escrita como uma função de similaridade, dada por

$$S(\mathbf{x}_i, \mathbf{x}_j) \triangleq \frac{\langle \mathbf{x}_i, \mathbf{x}_j \rangle}{\langle \mathbf{x}_i, \mathbf{x}_i \rangle + \langle \mathbf{x}_j, \mathbf{x}_j \rangle - \langle \mathbf{x}_i, \mathbf{x}_j \rangle}, \quad (4.6)$$

onde \mathbf{x}_i e \mathbf{x}_j são os vetores de características e $\langle \mathbf{x}_i, \mathbf{x}_j \rangle$, é o produto escalar, dado por

$$\langle \mathbf{x}_i, \mathbf{x}_j \rangle = \|\mathbf{x}_i\| \|\mathbf{x}_j\| \cos(\mathbf{x}_i, \mathbf{x}_j). \quad (4.7)$$

A aplicação do protocolo de experimentação será desenvolvida nos experimentos e em suas análises que estão apresentadas nas próximas seções.

4.2 Avaliação Experimental do Método de Rastreamento

A avaliação é realizada objetivando a análise da precisão das medidas estimadas e a performance do método de rastreamento. O método foi executado em um computador Pentium III com $800MHz$, com $256MB$ de memória, sob plataforma Windows XP. O dispositivo de captura é uma câmera digital Sony, padrão IEEE-1394 ($f = 8mm$ e $f = 16mm$), com taxa de captura de $15fps$ de imagens coloridas com 320×240 pixels. O método foi implementado em Visual C/C++. Para o interfaceamento da câmera foram utilizados os programas da biblioteca OpenCV.

4.2.1 Rastreamento e predição do movimento dos alvos

O primeiro experimento realizado teve como objetivo avaliar a precisão do método de rastreamento desenvolvido. Como forma de minimizar os eventuais erros sistemáticos, utilizou-se uma base fixa com dimensões previamente conhecidas. Sobre esta base, definiu-se uma trilha padrão sobre a qual o alvo descreve uma trajetória com medidas conhecidas *a priori*, tornando eficiente e mais prática a medida do erro.

4.2.1.1 Montagem do Aparato Experimental

Para a marcação da trilha T_r sobre a superfície de apoio, empregou-se um trilho circular de trem de brinquedo com raio $R = 29.00cm$ assentado sobre uma mesa. A Figura 4.4 mostra a fotografia do trilho sobre a mesa de experimentação e do trem envolvido por uma capa plástica vermelha para manter a uniformidade da cor durante o percurso sobre a trilha.

Uma câmera digital Sony (padrão IEEE-1394), calibrada, dotada de um conjunto de lentes com distância focal $f = 8mm$, foi posicionada a $1,50m$ a cima da base fixa. Para minimizar a ocorrência de falsos positivos devido, principalmente, à formação de sombras, uma lâmpada incandescente foi posicionada ortogonalmente em relação a mesa à mesma

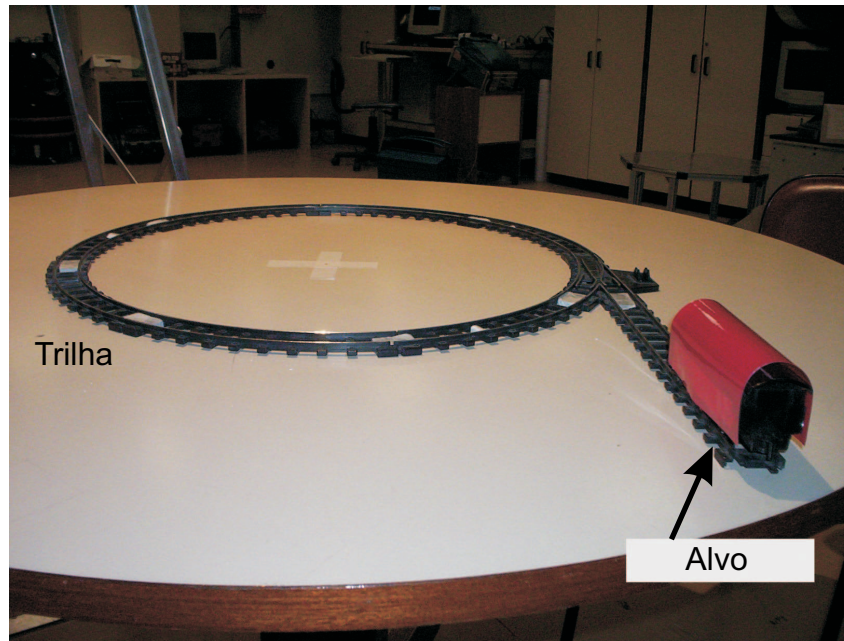


Figura 4.4: Um trilho de trem de brinquedo é montado sobre uma base fixa, servindo como a trilha de referência. O alvo é envolto em uma capa plástica de cor uniforme.

distância de $1.50m$ em relação a base. Detalhes da montagem podem ser observados nas fotografias mostradas na Figuras 4.4 e 4.5.

4.2.1.2 Determinação da Trilha de Referência no Espaço da Imagem

Como padrão para a análise do erro optou-se pela determinação empírica do valor do raio e do centro da trilha de referência no espaço da imagem. A determinação desses valores foi obtida por meio da aquisição de pontos amostrados no centro do trilho. Os pontos foram adquiridos em uma seqüência de imagens obtidas a partir do deslocamento do trem de brinquedo sobre a trilha de referência.

O procedimento para a obtenção dos pontos consistiu na segmentação da imagem do trem a cada quadro da seqüência, como mostrado na Figura 4.6, tomando-se o centro do *blob* como o ponto central da trilha. O conjunto dos n pontos $\mathcal{P} = \{(x_1, y_1), (x_2, y_2) \dots, (x_n, y_n)\}$ obtidos no espaço da imagem descrevem a trilha circular tomada como referência para a avaliação do método.

Com o conjunto de pontos \mathcal{P} , aplicou-se um procedimento implementado em *Matlab*[®] para a minimização quadrática do conjunto de pontos e a conseqüente obtenção dos parâmetros (o centro e o raio) do círculo de melhor ajuste no espaço da imagem. Os valores das coordenadas (a, b) do centro e o raio r do círculo de melhor ajuste são obtidos

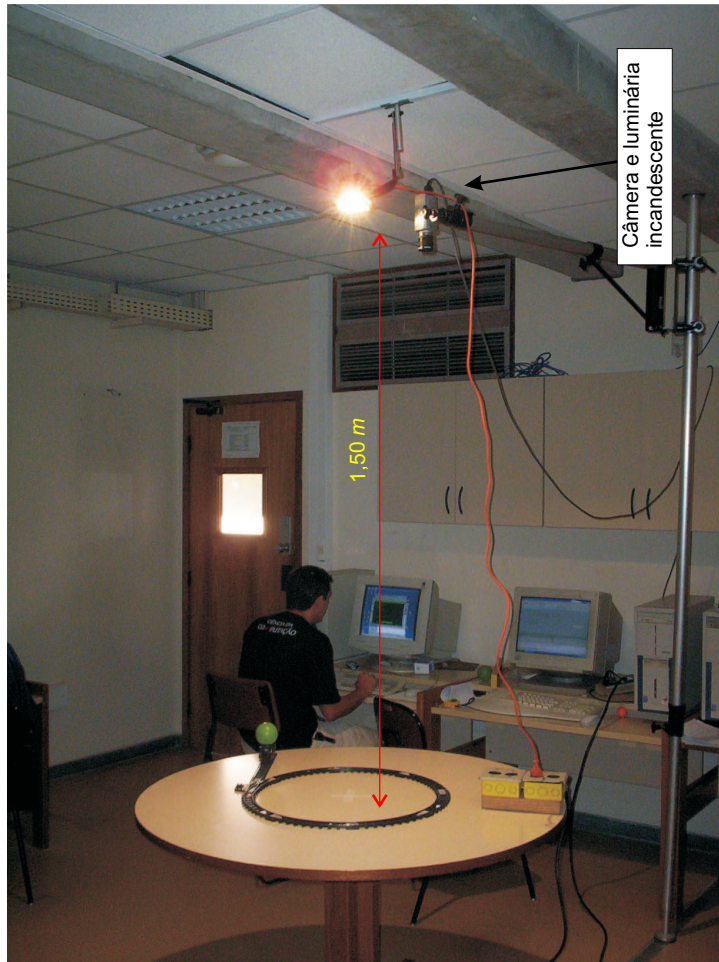


Figura 4.5: Montagem para a avaliação da precisão do método de rastreamento. A câmera está ortogonal ao plano da mesa, posicionada a uma altura de 1,50m em relação a esse plano. A mesma altura encontra-se posicionada a fonte de luz incandescente.

por meio da minimização do resíduo quadrático, dado por

$$\min\{(r_i^2 - (a - x_i)^2 + (b - y_i)^2)^2\}, \quad (4.8)$$

onde (x_i, y_i) são os pontos do conjunto \mathcal{P} , r_i é o valor do raio e a e b são os valores do centro da circunferência.

A opção pela obtenção empírica do raio e do centro do círculo no espaço da imagem deu-se devido a enorme dificuldade e da grande imprecisão que se obtém ao tentar se posicionar fisicamente o centro do círculo com o centro da imagem. Por outro lado, o valor do erro é obtido diretamente pelo emprego da Equação 4.2, em pixels, sem a necessidade da conversão entre unidades métricas.

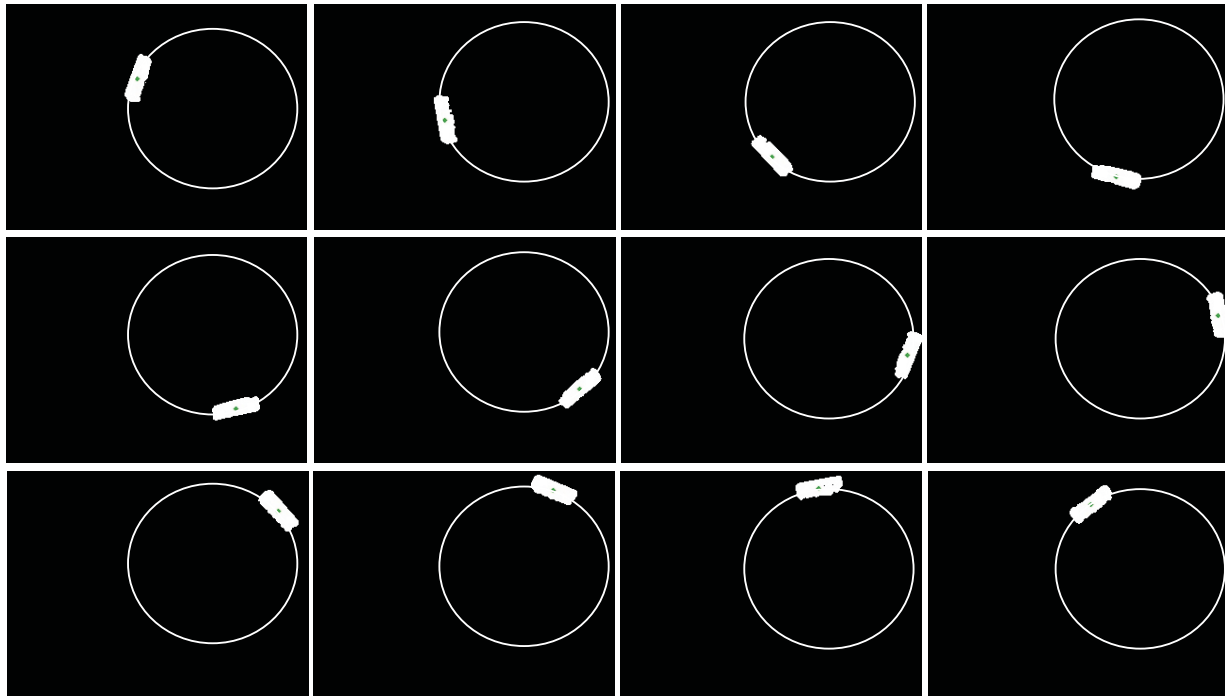


Figura 4.6: A segmentação do alvo e o posicionamento deste sobre a trilha de referência fornecem as medidas para a estimativa dos parâmetros da trajetória no espaço da imagem. As medidas são obtidas a partir do valor central do blob.

O círculo de melhor ajuste e os pontos medidos na trajetória do alvo sobre o trilho circular podem ser vistos na Figura 4.7.

Para um conjunto de 818 pontos medidos sobre a trilha, obteve-se um círculo com centro

$$C_0 = (224, 25; 144, 58)$$

e raio

$$R_0 = 85, 39\text{pixels.}$$

Estes valores serão usados como padrão para a avaliação e análise dos resultados obtidos pelas experimentações mostradas em seguida.

4.2.1.3 Análise do Erro Entre as Medidas dos Modelos de Predição e Observação em Relação a Base Fixa

Os valores de C_0 e R_0 foram utilizadas como o padrão para a identificação do erro nas predições e observações. A precisão do método é testada em função do aumento do número

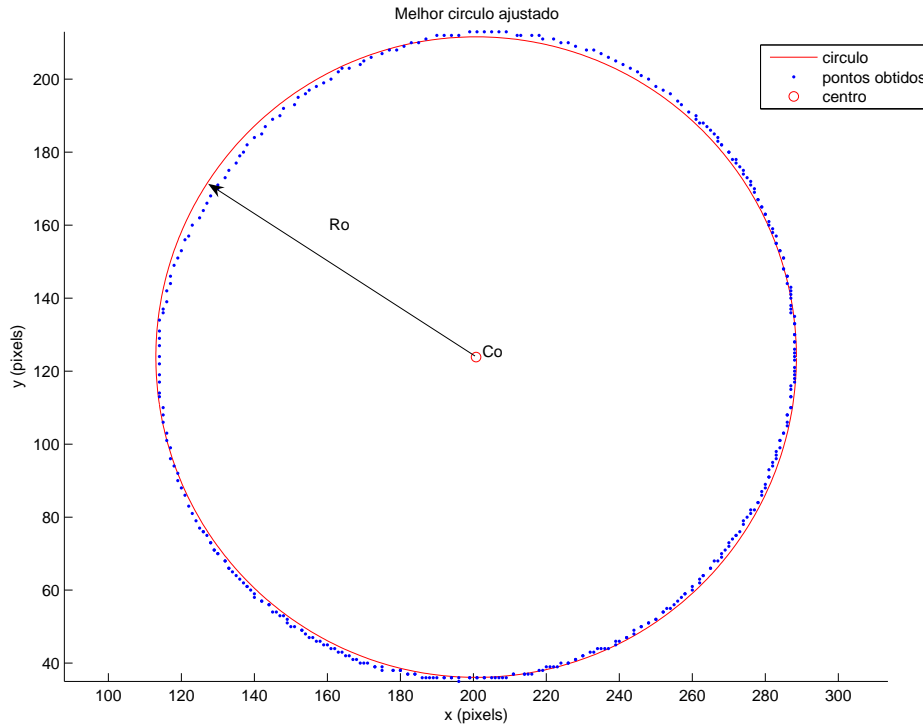


Figura 4.7: Círculo de melhor ajuste e pontos medidos no espaço da imagem.

de partículas consideradas $n_p = 50, 100, 200, 500, 1000$ e 2000 partículas.

Como métrica para o erro, utilizou-se o erro médio da diferença absoluta entre o tamanho do raio \bar{R}_{np} referente a cada um dos m pontos estimados $(\hat{x}_i, \hat{y}_i), i = 1, \dots, m$ e o valor do raio de referência $R_0 = 85,39$,

$$E_{np} = |\bar{R}_{np} - R_0|, \quad (4.9)$$

onde

$$\bar{R}_{np_i} = \sqrt{(224,25 - \hat{x})^2 + (144,58 - \hat{y})^2}. \quad (4.10)$$

Os resultados obtidos estão sumarizados no conjunto de gráficos apresentados a partir da Figura 4.8 até a Figura 4.13. A Tabela 4.1 apresenta as estatísticas encontradas para cada uma das séries referentes ao número de partículas np utilizadas.

Como pode ser observado nos resultados mostrados na Tabela 4.1, o valor do erro médio cai em relação ao aumento do número de partículas. O desvio e a mediana se mantêm com pouca variação e o erro cai consideravelmente de um valor máximo 45 pixels,

| np | Estimado | | | | | Medido | | | | |
|------|----------|----------|---------|-------|------|--------|----------|---------|------|------|
| | média | σ | mediana | Max | Min | média | σ | mediana | Max | Min |
| 50 | 3,87 | 10,57 | -11,15 | 45,87 | 1,58 | -0,09 | 1,49 | -3,01 | 2,46 | 0,14 |
| 100 | 0,39 | 1,68 | -3,10 | 3,30 | 0,39 | 0,22 | 1,57 | -3,06 | 3,11 | 0,54 |
| 200 | 0,20 | 1,82 | -4,38 | 4,05 | 0,32 | 0,20 | 0,97 | -2,88 | 2,62 | 0,15 |
| 500 | 0,15 | 1,75 | -3,74 | 3,82 | 0,55 | 0,18 | 1,72 | -3,16 | 2,74 | 0,62 |
| 1000 | 0,10 | 1,67 | -3,49 | 3,23 | 0,25 | 0,10 | 1,60 | -3,28 | 2,95 | 0,35 |
| 2000 | 0,03 | 1,62 | -2,87 | 3,03 | 0,08 | 0,07 | 1,65 | -3,16 | 2,95 | 0,11 |

Tabela 4.1: Estatísticas dos erros na estimação e na medida do raio.

com 50 partículas, para um valor mínimo, em termos absolutos, em torno 0.5 pixels com 2000 partículas.

O gráfico na Figura 4.8 mostra os valores estimados para o raio e erro absoluto obtido em uma série de rastreamentos com $np = 50$ partículas. Na parte de baixo da mesma figura observam-se os valores medidos e o erro em relação ao valor de referência R_0 .

Com 50 partículas é fácil notar a grande variabilidade do erro absoluto e a baixa precisão alcançada pelo estimador. O erro absoluto encontra valor máximo em 45 pixels. Por outro lado, a medida observada varia menos que o valor estimado, embora ainda com baixa precisão. Isso indica que, a pesar da medida ruim, o método sempre busca pela solução que torne menor a diferença entre a média e a medida observada.

Os resultados obtidos em relação ao erro absoluto para o rastreamento com $np = 100$ partículas para um único alvo sobre a trilha circular estão mostrados na Figura 4.9. As Figuras 4.9(a) e (b) mostram os resultados obtidos para o valor estimado pelo método para o raio e o erro absoluto, respectivamente. As Figuras 4.9(c) e (d) mostram o valor medido e o erro absoluto sobre o valor medido em cada instante de tempo.

Com 100 partículas, tanto os valores estimados para o raio quanto os valores observados (medidos) apresentam uma precisão mais razoável, com o erro médio quadrático caindo de 10,60 pixels com $np = 50$ partículas, para 1,70 pixels com $np = 100$ partículas. Todos as estatísticas estão sumarizados na Tabela 4.1.

De maneira semelhante as figuras anteriores, a Figura 4.10 mostra os valores estimados para o raio e erro absoluto obtido em uma série de rastreamento com $np = 200$ partículas. Percebe-se a diminuição do valor do erro médio para quase a metade do valor anterior, caindo de 0,40 pixels com $np = 100$ para 0,20 pixels, com $np = 200$ partículas. É importante perceber que embora haja uma redução significativa no erro médio, o desvio padrão mantém-se sem variação significativa, tanto para os valores estimados quanto para os valores medidos.

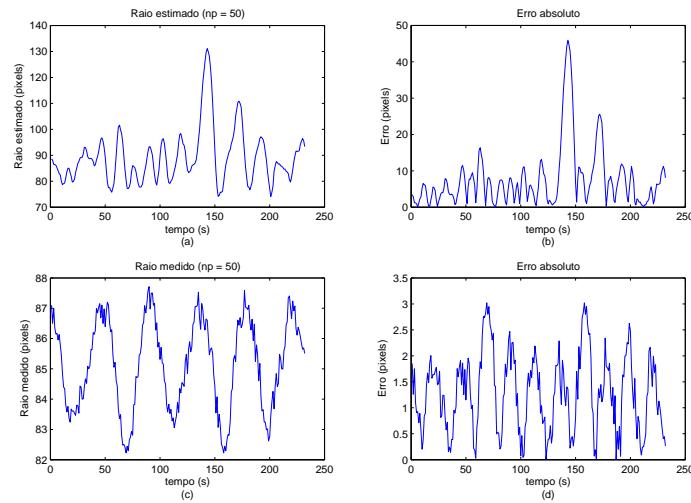


Figura 4.8: Valores estimados e medidos do raio e o erro em relação ao círculo de referência pelos modelos de predição e de medidas com 50 partículas. Os gráficos da coluna da direita da figura mostram a variação do valor do raio no modelo de predição (na parte de cima) e no modelo de medida ou observação (em baixo). Os gráficos da coluna da esquerda mostram o erro absoluto no modelo de predição (na parte de cima) e no modelo de medida (em baixo).

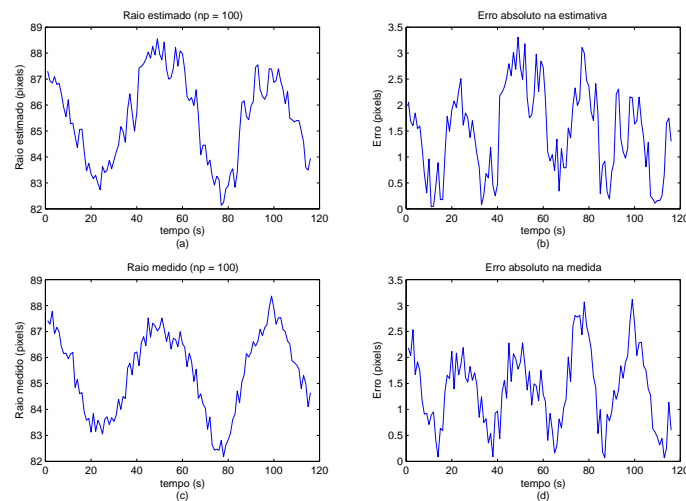


Figura 4.9: Valores estimados e medidos do raio e o erro em relação ao círculo de referência pelos modelos de predição e de medidas com 100 partículas. Os gráficos da coluna da direita da figura mostram a variação do valor do raio no modelo de predição (na parte de cima) e no modelo de medida ou observação (em baixo). Os gráficos da coluna da esquerda mostram o erro absoluto no modelo de predição (na parte de cima) e no modelo de medida (em baixo).

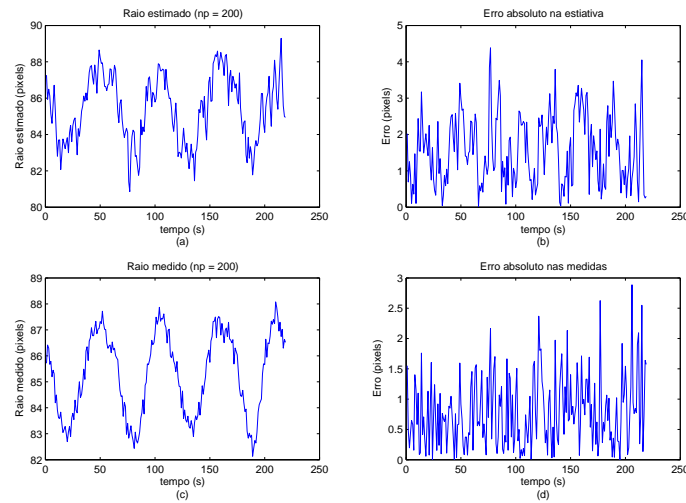


Figura 4.10: Valores estimados e medidos do raio e o erro em relação ao círculo de referência pelos modelos de predição e de medidas com 200 partículas. Os gráficos da coluna da direita da figura mostram a variação do valor do raio no modelo de predição (na parte de cima) e no modelo de medida ou observação (em baixo). Os gráficos da coluna da esquerda mostram o erro absoluto no modelo de predição (na parte de cima) e no modelo de medida (em baixo).

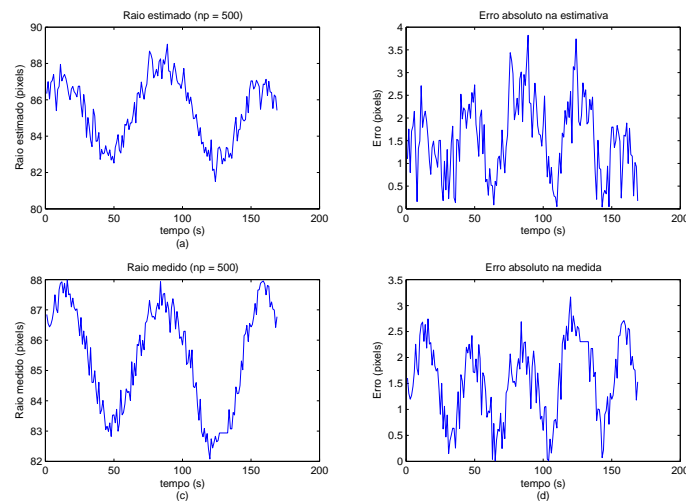


Figura 4.11: Valores estimados e medidos do raio e o erro em relação ao círculo de referência pelos modelos de predição e de medidas com 500 partículas. Os gráficos da coluna da direita da figura mostram a variação do valor do raio no modelo de predição (na parte de cima) e no modelo de medida ou observação (em baixo). Os gráficos da coluna da esquerda mostram o erro absoluto no modelo de predição (na parte de cima) e no modelo de medida (em baixo).

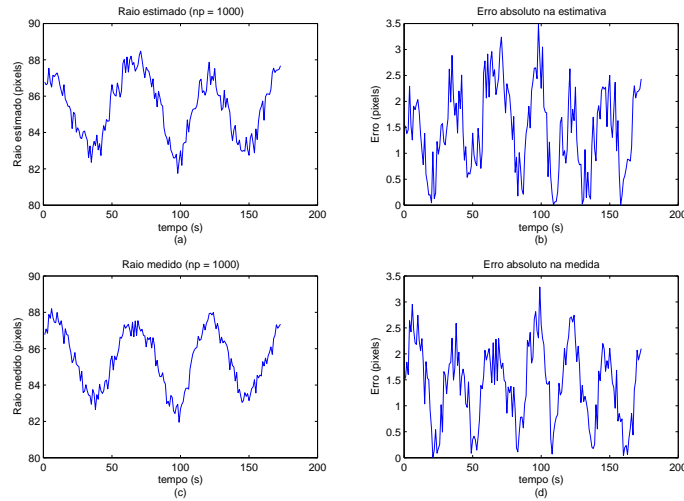


Figura 4.12: Valores estimados e medidos do raio e o erro em relação ao círculo de referência pelos modelos de predição e de medidas com 1000 partículas. Os gráficos da coluna da direita da figura mostram a variação do valor do raio no modelo de predição (na parte de cima) e no modelo de medida ou observação (em baixo). Os gráficos da coluna da esquerda mostram o erro absoluto no modelo de predição (na parte de cima) e no modelo de medida (em baixo).

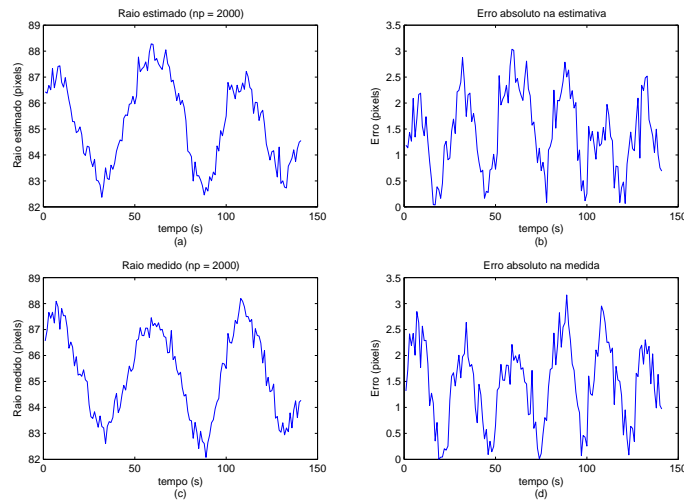


Figura 4.13: Valores estimados e medidos do raio e o erro em relação ao círculo de referência pelos modelos de predição e de medidas com 2000 partículas. Os gráficos da coluna da direita da figura mostram a variação do valor do raio no modelo de predição (na parte de cima) e no modelo de medida ou observação (em baixo). Os gráficos da coluna da esquerda mostram o erro absoluto no modelo de predição (na parte de cima) e no modelo de medida (em baixo).

Usou-se o valor de 200 partículas como o valor ideal para o emprego no método de rastreamento. Como pode ser observado, devido as altas frequências observadas na variação do erro, com 200 partículas se obtém um erro médio suficientemente pequeno (0,20 pixels) com variabilidade relativamente baixa de 1,5 pixels.

O gráfico na Figura 4.11 mostra os valores estimados para o raio e o erro absoluto obtido para o rastreamento do alvo com $np = 500$ partículas. Como nos casos anteriores, verifica-se o decaimento do erro médio, de 0,20 pixels para 0,10 pixels, e o valor do desvio padrão próximo ao valor anterior, garantindo a queda do valor erro médio estimado e a baixa dispersão do erro em torno da média obtida.

Os gráficos mostrados na Figura 4.12 apresentam os valores estimados para o raio e erro absoluto obtido no rastreamento de um alvo sobre a trilha circular com $np = 1000$ partículas. Na parte de baixo da mesma figura observam-se os valores medidos e o erro absoluto em relação ao valor do raio de referência R_0 . Nota-se uma pequena queda do valor do erro médio em relação à medida anterior. Porém, o desvio padrão não apresenta variação significativa, mostrando que a variabilidade do método é quase constante.

Na Figura 4.13, pode-se observar o menor valor obtido para o erro absoluto. Este conjunto de gráficos mostra que os erros realmente variam em função do valor de np e garantem a redução do erro com o aumento do número de partículas. Percebe-se que o desvio padrão manteve-se sem variações discrepantes, a exceção do caso em que se empregou $np = 50$ partículas. Para valores de $np \geq 100$, percebe-se ainda que os valores máximos e mínimos do erro absoluto se mantêm dentro de uma faixa entre 4 e 0.5 pixels.

Os gráficos apresentaram um comportamento periódico tanto para as medidas do raio quanto para a estimativa do erro. Esse comportamento deve-se, principalmente, a tomada das medidas sobre um círculo e a variação frequente do centro do *blob*. Observou-se, ainda, que as altas frequências observadas refletem o processo de oscilação do erro na medida provocado pela diferença entre as dimensões horizontais e verticais dos pixels e pela variação da posição do centro do *blob*, que ocorre devido a mudança de posição do alvo em relação a fonte de luz e a conseqüente ocorrência de sombras durante o percurso do alvo sobre o trilho.

O valor obtido para o erro médio E_{np} está mostrado na Figura 4.14. Como pode ser notado, o erro decai conforme o número de partículas n_p aumenta. Esse fato mostra que a abordagem obedece o decaimento teórico do erro nos filtros de partículas, que é limitado assintoticamente por $1/\sqrt{np}$, sendo np o número de partículas.

A precisão do método é estabelecida em termos do erro médio quadrático. O erro médio quadrático E_{RMS} é uma estatística que caracteriza o desvio médio total das medidas em relação a estimativa média de referência. O E_{RMS} é calculado por meio de:

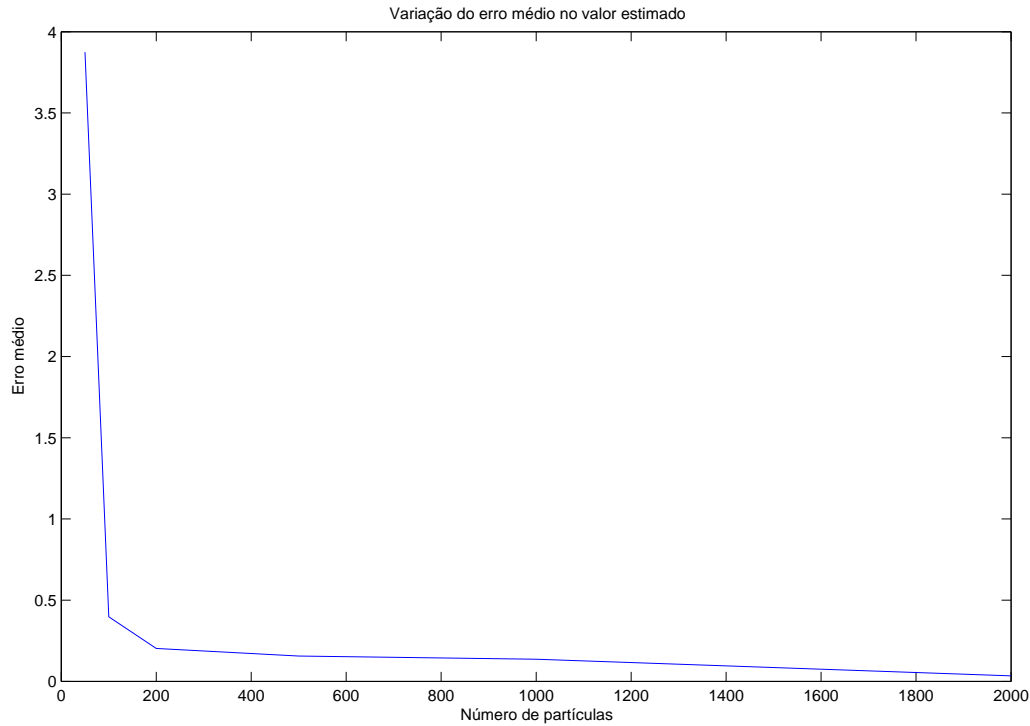


Figura 4.14: Variação do valor do erro médio na estimativa do raio. O erro decai com o aumento do número de partículas.

$$E_{RMS} = \sqrt{\frac{1}{N} \sum_{j=1}^N |\bar{R}_j - R_0|^2}, \quad (4.11)$$

onde N é o número de execuções (corridas) do experimento. Tomando $N = 6$, \bar{R}_j os valores das médias obtidas para a predição (Tabela 4.1), obteve-se

$$E_{RMS} = 1.98 \text{ pixels}.$$

O valor do $E_{RMS} = 1.98 \text{ pixels}$ caracteriza uma boa precisão estatística do método, garantindo, na média, um erro na estimativa da posição do alvo no espaço da imagem de aproximadamente dois pixels. Assim, para 95% de confiança obteve-se um limite para o intervalo de confiança de 0,80 pixels, tornando a estimativa do erro médio quadrático no intervalo $E_{RMS} = 1.98 \pm 0,80 \text{ pixels}$. Para 99% de confiança, o limite do intervalo é de 1,20 pixels, obtendo-se $E_{RMS} = 1.98 \pm 1,20 \text{ pixels}$.

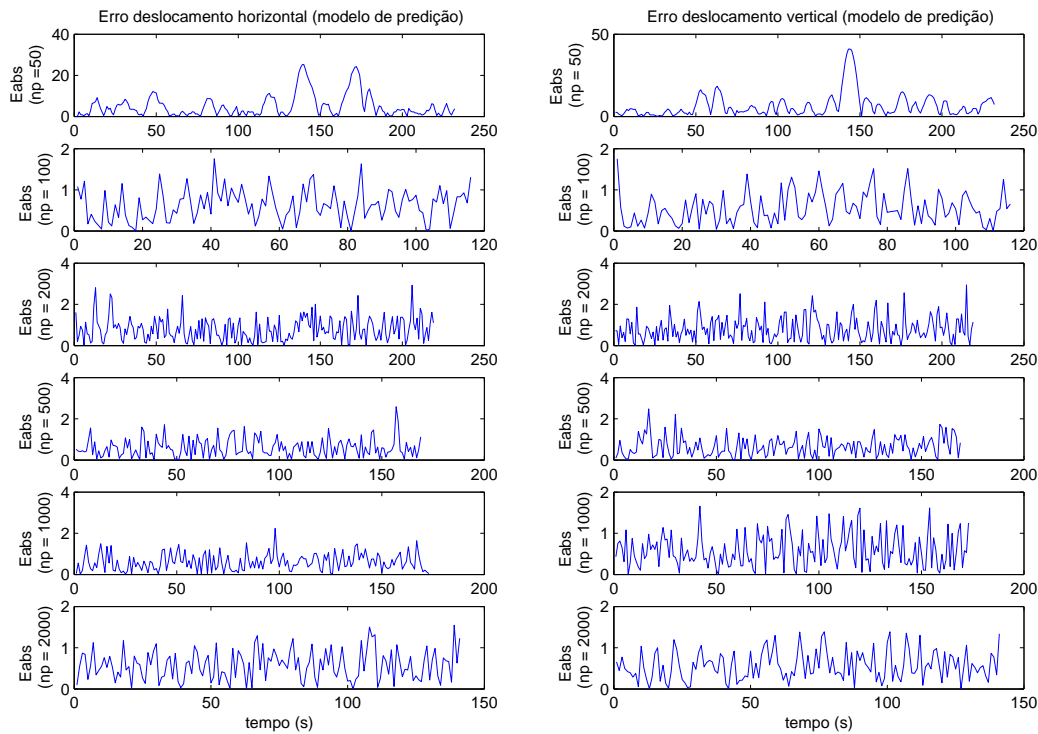


Figura 4.15: Erro na translação horizontal e vertical nas medidas de predição e observação do modelo de rastreamento para diferentes números de partículas geradas. As altas frequências tendem a se manter mais próximas da média com o aumento do número de partículas.

4.2.1.4 Análise do Erro na Translação Horizontal e Vertical

Analisou-se ainda o erro entre o deslocamento horizontal Δx e vertical Δy das medidas obtidas pelo modelo de predição e por meio do modelo de observação. Os erros referentes ao deslocamento foram obtidos por meio da Equações 4.4 e 4.5, respectivamente. A Figura 4.15 mostra os resultados para o erro entre os modelos

Essa análise mostra quanto a medida prevista se afasta, em termos absolutos, da medida observada nas direções horizontal e vertical respectivamente. As estatísticas estão sumarizadas na Tabela 4.2 mostrada a seguir.

Os valores obtidos para $np = 50$ partículas continuam mostrando a baixa precisão do método para um conjunto de poucas partículas. Isso pode ser notado por meio do valor médio de 7,50 pixels obtido para o deslocamento horizontal. Porém, a partir de $np = 100$ partículas o deslocamento em valor absoluto se mantém praticamente constante, variando

| np | Horizontal | | | | | Vertical | | | | |
|------|------------|----------|---------|-------|------|----------|----------|---------|------|------|
| | média | σ | mediana | Max | Min | média | σ | mediana | Max | Min |
| 50 | 7,50 | 8,33 | 5,78 | 45,81 | 0,03 | 1,54 | 0,80 | 1,24 | 3,27 | 0,01 |
| 100 | 1,57 | 0,90 | 1,39 | 4,61 | 0,01 | 1,37 | 0,76 | 1,38 | 3,34 | 0,01 |
| 200 | 1,54 | 0,97 | 1,49 | 3,37 | 0,01 | 1,41 | 0,75 | 1,47 | 3,33 | 0,01 |
| 500 | 1,53 | 0,89 | 1,46 | 4,01 | 0,01 | 1,29 | 0,76 | 1,53 | 3,34 | 0,00 |
| 1000 | 1,47 | 0,85 | 1,42 | 3,76 | 0,00 | 1,39 | 0,76 | 1,46 | 3,32 | 0,00 |
| 2000 | 1,43 | 0,79 | 1,31 | 3,05 | 0,00 | 1,40 | 0,78 | 1,55 | 3,44 | 0,00 |

Tabela 4.2: Estatísticas dos erros na translação horizontal e vertical.

entre 1,57 a 1,43 pixels, com o desvio padrão variando em torno de 0,90 pixels. Esse resultado já era previamente esperado, uma vez que as medidas foram tomadas em relação a uma base circular fixa, com variações de translação constantes e uniformes.

A Figura 4.16 mostra a uniformidade do erro nas medidas horizontais e verticais. Como pode ser observado, em ambos os gráficos ocorre uma redução abrupta do erro de 50 para 100 partículas. A precisão do erro com 50 partículas é muito pobre. A partir de 100 partículas, o erro tende a se estabilizar em um valor médio.

No gráfico a esquerda, as barras representam o valor do erro médio absoluto do deslocamento horizontal entre as medidas do modelo. Da mesma maneira, o gráfico da direita mostra a uniformidade entre as medidas verticais. Em ambas figuras, nota-se que a variabilidade é baixa, mesmo aumentando o número de partículas, e mantém-se sempre coesa em torno da média e da mediana.

O erro absoluto ser maior na dimensão vertical que na horizontal explica-se devido as dimensões horizontais e verticais dos pixels serem diferentes. Como o a dimensão vertical do pixel é maior que sua dimensão horizontal, o erro na vertical é relativamente maior.

4.2.2 Avaliação do Rastreamento Simultâneo de Múltiplos Alvos

Para se avaliar o rastreamento de múltiplos objetos, levou-se em conta a análise da associação entre padrão de cor e movimento e a performance deste frente a detecção e estimação correta da posição do objeto.

A seqüência de imagens mostradas na Figura 4.17 exemplifica o caso típico do rastreamento de múltiplos objetos. Nos primeiros quadros se apresenta um único objeto entrando em cena. Na seqüência, as partículas são espalhadas por uma região em torno da observação efetivada pela câmera. As partículas tendem a se agruparem próximo ao valor observado, por meio do cálculo do erro entre a posição estimada e a posição observada e o

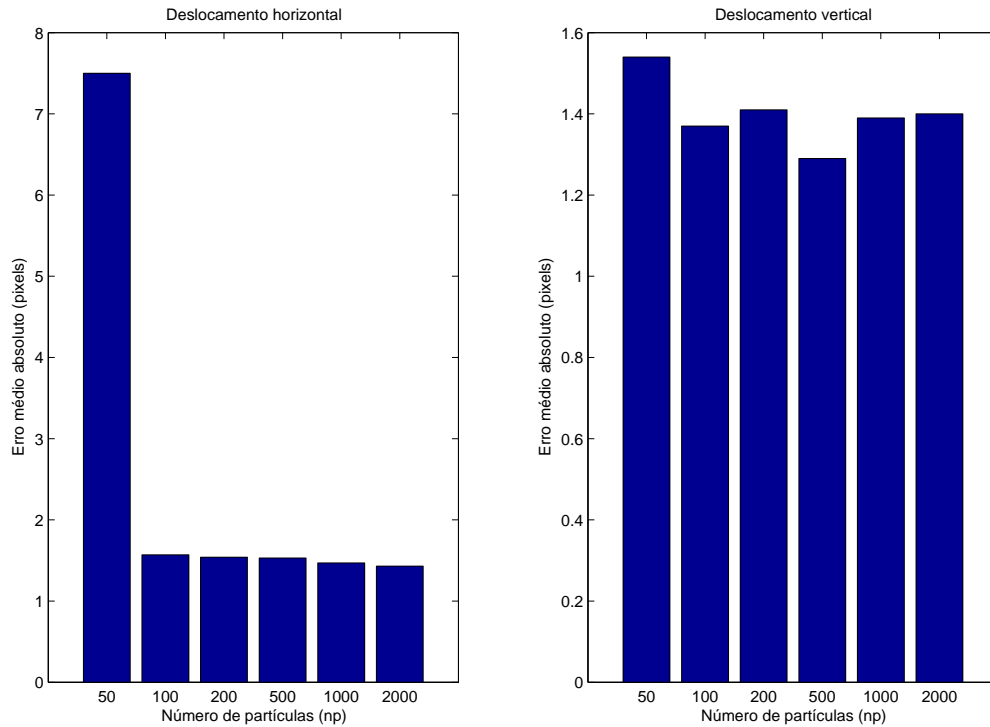


Figura 4.16: Erro médio nas medidas horizontais (a esquerda) e verticais (a direita). A maior variação do erro nas medidas verticais é devido a maior dimensão do pixel nesta direção.

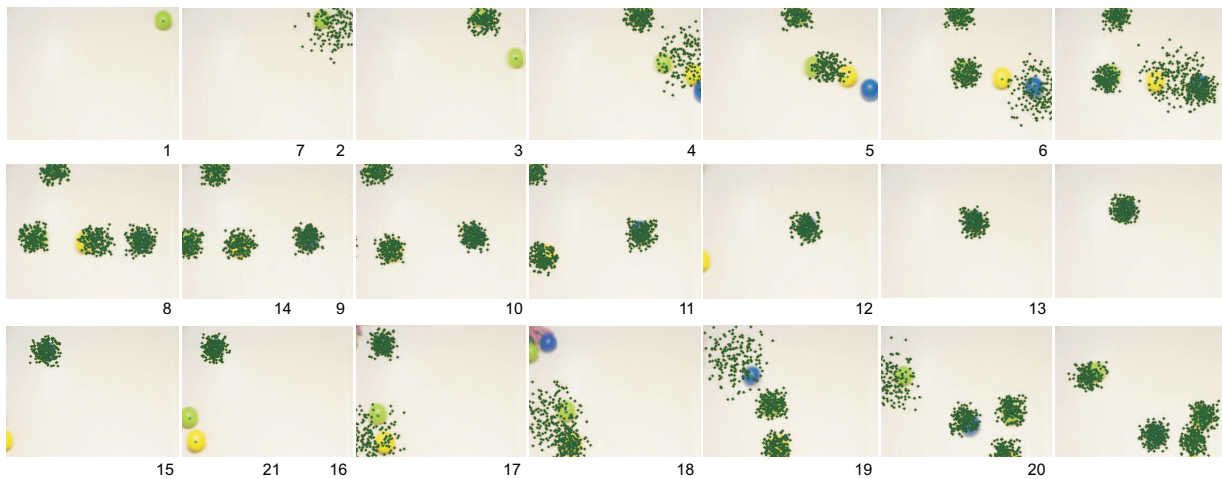


Figura 4.17: Rastreamento de múltiplos alvos. O quadro 1 mostra a entrada de um objeto em cena. Os outros quadros mostram a propagação das partículas sobre quatro objetos, rastreados simultaneamente.

| Cor | Matiz | Desvio |
|--------------|--------|--------|
| Azul escuro | 2.3521 | 0.3138 |
| Amarelo | 0.7564 | 0.0182 |
| Verde escuro | 0.9532 | 0.1000 |
| Verde claro | 0.8694 | 0.2077 |
| Vermelho | 4.3987 | 0.2193 |

Tabela 4.3: Valor Médio Observado da Cor.

valor de cor associado ao objeto. Em seguida, outros alvos entram em cena, o processo de iniciação e associação se repete para cada novo objeto. Ao final, pode-se observar quatro objetos sendo simultaneamente rastreados.

4.2.2.1 Montagem do Aparato Experimental

O experimento foi executado em um computador Pentium III com $800MHz$, com $256MB$ de memória, sob plataforma Windows $XP^{\text{®}}$. O dispositivo de captura é uma câmera digital Sony, padrão IEEE-1394 ($f = 8mm$), com taxa de captura de $15fps$ de imagens coloridas com 320×240 pixels. O método está implementado em Visual $C/C++^{\text{®}}$. Para o interfaceamento da câmera foram utilizados os programas da biblioteca $OpenCV^{\text{®}}$.

4.2.2.2 Performance do Método – A Detecção de Falsos Positivos

A cor é a principal característica visual obtida pelo método de rastreamento. Uma preocupação constante no método desenvolvido foi a acuidade no método de extração e classificação das cores dos alvos. Um falso positivo ocorre quando o filtro detecta o centro de um blob de cor diferente do objeto rastreado ou quando a estimativa está muito afastada (maior do que a variância de 10 pixels ajustada na fdp do filtro de partículas).

A Tabela 4.3 mostra o valor médio e o desvio da matiz das cores obtidas em um conjunto de 8940 observações de cor tomadas em todo o conjunto de imagens disponíveis dos alvos.

Em uma única câmera, os valores da cor mostraram-se persistentes, apresentando variação média de 0.05 radianos entre quadros, caracterizando-se como um padrão persistente e válido para a implementação da metodologia.

A ocorrência de falsos positivos pode ser notada na Figura 4.18, onde percebe-se que posições estimadas como corretas são avaliadas em locais errados, indicados pelas setas.

Considera-se que um falso positivo é detectado quando a diferença entre o valor

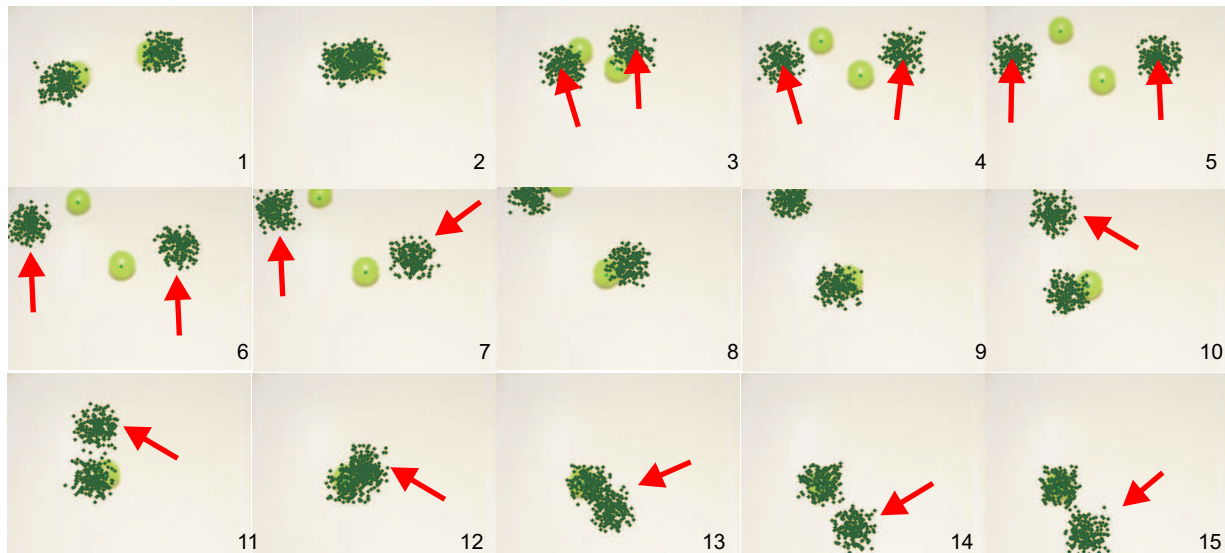


Figura 4.18: Detecção de falsos positivos. As setas indicam regiões de ocorrências de falsos positivos.

observado e o valor medido é maior que um limiar pré-definido. Para os experimentos desenvolvidos foram analisadas 12 seqüências de vídeo, mostradas na Tabela 4.4. O valor do limiar foi ajustado para 15 pixels devido a este valor ser maior ao valor do desvio padrão normalmente aqui empregado como parâmetro para a distribuição de probabilidade no método de Monte Carlo (10 pixels).

A análise das seqüências de vídeo foi *off-line*. Empregou-se o programa VirtualDub® 1.6.5 para separar os quadros das seqüências de vídeo e o programa MatLab® R.12 para a análise dos quadros e detecção dos falsos positivos.

Foram considerados como quadros válidos para o cálculo do percentual de falsos positivos, P_{fp} , somente os quadros das seqüências de vídeo que apresentavam movimento de alvos. A relação é dada por

$$P_{fp} = \frac{Nq}{Nv} \times 100, \quad (4.12)$$

onde Nq é o número de falsos positivos detectados em cada quadro e Nv é a quantidade de quadros válidos de uma seqüência de vídeo. A Tabela 4.4 mostra 12 seqüências de vídeo analisadas em relação a ocorrência de falsos positivos.

Para um total de 596 quadros com objetos em movimento analisados, foram detectadas 72 ocorrências de falsos positivos, fornecendo uma taxa de 12.08% de detecções errôneas. Este valor indica que, em geral, a detecção de falsos positivos apresenta-se com uma freqüência relativamente alta, uma vez que em toda seqüência sempre ocorrerá falsos

| Vdo. | Nv | Nq | $P_{fp}(\%)$ |
|-------|------|------|--------------|
| 1 | 52 | 6 | 11,53 |
| 2 | 25 | 3 | 12,00 |
| 3 | 103 | 6 | 5,83 |
| 4 | 65 | 4 | 6,15 |
| 5 | 83 | 8 | 9,64 |
| 6 | 44 | 4 | 9,10 |
| 7 | 56 | 4 | 7,14 |
| 8 | 48 | 4 | 8,33 |
| 9 | 92 | 18 | 19,56 |
| 10 | 34 | 4 | 11,76 |
| 11 | 55 | 6 | 10,91 |
| 12 | 32 | 5 | 15,62 |
| Total | 596 | 72 | 12,08 |

Tabela 4.4: Ocorrência de falsos positivos em seqüências de vídeo.

positivos nos primeiros quadros, no processo de iniciação dos alvos, antes da formação completa do blob. Se esses primeiros quadros não forem levados em consideração, a taxa cai consideravelmente para valores em torno de 4.8%.

Algumas seqüências apresentaram valores mais elevados do P_{fp} (Vdo. 9, na Tabela 4.4) devido a situações provocadas. A Figura 4.18 mostra um exemplo no qual alvos de cor semelhante se chocam e provocam a ocorrência de falsos positivos.

A característica de cor tendem a ser persistentes e robustas, levando a taxas muito baixas de ocorrência de falsos positivos provocadas pela variação da cor. Casos extremos ocorrem com a mudança da iluminação e a formação de sombras.

4.2.3 Discussão

O método de rastreamento se comportou como esperado em relação ao erro. Na primeira etapa de experimentação, a determinação do erro médio sob condições controladas forneceu uma estimativa segura, que confrontada com os resultados teóricos garantem a precisão dos resultados obtidos.

Um fato curioso é a variação periódica do erro em todas as análises realizadas nessa primeira fase. Uma explicação para esse fato é que as medidas previstas e observadas são tomadas sobre uma circunferência, que pode eventualmente manter a característica periódica das medidas em seus respectivos erros. Observou-se ainda que a frequência do

erro é alta em relação a média, isso é devido as variações periódicas do centro do *blob* por causa da iluminação e o conseqüente surgimento de sombras.

Na segunda fase de experimentações, quando se relaxou as condições de iluminação do ambiente e se rastreou múltiplos objetos simultaneamente, percebeu-se que o método mantém a precisão muito próxima ao valor de 1,90 pixels do erro médio quadrático encontrado no experimento controlado. Isso é uma garantia de robustez do método e de viabilidade de aplicação em experimentações menos controladas.

Alguns resultados na associação entre trajetórias foram afetados na precisão devido a distorção das lentes o que comprometia uma medida precisa. Estas falhas foram sendo analisadas e corrigidas na desenrolar dos experimentos, conforme iam sendo detectadas.

4.3 Avaliação do Método de Associação entre Medidas e Trajetórias em uma Rede de Câmeras com Campos de Visão Disjuntas

O objetivo deste experimento é avaliar a re-identificação do alvo e a associação deste a uma trajetória em câmeras com campos de visão sem sobreposição.

A abordagem foi experimentada nas dependências do Laboratório de Visão Computacional e Robótica (VeRLab). Foi montada uma rede TCP/IP formada por três câmeras digitais ligadas a computadores Pentium IV sob a plataforma Windows XP, todos com $800MHz$ e $256Mb$ de memória.

Utilizou-se uma câmera digital Sony, padrão IEEE-1394 com jogo de lentes com distância focal $f = 8mm$, com capacidade de captura de imagens coloridas com tamanho de 320×240 pixels a uma taxa de 15 quadros por segundo e duas câmeras digitais tipo webcam com interface padrão USB, lentes com ajuste de foco manual com $f = 6mm$, com capacidade de aquisição de imagens com 320×240 pixels e taxa de aquisição de 30 quadros por segundo.

Os parâmetros intrínsecos da câmera foram obtidos por meio de calibração. Os parâmetros extrínsecos, como a posição absoluta e orientação foram ajustados manualmente por meio da utilização de réguas e trenas métricas.

4.3.1 Montagem do Aparato Experimental

As câmeras foram posicionadas sobre o piso do laboratório com alturas e direções de visada referenciadas em relação a um sistema de referência fixo, marcado sobre a mesma base.

Nesta montagem, a câmera 1, está localizada em $C_1(-1.20m, 0.0m, -1.35m)$, a câmera 2 em $C_2(-1.42m, 2.40m, 1.32m)$ e a câmera 3 em $C_3(0.15m, 3.27m, 1.47m)$ em relação ao sistema de eixos mostrado na Figura 4.19.

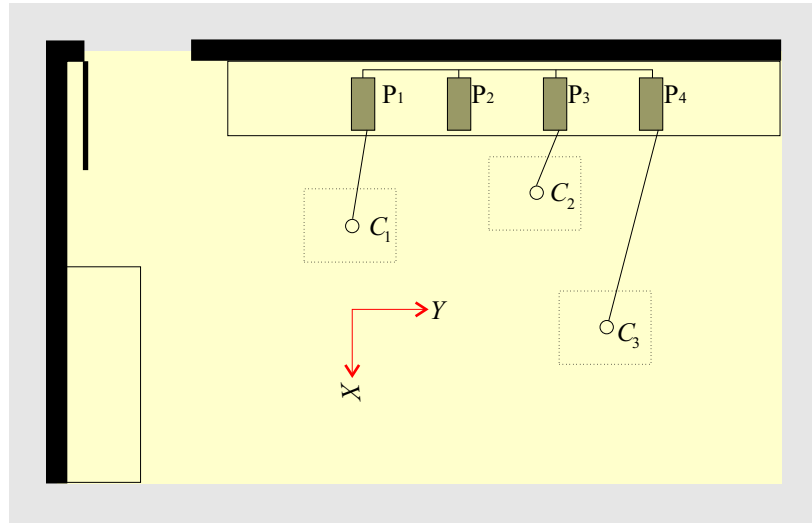


Figura 4.19: Disposição física das câmeras C_1 , C_2 e C_3 no ambiente do VeRLab. A rede é formada por quatro computadores P_1 , P_2 , P_3 e P_4 .

As câmeras C_1 e C_2 estão orientadas em 0° em relação ao eixo vertical do sistema de referência global, e a Câmera C_3 está orientada 45° em relação ao mesmo eixo.

4.3.2 Experimentação

A metodologia foi testada sobre 12 seqüências de imagens reais. Como alvos foram usadas bolas coloridas (duas bolas amarelas, duas verdes, uma azul e uma vermelha), todas com diâmetro $d = 5cm$. As condições de iluminação são as do próprio ambiente e os parâmetros para o filtro de partículas foram ajustados para $n = 200$ partículas e o desvio padrão entre as medidas obtidas pelo rastreador em $\sigma = 20$ pixels.

O cenário é descrito como uma via onde os alvos trafegam em diferentes sentidos e são monitorados por três câmeras sem sobreposição dos campos de visão. A via de $4,50m$ é descrita por duas retas paralelas distanciadas de $0.90m$, marcadas no chão do laboratório com fita crepe.

Os dados são adquiridos assincronamente e processados em tempo real. O sistema inicia o processo de rastreamento por meio da execução do filtro de partículas e envia a informação obtida do rastreamento para a rede.

Neste experimento, três alvos são rastreados e descrevem suas trajetórias sobre o

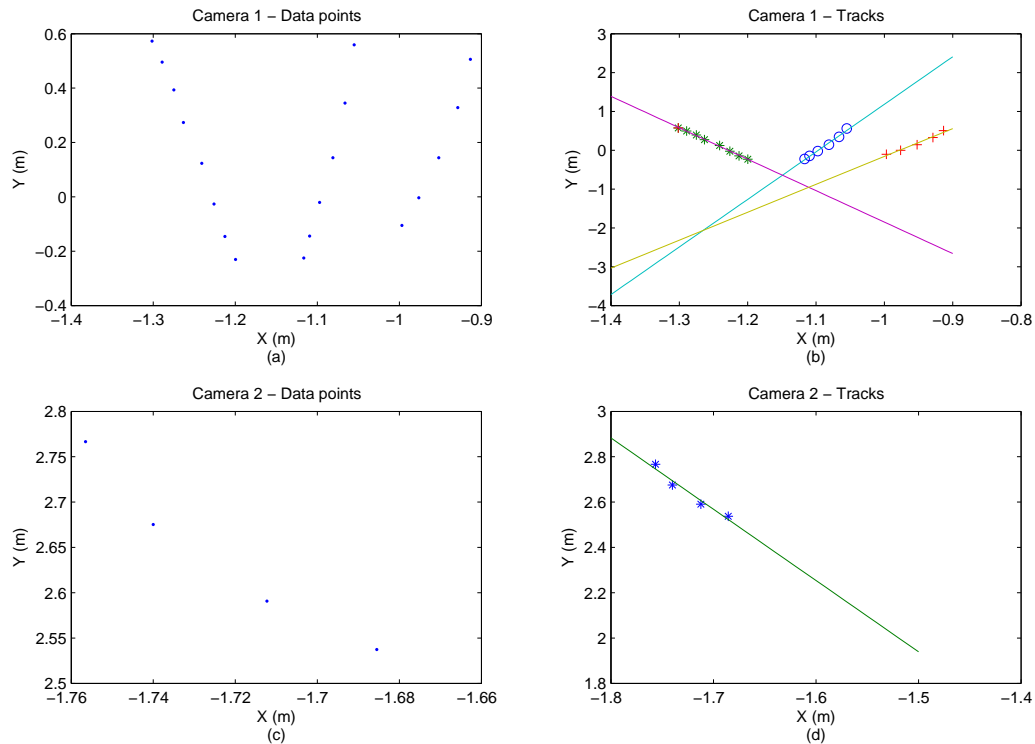


Figura 4.20: Conjunto de dados da trajetória estimada obtidos pela câmera 1 (a). Trajetórias e alvos identificados na câmera 1 (b). Pontos da trajetória estimada obtidos pela câmera 2 (c). Trajetórias e alvos identificadas na câmera 2 (d).

espaço de observação de duas câmeras. Os dados iniciais são obtidos e enviados para o processamento no módulo de associação de dados (Figura 4.20 (a) e (c)).

O algoritmo EM calcula no passo-E as probabilidades associadas a classificação de cada parâmetro de cor e movimento. O sistema então identifica o alvo e associa um peso (a probabilidade) para cada ponto obtido (Figura 4.20 (b) e (d)).

No passo-M, o método encontra a reta de melhor ajuste. Estes valores são estimados por meio de um processo de minimização quadrática. As Figuras 4.20 (b) e (d) apresentam a reta estimada e a conseqüente identificação da trajetória do alvo. A Tabela 4.5 apresenta os valores estimados como parâmetros da trajetória e o erro obtido pelo processo de associação entre os parâmetros.

Finalmente, os parâmetros são avaliados para se encontrar a melhor associação entre as trajetórias. O algoritmo EM fornece automaticamente o melhor ajuste entre os parâmetros. Porém, para uma avaliação independente do método, escolheu-se medir a associação entre as trajetórias por meio de uma métrica de similaridade.

| | a_i | b_i | h_i | Erro |
|-------|---------|---------|--------|--------|
| T_1 | 12.2478 | 13.4300 | 1.3024 | 0.0785 |
| T_2 | -8.0965 | -9.9468 | 1.2727 | 0.0393 |
| T_3 | 7.1880 | 7.0271 | 0.8156 | 0.0731 |
| T_4 | -3.1441 | -2.7766 | 1.4667 | 0.0354 |

Tabela 4.5: Resultados da minimização quadrática (M-Step).

| | T_1 | T_2 | T_3 | T_4 — |
|-------|--------|--------|--------|---------|
| T_4 | 2.7200 | 0.2863 | 2.7212 | - |

Tabela 4.6: Similaridade das medidas obtidas entre as trajetórias.

A similaridade entre as trajetórias estão sumarizadas na Tabela 4.3.2. Na análise deste experimento, encontrou-se o melhor ajuste entre as trajetórias T_4 na câmera 1 e T_2 na câmera 2.

A Figura 4.22 mostra todos os campos de visão das câmeras e a trajetória estimada. As trajetórias T_4 and T_2 formam uma trajetória de melhor ajuste, previamente determinada pelo algoritmo EM.

4.3.2.1 Análise do Reconhecimento da Trajetória por Similaridade de Padrões

Realizou-se um conjunto de 10 experimentos semelhantes ao que foi descrito nesta seção. Foram analisados um total de 1485 quadros capturados por três câmeras. A Tabela 4.7 mostra os valores obtidos para as associações entre as trajetórias.

O objetivo neste experimento foi analisar o reconhecimento da trajetória e do alvo e identificar o fator que mais influencia no processo de associação entre trajetórias e alvos. Particularmente, o Algoritmo EM é uma abordagem robusta, que garante a separabilidade do conjunto de dados e fornece um modelo de associação para a trajetória. Porém, essa associação é possível devido a análise dos resíduos, que na metodologia incorporam os parâmetros do modelo da trajetória e a informação de cor.

Para a condução desse experimento, considerou-se que cada trajetória T_i representa uma classe observada por três câmeras, mantendo uma câmera como referência. Cada trajetória fornece uma observação k_{ij} , referente ao objeto j observado pela câmera i , que agrega os parâmetros a e b do modelo da trajetória e o valor h , a matiz da cor. Partindo-se do princípio que os objetos descrevem uma trajetória retilínea e que a cor não varia,

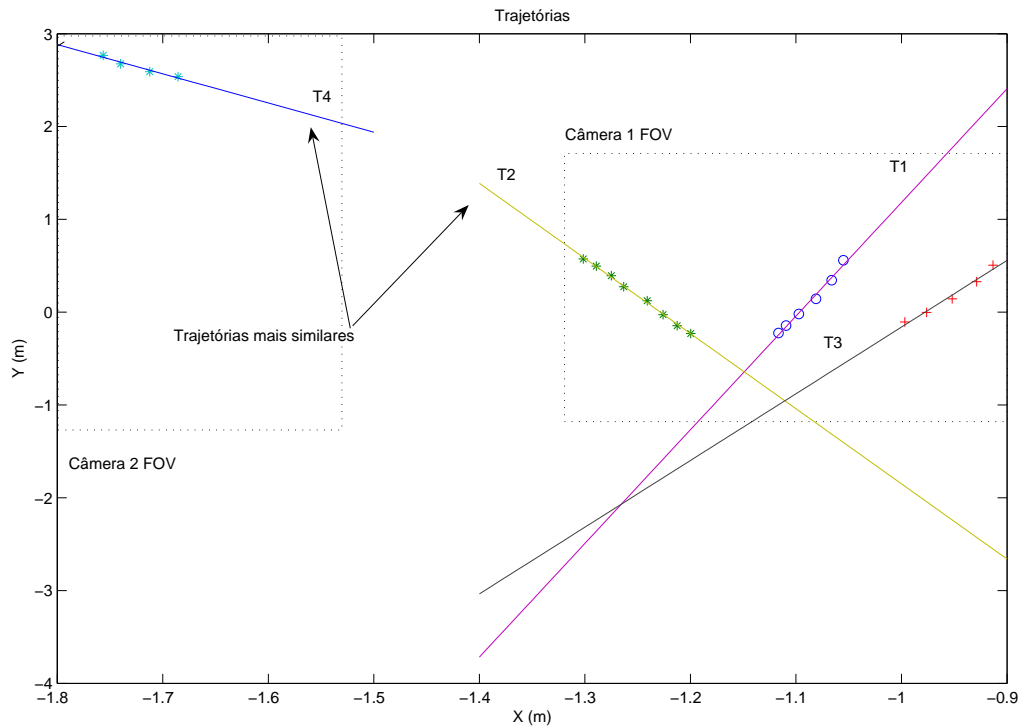


Figura 4.21: Trajetórias sobre uma região de monitoramento. As retas são as aproximações da trajetória em cada campo de visão das câmeras. Os retângulos tracejados representam uma aproximação grosseira do campo de visão das câmeras.

pode-se inferir que os parâmetros do modelo da trajetória e a cor do objeto manterão uma relação de similaridade nos campos de visão de câmeras diferentes.

A Tabela 4.7 apresenta o resultado da condução de dez experimentos. Cada experimento é composto pelo lançamento de quatro bolas de cores diferentes, formando quatro trajetórias que são observadas por três câmeras, porém, nem todas as trajetórias entram no campo visual de todas as câmeras. O objetivo é então descobrir quais trajetórias são observadas por câmeras diferentes.

A linha da tabela que especifica o experimento destaca a trajetória na câmera de referência. As outras linhas mostram o valor da relação de similaridade $S(k_{ij}, k_{jl})$ entre as observações das trajetórias na câmera de referência k_{1j} nas outras câmeras $k_{lj}, l = 1, 2$ em que alguma trajetória tenha sido observada. A similaridade entre as observações é dada pela Equação 4.13

$$S(k_{ij}, k_{jl}) \triangleq \frac{\langle k_{ij}, k_{lj} \rangle}{\langle k_{ij}, k_{ij} \rangle + \langle k_{lj}, k_{lj} \rangle - \langle k_{ij}, k_{lj} \rangle}, \quad (4.13)$$

onde k_{ij} e k_{lj} são as observações sobre o objeto i tomados pelas câmeras j e l , respectivamente e $\langle k_{ij}, k_{lj} \rangle$, é o produto escalar usual entre vetores.

Por meio dos resultados obtidos, verificou-se que trajetórias que mantêm uma relação de similaridade $S(k_{ij}, k_{jl}) < 0.3$ são as mesmas trajetórias observadas por câmeras separadas. O valor que mais influencia na determinação dessa relação é a cor. Os parâmetros do modelo linear da trajetória tendem a elevar essa relação e descaracterizar-la. Esse problema ocorre devido a má calibração das câmeras, que levam a determinação da posição errada do alvo no espaço, e a distorção radial das lentes. O experimento apresentado a seguir mostra a persistência da cor no sistema distribuído.

4.3.2.2 Persistência da Cor no Sistema

Este experimento usou como base toda a montagem do experimento anterior. Apenas a informação de cor foi obtida, separada em classes e armazenada para processamento posterior. O objetivo deste experimento é verificar se a informação de cor é persistente em todo o sistema, ou seja, se o padrão de cor muda e quanto muda entre câmeras diferentes.

Para desenvolver essa experiência, determinou-se um valor da matiz da cor representativa para cada classe de alvo, conforme apresentado na Tabela 4.8. A segunda coluna mostra o valor médio global representativo das cores, calculado para cada classe. As colunas seguintes mostram o valor médio da matiz para cada alvo observado nas câmeras C_1 , C_2 e C_3 , durante os rastreamentos desenvolvidos no experimento anterior. O erro foi calculado como a diferença absoluta entre o valor representativo da classe (coluna Matiz) e os valores obtidos para matiz em cada câmera (colunas Matiz C_1 , Matiz C_2 e Matiz C_3), sendo obtidos por meio de:

$$\text{Erro } C_1 = |\text{Matiz} - \text{Matiz } C_1|;$$

$$\text{Erro } C_2 = |\text{Matiz} - \text{Matiz } C_2|;$$

$$\text{Erro } C_3 = |\text{Matiz} - \text{Matiz } C_3|.$$

Os resultados mostram que a cor se mantém persistente em todas as câmeras. O desvio dos valores obtidos em relação aos pontos centrais das classes são baixos, apresentando valor mínimo de 0,0001 e máximo de 0,0153. A variação interclasse também é pequena, ficando com média de 0,05.

| | | | | |
|---------|---------|--------|--------|--------|
| Exp. 1 | T_1 | T_2 | T_3 | T_4 |
| T_4 | 2.7200 | 0.2863 | 2.7212 | - |
| Exp. 2 | T_1 | T_2 | T_3 | T_4 |
| T_1 | - | 0.1342 | 2.872 | 2.5434 |
| T_3 | 1.5292 | 2.3478 | - | 0.1987 |
| Exp. 3 | T_1 | T_2 | T_3 | T_4 |
| T_2 | 0.1022 | - | 2.9933 | 1.8762 |
| T_4 | 2.6532 | 0.2863 | 2.7212 | - |
| Exp. 4 | T_1 | T_2 | T_3 | T_4 |
| T_2 | 2.8232 | - | 2.0239 | 0.1923 |
| T_3 | 0.8762 | 1.5672 | - | 1.7888 |
| Exp. 5 | T_1 | T_2 | T_3 | T_4 |
| T_1 | - | 0.2362 | 1.3293 | 1.8732 |
| T_3 | 1.7430 | 0.2863 | - | 0.1283 |
| Exp. 6 | T_1 | T_2 | T_3 | T_4 |
| T_2 | 2.8220 | - | 0.0323 | 1.6543 |
| Exp. 7 | T_1 | T_2 | T_3 | T_4 |
| T_3 | 2.3425 | 0.1163 | - | 2.3493 |
| Exp. 8 | T_1 | T_2 | T_3 | T_4 |
| T_3 | 2.5430 | 0.1324 | - | 2.9982 |
| T_4 | 0.2323 | 0.8932 | 2.8331 | - |
| Exp. 9 | T_1 | T_2 | T_3 | T_4 |
| T_1 | - | 0.2863 | 2.8531 | 2.1837 |
| Exp. 10 | T_1 | T_2 | T_3 | T_4 |
| T_2 | 0.07649 | - | 2.3520 | 3.0196 |
| T_3 | 2.8881 | 2.2863 | - | 0.1923 |

Tabela 4.7: Experimentos para análise da similaridade entre observações de trajetórias.

| Cor | Matiz | Matiz C_1 | Erro C_1 | Matiz C_2 | Erro C_2 | Matiz C_3 | Erro C_3 |
|--------------|--------|-------------|------------|-------------|------------|-------------|------------|
| Azul escuro | 2,3521 | 2,3520 | 0,0001 | 2,3383 | 0,0138 | 2,3518 | 0,0003 |
| Amarelo | 0,7564 | 0,7544 | 0,0020 | 0,7021 | 0,0543 | 0,7549 | 0,0015 |
| Verde escuro | 0,9532 | 0,9500 | 0,0032 | 0,9212 | 0,0320 | 0,9499 | 0,0033 |
| Verde claro | 0,8694 | 0,8606 | 0,0088 | 0,8582 | 0,0112 | 0,8600 | 0,0094 |
| Vermelho | 4,3987 | 4,3843 | 0,0144 | 4,4077 | 0,0090 | 4,3834 | 0,0153 |

Tabela 4.8: Valor Médio Observado da Cor.

4.3.3 Discussão

O método mostrou-se suficientemente robusto no processo de detecção e associação das trajetórias. O Algoritmo EM garante a robustez no processo de associação ao determinar um modelo de regressão linear sobre o conjunto de dados. Porém, devido aos problemas decorrentes de calibração os parâmetros do modelo da trajetória avaliados errados, podem levar a não associação das trajetórias e comprometer o sistema. A solução adotada é usar um valor de limiar mais alto na (em torno de 20% no cálculo do resíduo médio) ou ampliar a importância da cor.

Para todo o conjunto de imagens analisadas, observou-se pouca variação na similaridade entre cor, movimento e trajetória. As cores, em particular, tendem a ser persistentes tornando o método suficientemente robusto para os casos onde não haja intercessões bruscas entre hipóteses sobre objetos caracterizados pela mesma distribuição de cor.

A aproximação da trajetória dos alvos por um modelo linear mostrou-se satisfatória na determinação das características geométricas das trajetórias descritas por alvos reais, que tentem a caminhar em direção seguindo o sentido das ruas, avenidas, estradas, calçadas ou corredores. Porém, o cuidado com a calibração da câmera é muito importante. Valores errados na determinação do modelo, podem levar a trajetórias completamente diferentes.

Os resultados obtidos mostraram a viabilidade do rastreamento mesmo quando a observação dos objetos não é possível por um relativo período de tempo entre câmeras com campos de visão disjuntos. Características visuais associadas as informações do movimento do alvo e a geometria da trajetória foram modeladas como padrões para o estabelecimento de correspondência entre padrões. Estas características são combinadas em um processo de associação e otimização, estimando-se um modelo da trajetória dos alvos pelo espaço monitorado.

Os experimentos descritos nas seções seguintes usam os modelos de trajetória estimadas como a informação a ser compartilhada pelo grupo de câmeras para a obtenção do posicionamento e observação dos alvos.

4.4 Avaliação do Método de Posicionamento

As aplicações experimentais reais objetivaram verificar a precisão do posicionamento de um robô em relação a uma trilha pré-determinada marcada no piso do laboratório. Trata-se de uma avaliação de cenário da metodologia. Supõe-se que o robô disponha de uma câmera e deve posicioná-la em relação a uma trajetória observada por outra câmera do sistema. A observação inicial é realizada por uma câmera fixa, que encontra os parâmetros

da trajetória e transmite, via rede, ao robô.

4.4.1 Posicionamento do Robô em Relação à Trajetória

Foi empregada uma câmera digital Sony, padrão IEEE-1394 com um jogo de lentes com distância focal $f = 8mm$, com capacidade de captura de imagens coloridas com tamanho de 320×240 pixels a uma taxa de 15 quadros por segundo. O processamento foi conduzido em dois computadores Pentium IV com $800MHz$, $256MHz$, na plataforma Windows XP®, ligados em rede. Como agente móvel utilizou-se um robô Pioneer P3-AT de quatro rodas controlado via porta serial por uma CPU do sistema, dotado de giroscópio e odômetro utilizados para a medir as rotações e translações executadas.

O objetivo do experimento é avaliar a performance do modelo em relação ao posicionamento da câmera, usando, em conjunto, os módulos de rastreamento, associação de dados e de planejamento.

Foi marcada com uma fita negra um trilha no chão do laboratório formando uma linha reta a 45° em relação ao referencial fixo, com equação $y = x + 0.5$. A câmera foi calibrada e posicionada sobre o tripé na origem do sistema de referência a uma altura de $1,30m$ com inclinação de 0° apontado para o piso.

Os parâmetros da posição da câmera são obtidos pela aplicação direta das equações de posicionamento. A rotação planejada é obtida por meio da avaliação Equação 4.14.

$$\theta_f = \theta_i - \arccos \left(\frac{\langle \mathbf{v}, \mathbf{v}_d \rangle}{\|\mathbf{v}\| \|\mathbf{v}_d\|} \right), \quad (4.14)$$

onde ϕ é o ângulo de rastreamento, θ_i é o ângulo de orientação inicial, θ_f é o ângulo de orientação final, \mathbf{v} é o vetor direção, \mathbf{v}_d é o vetor entre a posição inicial e o ponto de referência e $\langle \mathbf{v}, \mathbf{v}_d \rangle$ é o produto interno entre os vetores.

A translação por meio da Equação 4.15,

$$\mathbf{t} = \mathbf{v}_d - \mathbf{n}. \quad (4.15)$$

Os parâmetros de entrada para avaliação das equações são as coordenadas da posição do robô e sua orientação em relação ao eixo das abscissas do sistema de referência global. O robô foi posicionado inicialmente sobre a origem do sistema para marcar o seu referencial e em seguida foi direcionada para outras posições do ambiente.

A análise foi feita por meio do posicionamento do robô em um ponto pré-determinado e executada pela comparação entre resultados obtidos pelo planejador e o executado pelo robô para o seu posicionamento ortogonal em relação à trajetória.

| $\theta_{est}(rad)$. | $\theta_{med}(rad)$ | $Erro_{\theta}(rad)$ | $T_{est}(m)$ | $T_{med}(m)$ | $Erro_T(m)$ |
|-----------------------|---------------------|----------------------|--------------|--------------|-------------|
| 0,1745 | 0,2099 | 0,0354 | 1,3000 | 1,4218 | -0,1218 |
| 0,3491 | 0,3827 | 0,0336 | 1,1000 | 1,2362 | -0,1362 |
| 0,5236 | 0,4461 | -0,0775 | 1,7000 | 1,7235 | -0,0235 |
| 0,6109 | 0,5439 | -0,0670 | 0,8000 | 0,8183 | -0,0183 |
| 0,7854 | 0,7940 | 0,0086 | 1,0000 | 1,1144 | -0,1144 |
| 1,0472 | 1,0480 | 0,0008 | 1,0000 | 1,1083 | -0,1083 |
| 1,5708 | 1,6261 | 0,0553 | 1,2000 | 1,2977 | -0,0977 |
| 1,6581 | 1,6982 | 0,0401 | 1,5000 | 1,6131 | -0,1131 |
| 2,3562 | 2,3975 | 0,0413 | 1,7000 | 1,7995 | -0,0995 |
| 3,1416 | 3,2091 | 0,0675 | 1,5000 | 1,6325 | -0,1325 |
| 4,7124 | 4,7425 | 0,0301 | 1,5000 | 1,5408 | -0,0408 |
| 0,2618 | 0,3340 | 0,0722 | 1,0000 | 1,0629 | -0,0629 |

Tabela 4.9: Planejamento da posição e orientação das câmeras.

Bolas coloridas foram lançadas sobre a trajetória fixa com o objetivo do sistema de visão gerar a informação sobre a trajetória e transmitir para o robô. A partir de sua posição, o robô rotaciona, ficando em posição ortogonal em relação a trajetória e desloca-se a distância suficiente para cobri-la.

Foram realizados um conjunto de doze lançamentos de bolas sobre a trajetória fixa com o robô posicionado em locais e orientações diferente e pré-determinadas. A Tabela 4.9 sumariza os resultados obtidos.

O erro médio na orientação foi de 0,020 radianos, equivalente a aproximadamente $1,15^\circ$, que é satisfatório, principalmente se for considerada a pequena distância que o robô se desloca. O maior erro obtido foi de 0,0722 radianos, equivalente a 4° , e o menor erro em termos absolutos de 0,0008 radianos.

Com relação a translação, foi obtido um erro médio de $-0,0891m$, aproximadamente $9,0cm$ para menos. Devido as pequenas distâncias colocadas entre o robô e a trajetória, o sistema mantém uma boa precisão na hometria. Isso explica a ocorrência de um erro médio relativamente pequeno na translação do robô. O maior e o menor valor obtido em termos absolutos foi de $0,1362m$ e $0,0183m$ respectivamente. O desvio padrão ficou em $0,0419m$.

4.5 Considerações

Os experimentos avaliaram a metodologia sob o enfoque tecnológico e de cenário. O primeiro grupo de experimento caracterizou a precisão dos resultados e avaliou a performance dos métodos de rastreamento de alvos sob um contexto controlado, garantindo a definição da precisão e aplicabilidade dos métodos avaliados.

Um segundo grupo de experimentos caracterizou a performance da metodologia por meio da construção de um cenário formado por um aparato computacional montado sob uma rede câmeras onde alvos eram lançados sob uma região de monitoramento. Por fim, avaliou-se o posicionamento de um robô frente a informação obtida por meio de uma câmera.

Cada componente do sistema mostrou sua efetiva atuação no rastreamento, na associação entre observações e trajetórias e no planejamento do posicionamento das câmeras. As métricas e os procedimentos experimentais foram previamente definidos, compondo um protocolo experimental que favoreceu a análise e avaliação dos métodos que compõem o sistema.

O método de rastreamento foi avaliado sobre diferentes aspectos. No primeiro instante, procurou-se analisar o comportamento do erro médio quadrático frente o aumento do número de partículas. Para a efetivação desse experimento, montou-se um aparato baseado na determinação de uma trilha fixa, composta por um trilho circular de trem de brinquedo com dimensões conhecidas. Essa trilha é mapeada no espaço da imagem e os erros são estimados por meio da comparação entre as distâncias estimadas e o valor fixo do raio do trilho, em pixels. Os resultados obtidos mostraram o decaimento do valor do erro médio quando se aumenta o número de partículas, comprovando os resultados teóricos e garantindo a aplicabilidade do método.

A performance do método para o rastreamento de múltiplos objetos também foi analisada e avaliada frente o aumento do número de objetos rastreados simultaneamente. Os resultados obtidos comprovaram a eficácia da aplicação do método, garantindo sua utilização nos propósitos da ODC. A experimentação deu-se de forma mais livre, sem controle de parâmetros externos, sob iluminação do ambiente e sem determinação prévia de trilhas. Fundamentalmente, o conjunto de experimentos desenvolvidos mostrou que método de rastreamento garante:

- iniciação automática do processo de rastreamento e formação da trajetória;
- a finalização automática do processo de rastreamento;
- o rastreamento de múltiplos alvos em tempo real; e

- a robustez suficiente para o rastreamento de múltiplos alvos para um conjunto grande de dados.
- precisão no posicionamento em relação às trajetórias estimadas.

O funcionamento conjunto do rastreador e da identificação de trajetórias nos campos visuais disjuntos das câmeras foi avaliado por meio da análise da correspondência dos parâmetros de cor e da trajetória obtidos sobre os diferentes FOVs. Os resultados são satisfatórios, porém, em alguns casos, distorção radial das lentes das câmeras e alguns problemas na calibração levaram a obtenção de valores com precisão mais baixa, porém aceitável frente aos testes desenvolvidos. Como a associação está sujeita a aceitação de um valor menor ou igual que um dado limiar, o ajuste do parâmetro de um limiar garante a robustez necessária para a associação das trilhas.

A atuação dos robôs na observação dinâmica também foi analisada por meio da experimentação real. Empregou-se um robô Pioneer IV em uma série de verificações do seu posicionamento a partir das informações capturadas por uma câmera monitoradora de uma determinada região no ambiente. Verificou-se que o modelo de planejamento para o posicionamento do robô se adaptou muito bem nos casos analisados, garantindo a precisão suficiente para a validação do modelo conceitual proposto.

No geral, o processo experimental foi complexo e trabalhoso. Buscou-se por meio das análises a garantia da validade estatística do modelo computacional/matemático proposto. Com base nos resultados obtidos e em todo o conjunto de programas produzidos, novos experimentos serão conduzidos visando a aplicabilidade da metodologia em sistemas de segurança. O capítulo seguinte irá apresentar as considerações finais e as extensões para trabalhos futuros.

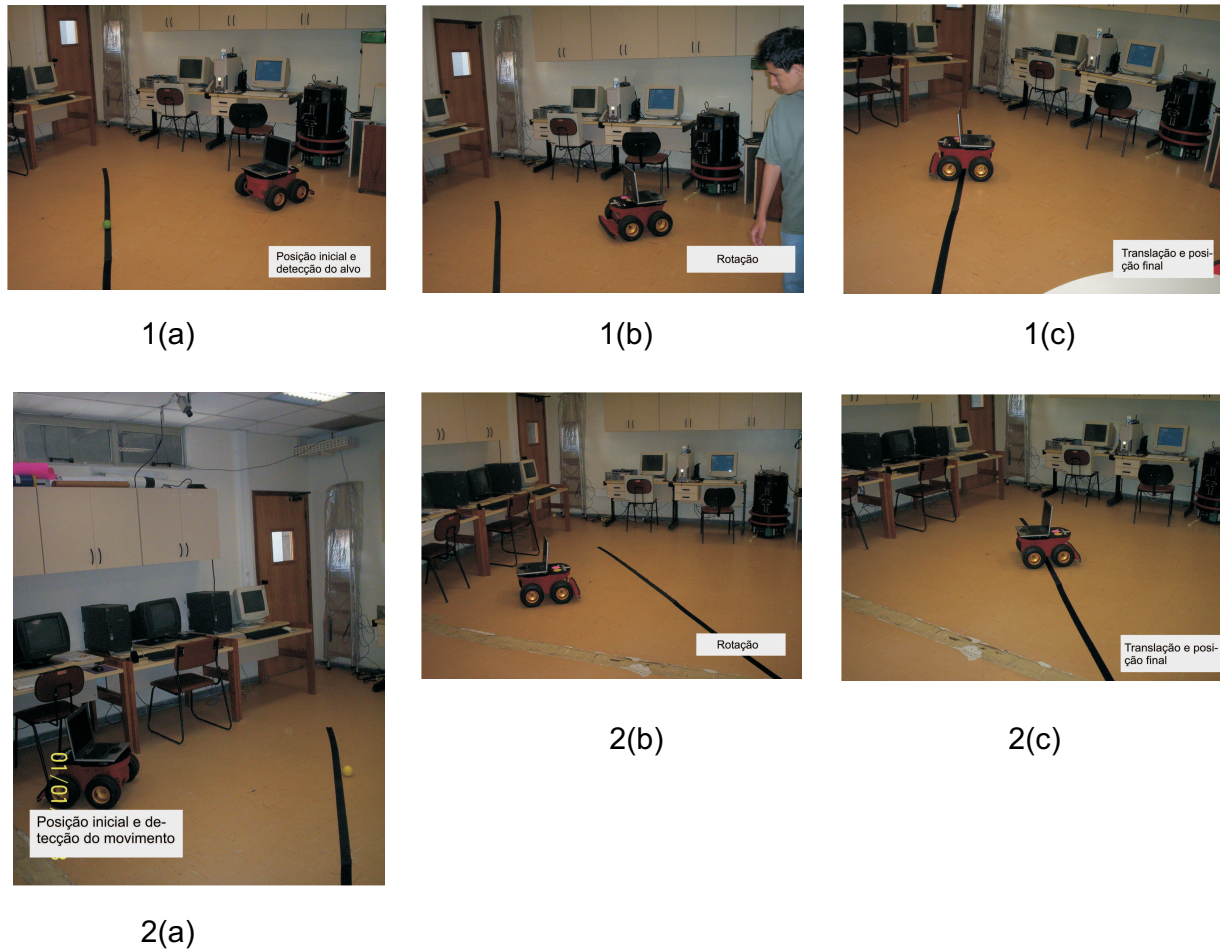


Figura 4.22: Fotografias do posicionamento do robô em relação à trajetória. Como trajetória de referência e guia para lançamento dos alvos, usou-se a fita preta marcada sobre o chão do laboratório. A bola colorida é lançada e o seu movimento capturado pela câmera (foto 1(a)). Os parâmetros obtidos pela câmera são transmitidos ao robô que planeja sua posição final, executa o movimento de rotação (foto 1(d)) e em seguida a translação para a posição final (foto 1(c)). A seqüência de fotos de 2(a) a 2(c) ilustra o mesmo procedimento com o robô em outra posição.

Capítulo 5

Conclusão e Trabalhos Futuros

C'est l'histoire d'un mec qui tombe d'un immeuble de cinquante étages; au fur et à mesure de sa chute il se répète sans cesse pour se rassurer: jusqu'ici tout va bien, jusqu'ici tout va bien, jusqu'ici tout va bien... mais l'important, c'est pas la chute c'est l'atterrissage.

*Trecho do diálogo entre personagens do filme “La Haine”
de Mathieu Kassovitz, cineasta francês.*

Nesta tese abordou-se o problema da observação da trajetória de objetos por meio de um sistema de câmeras distribuídas. Problemas dessa natureza são complexos e requerem soluções eficientes para o tratamento e processamento da informação visual que viabilizem atuação segura e rápida do agente tomador de decisão.

A abordagem foi desenvolvida por meio de uma metodologia que enfoca o problema sob o ponto de vista distribuído em um contexto da visão computacional ativa. Desenhou-se, assim, uma arquitetura funcional para a coordenação da ação entre as câmeras baseada no compartilhamento informação visual. A solução desenvolvida veio por meio da implementação de uma metodologia que integra a informação visual em uma estrutura comum a todas os membros do sistema, viabilizando o posicionamento das câmeras na tarefa de observar a trajetória dos alvos.

Desenvolveu-se, ainda, uma abordagem eficiente para a fusão da informação no espaço de probabilidade de um filtro de partículas e o consequente aprimoramento da iniciação de objetos no rastreamento por meio dos métodos seqüenciais de Monte Carlo. Essa melhoria garante menor custo computacional e viabilidade em aplicações tempo real.

A estrutura computacional implementada propiciou um esquema de rastreamento distribuído em uma rede de câmeras sem sobreposição dos campos de visão. Essa estrutura favorece a captura de diferentes vistas simultâneas sobre a área de observação, fornecendo informações que facilitam e entendimento da dinâmica dos alvos na cena global.

O movimento das câmeras é orientado por meio de uma estratégia para planejamento de posição montada sobre uma estrutura topológica que agrega a capacidade de percepção sobre toda a região monitorada, tornando possível o posicionamento eficiente do sensor a partir da otimização de critérios pré-definidos impostos sobre a capacidade óptica do sensor e as restrições providas do ambiente.

As aplicações experimentais desenvolvidas em nível de avaliação tecnológica mostraram a viabilidade da metodologia, principalmente sobre o contexto do rastreamento visual de múltiplos objetos e da integração da informação sensorial produzida pelo grupo de câmeras. Os experimentos reais desenvolvidos levaram a análise criteriosa da precisão do método de rastreamento e do procedimento de reconhecimento de trajetórias sobre câmeras com campo de visão disjuntos. O esquema de posicionamento foi avaliado em um cenário restrito, porém mostrou-se a aplicabilidade do método para casos mais gerais devido aos resultados produzidos e a robustez comprovada.

5.1 Limitações

As maiores limitações do trabalho são inerentes à necessidade de processos de calibração, a falta de tratamento para a ocorrência simultânea de movimento entre as câmeras e os alvos, a necessidade da integração de mais características visuais e, principalmente, a falta de experimentação real em nível operacional.

A forte dependência da calibração das câmeras torna todo o processo de montagem complexo e oneroso. Gasta-se muito tempo na determinação dos parâmetros intrínsecos e no posicionamento das câmeras no ambiente de monitoramento. A condução errada da montagem do aparato experimental ou na determinação dos parâmetros das câmeras levam a erros de difícil identificação e tratamento. Um erro na calibração da câmera leva à determinação de medidas de profundidade (distância entre a câmera e o alvo) erradas e a conseqüente degeneração de todo o sistema. A deformação radial das lentes torna complicado o alinhamento das trajetórias em câmeras com campo de visão separados.

Embora o sistema preveja a dinâmica das câmeras pelo ambiente, as imagens só podem ser obtidas a partir da câmera estacionada em uma posição fixa. O sistema identifica o movimento dos alvos por meio da técnica convencional de subtração de referência, necessitando de um modelo de fundo estático e sem objetos em movimento. A suposição da câmera iniciar o processo de aquisição de imagens somente a partir de uma posição estática é inviável na maioria das aplicações reais.

Na metodologia, limita-se a obtenção da cor como característica visual para a identi-

fição do alvo. Os alvos são de tamanho fixo, previamente conhecidos e com cor uniforme. Alvos reais são, na maioria dos casos, pessoas ou veículos que apresentam um número maior de características visuais. Para agregar robustez ao sistema e torná-lo aplicável no contexto da segurança é necessário que seja usado um conjunto maior de características visuais como a forma e a textura, por exemplo.

Os procedimentos experimentais foram desenvolvidos no contexto da avaliação tecnologia e de cenário. Para a utilização prática da metodologia faz-se necessário que avaliações operacionais sejam implementadas e verificadas sob a especificação concreta de uma tarefa definida por um usuário final. Procedimentos experimentais mais elaborados devem requerer um número maior de câmeras, coordenação e uso de robôs e a aplicação em ambientes com obstáculos.

5.2 Extensões

Muitas das extensões deste trabalho estão concentradas na eliminação ou relaxamento das limitações, mencionadas na seção anterior. A extensão natural é considerar maior grau de liberdade as câmeras e as implicações do movimento dos objetos contidos em um espaço tridimensional.

A aplicação de técnicas de autocalibração podem ser aplicadas e viabilizar maior robustez ao sistema. São várias as técnicas de visão computacional podem ser aplicadas para resolver o problema da aquisição da informação de profundidade. Outras tantas podem ser aplicáveis ao processo de autocalibração, tornando o sistema mais versátil em termos de aquisição e determinação dos parâmetros da câmeras. O emprego efetivo de robôs móveis e de seus diferentes sensores amplia a capacidade sensorial do sistema e a obtenção mais acurada da informação tridimensional, garantindo maior precisão aos resultados sensoriados.

Finalmente, enfatiza-se a necessidade de uma análise mais profunda do funcionamento da metodologia sob cenários mais complexos, em condições adversas e com alvos reais formados por pessoas ou veículos. A execução de uma avaliação operacional de metodologia reforça a definição precisa de requisitos para aplicações reais de seu uso. Esse conjunto de avaliações devem usar dados reais obtidos sem o controle específico ou parametrização prévia e não necessariamente repetitivas, porém conclusivas na realização de objetivos.

5.3 Trabalhos Futuros

Os trabalhos a serem desenvolvidos em futuro próximo objetivam seguir pelas três linhas de pesquisa evidenciadas nesta tese:

- Sistemas multi-câmeras e visão distribuída.

O interesse pelos sistemas formados por múltiplas câmeras para o monitoramento de regiões muito grandes tem crescido bastante nos últimos anos. Câmeras de segurança são percebidas em todos os lugares: locais públicos, shopping centers, escritórios, bancos, estacionamentos, residências, escolas, aeroportos, e em uma variedade de outros locais. A quantidade de informação obtida e armazenada por essas redes de câmeras é enorme, sugerindo um campo de pesquisa relativamente novo e carente de desenvolvimento de novas tecnologias.

As aplicações recentes de redes de câmeras em sistemas de segurança sugerem o desenvolvimento de metodologias que vão da simples análise de imagens, passam por técnicas sofisticadas de reconhecimento de face, por exemplo, e chegando até ao desenvolvimento de complexos sistemas que analisam o comportamento de pessoas em um ambiente monitorado. Os sistemas de controle de tráfego são bons exemplos da necessidade de sistemas de múltiplas câmeras com visão distribuída. A automatização das tarefas de rastreamento, identificação, classificação e contagem de veículos são evidentes para o aprimoramento e melhora do controle de tráfego.

- Visão ativa.

Os sistemas de visão ativa em vez de observar passivamente o ambiente, operem em tempo real e adquirem informação visual com maior precisão. As técnicas de visão ativa interagem diretamente com os objetos, modificando parâmetros ópticos, como o zoom ou a abertura automática da íris da câmera, ou modificando a direção da visada garantindo a melhor posição da câmera para a aquisição de imagens. A visão ativa caracteriza-se como uma linha de pesquisa desafiante, onde novos problemas podem ser identificados, tratados e adequadamente aplicados em áreas como a robótica móvel, em particular no auxílio a navegação de veículos robóticos por meio da visão. A necessidade de adaptação do sistema de percepção visual dos robôs em diferentes ambientes sob diferentes condições de luminosidade é um problema no qual a utilização dos métodos e das técnicas de visão ativa se fazem necessários.

A seguir, são delineadas propostas de trabalho que se enquadram nas linhas de pesquisa apresentadas nos parágrafos anteriores.

5.3.1 Observação Cooperativa Tridimensional

A observação cooperativa tridimensional é uma extensão da metodologia desenvolvida nesta tese. Objetiva-se nesta proposta a identificação da estrutura tridimensional de objetos que compõe uma cena utilizando as habilidades de uma rede de câmeras móveis, assentadas sobre robôs. O problema aqui é encontrar a posição das câmeras para se obter a estrutura tridimensional dos objetos sob observação. A idéia geral desta proposta é tornar mais eficiente o processamento da correspondência entre diversos pares de imagens capturados pelas câmeras e que compartilham as sobreposições de diferentes campos de visão.

Na técnica convencional de estéreo com ampla linha base se utilizam vários pares estéreo com diferentes tamanhos de linhas base, geradas por um deslocamento lateral de uma única câmera previamente calibrada. Em geral, o processo de correspondência é realizada por meio da soma da diferença dos quadrados (SSD – *Sum of Squared Difference*) entre os múltiplos pares estéreo. As funções SSD para pares estéreo individuais são representadas em relação ao inverso da distância, que são simplesmente somadas para produzir a soma dos SSDs. A estimativa da profundidade é calculada encontrando-se o mínimo da função resultante [Forsyth e Ponce (2002)].

A proposta de trabalho futuro é a inserção ao método original da aquisição de pares estéreo com deslocamentos guiados por meio de uma função objetivo que garanta pontos de visada ótimos através de deslocamentos horizontais, verticais ou transversais executados por robôs móveis que agem de forma cooperativa na troca de informações visuais obtidas durante o deslocamento.

A construção da função objetivo pode ser conduzida por meio da determinação da geometria epipolar estimada a partir de duas câmeras separadas e o consecutivo mapeamento dessa informação em um modelo de mapa tridimensional. Mais precisamente, se I_1 e I_2 são duas imagens separadas por uma rotação e translação \mathbf{R} e \mathbf{t} , a posição relativa das câmeras é obtida por meio da minimização do erro relacionado com a rotação \mathbf{R} e o vetor de translação \mathbf{t} , que é dado por

$$E(\mathbf{R}, \mathbf{t}) = \sum_i d(F_i^1, T_{\mathbf{R}, \mathbf{t}}(F_i^2)) \quad (5.1)$$

Onde F_i^1 e F_i^2 são as características obtidas nas imagens I_1 e I_2 respectivamente, e $T_{\mathbf{R}, \mathbf{t}}(F_i^2)$ é a projeção de F_i^2 em I_1 para uma dada a rotação \mathbf{R} e uma translação \mathbf{t} .

Em geral, se as características obtidas são pontuais d é a distância usual entre os pontos da imagem e as linhas epipolares. Se, entretanto, as características são densas, obtidas a partir de retalhos planares, $T_{\mathbf{R}, \mathbf{t}}$ é uma homografia e d é a correlação entre os

retalhos planares considerados.

O objetivo geral desta proposta é, em síntese, a remoção da ambigüidade das informações visuais e o aumento da precisão do método estéreo por meio da utilização de câmeras movidas por robôs, evitando os problemas associados com a orientação e a oclusão dos elementos da cena. O maior desafio desta proposta é o desenvolvimento uma abordagem cooperativa na qual o posicionamento dos robôs favoreça a minimização do erro E , criando linhas base ajustáveis e integrando a informação visual tridimensional da cena.

5.3.2 Visão Distribuída em Rede de Sensores Visuais sem Fio

O problema que aqui proposto consiste na especificação e no desenvolvimento de um mecanismo de alocação de tarefas a serem executadas em cada camada da hierarquia da rede, baseada na informação visual sensoriada. O objetivo é agregar a informação processada nas diferentes camadas com a informação visual provida pelas câmeras, por meio de um mecanismo de alocação de tarefas, obtendo-se uma representação mais confiável do evento monitorado. O problema pode ser caracterizado da seguinte maneira:

Seja R uma rede de sensores sem fio multicamada e multimodal. As camadas possuem capacidade de processamento e funcionalidades específicas. Elementos de R são assumidos como homogêneos em suas camadas e são capazes de comunicarem entre si em uma mesma camada ou em camadas distintas. Seja ainda A uma aplicação sobre esta rede. Esta aplicação necessita que um conjunto de tarefas, $T = \{t_1, t_2, \dots, t_n\}$, sejam executadas pelas camadas da rede, $C = c_1, c_2, \dots, c_n$. Uma camada pode ter capacidade de realizar mais de uma tarefa e uma mesma tarefa pode ser realizada por mais de uma camada. Como alocar dinamicamente tarefas as camadas em função dos objetivos da aplicação A considerando critérios de qualidade de serviço como, redução de latência, economia de energia, balanceamento de carga, confiabilidade do serviço prestado, dentre outras.

A idéia principal é que alocação de tarefas as camadas de sensores é determinada por meio das interações localizadas entre nós de uma mesma camada, programados para reduzir tráfego, coordenar sensoriamento e direcionar interesses.

5.4 Considerações Finais

Esta tese teve como objetivo científico o estudo da integração entre percepção visual do movimento, ação e cooperação em sistemas distribuídos de visão computacional ativa, procurando enfatizar a aplicação no contexto demandado pelos sistemas de segurança. Especificamente, tratou-se da observação de alvos por meio de um sistema distribuído uti-

lizando um esquema de cooperação entre as câmeras e uma estratégia de planejamento de posição que leva em consideração as restrições da câmera e do ambiente.

Mostrou-se que é possível a utilização conjunta de um grupo de câmeras na observação da trajetória de alvos por meio compartilhando informação visual, com o ganho do conseqüente aumentando do espaço perceptivo do grupo.

Porém, entende-se que este trabalho é apenas um começo, uma etapa inicial a partir da qual novas pesquisas sobre temas correlacionados serão derivadas e desenvolvidas. Com a finalização desta etapa, se buscará pela inclusão das soluções tecnológicas providas aqui em sistemas de proteção e segurança reais, como o SIVAM, mostrando a flexibilidade da integração de sistemas de visão computacional distribuída em sistemas de proteção reais de grande porte.

É certo que a efetivação futura da integração da tecnologia aqui desenvolvida irá aprimorar consideravelmente os resultados e o uso de sistemas de segurança em regiões grandes e de difícil acesso. Porém, é mais evidente ainda que este trabalho abre uma linha de atuação que influenciará decisivamente o desenvolvimento tecnológico na Amazônia Brasileira.

Apêndice A

Métodos Seqüenciais de Monte Carlo

Os métodos de predição conhecidos como *Métodos Seqüencias de Monte Carlo* (MSMC), ou *Filtro de Partículas*, fornecem soluções gerais para os problemas de estimação nos quais a aproximação Gaussiana ou a linearização do modelo não são possíveis ou levam a degradação da performance do sistema.

A idéia fundamental do filtro de partículas é a utilização do método de Monte Carlo como solução para o problema de estimação ótima. Um filtro de partículas fornece uma solução Bayesiana aproximada para um modelo de tempo discreto por meio da atualização de uma descrição aproximada da função de densidade de probabilidade posterior [A. Doucet e Andrieu (2000); Hue (2002)].

Para se formalizar a definição de um filtro de partículas, assume-se que os estados do sistema são modelados como Markovianos¹, não lineares e não Gaussianos. Assim, sejam:

$$\mathbf{x}_{0:k} \triangleq \{x_0, x_1, \dots, x_k\},$$

o vetor de estado do sistema a partir do tempo inicial até o tempo k e

$$\mathbf{z}_{1:k} \triangleq \{z_1, z_2, \dots, z_k\},$$

o vetor de observações obtidas até o tempo k para qualquer função de interesse $f_k : \mathbf{x}^{k+1} \rightarrow \mathbb{R}^n$ integrável em relação a $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$.

Em um filtro de partículas, as funções de densidade de probabilidade de variáveis aleatórias não paramétricas são representadas por um conjunto de partículas $x_{0:k,i=1}^{(i)N}$ (amostras aleatórias identicamente distribuídas) de acordo com

¹Um sistema é dito Markoviano se este preserva a propriedade de Markov, ou seja, a distribuição de probabilidades dos estados futuros, dado o estado corrente, é independente dos estados passados.

$$p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}).$$

Uma estimação desta distribuição é dada por

$$p_N(d\mathbf{x}_{0:k} | \mathbf{z}_{1:k}) = \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x}_{0:k} - \mathbf{x}_{0:k}^{(i)}), \quad (\text{A.1})$$

onde δ denota a função delta de Dirac.

Devido a propriedade da invariância do método de Monte Carlo [Frank (2003)], sabe-se que

$$I_N(f_k) = \int f_k(\mathbf{x}_{0:k}) p_N(d\mathbf{x}_{0:k} | \mathbf{z}_{1:k}) = \frac{1}{N} \sum_{i=1}^N f_k(\mathbf{x}_{0:k}^{(i)}). \quad (\text{A.2})$$

A vantagem do método de Monte Carlo é óbvia: pode-se determinar qualquer estimação $I_N(f_k)$ com taxa de convergência independente da dimensão do integrando. Porém, infelizmente, obter amostras de uma distribuição conjunta *a posteriori* $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$ de forma eficiente é uma tarefa, em geral, impossível [Bar-Shalom et al. (2001); Frank (2003)].

Na maioria dos casos, empregam-se métodos de Monte Carlo via cadeias de Markov (MCCM), ainda com o objetivo de se obter uma amostra da distribuição *a posteriori* e calcular estimativas amostrais de características desta distribuição [Frank (2003)]. Porém, os MCCM são métodos iterativos e pouco aplicáveis em problemas de estimação recursiva [Frank (2003); Bar-Shalom (1992)].

Uma alternativa aos MCCM é a utilização das técnicas de amostragem e reamostragem. Estas técnicas objetivam gerar valores em duas etapas: na primeira etapa geram-se valores de uma distribuição auxiliar conhecida; na segunda etapa utiliza-se um mecanismo de correção para que os valores sejam representativos (ao menos aproximadamente) da distribuição *a posteriori*. As principais técnicas de amostragem estão apresentadas em seguida.

A.0.0.1 Amostragem por Importância

Na amostragem por importância, introduz-se uma distribuição arbitrária $q(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$ chamada de *distribuição de amostragem por importância*, *função de importância* ou *densidade de importância*. Sabendo-se que

$$I(f_k) = \int f_k(\mathbf{x}_{0:k}) \frac{p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})}{q(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})} q(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}) d\mathbf{x}_{0:k}, \quad (\text{A.3})$$

$$= \int f_k(\mathbf{x}_{0:k}) w^*(\mathbf{x}_{0:k}) q(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}) d\mathbf{x}_{0:k}, \quad (\text{A.4})$$

onde

$$w^*(\mathbf{x}_{0:k}) = \frac{p(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})}{q(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})} = \frac{p(\mathbf{z}_{1:k}|\mathbf{x}_{0:k})p(\mathbf{x}_{0:k})}{p(\mathbf{z}_{1:k})q(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})} \quad (\text{A.5})$$

é conhecido como *peso de importância verdadeiro*. Assim, podem-se simular N amostras independentes e identicamente distribuídas (*iid*) de acordo com $q(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$. Uma possível estimativa de $I(f_k)$ é

$$\hat{I}_N^*(f_k) \triangleq \frac{1}{N} \sum_{i=1}^N f_k(\mathbf{x}_{0:k}^{(i)}) w_k^{*(i)}. \quad (\text{A.6})$$

Em geral, esta estimativa não pode ser empregada diretamente, devido a necessidade da constante de normalização $p(\mathbf{z}_{1:k})$.

Se tomarmos o peso de importância não normalizado

$$w(\mathbf{x}_{0:k}) = \frac{p(\mathbf{z}_{1:k}|\mathbf{x}_{0:k})p(\mathbf{x}_{0:k})}{q(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})} \propto w_k^*(\mathbf{x}_{0:k}), \quad (\text{A.7})$$

fica-se com

$$I(f_k) = \int f_k(\mathbf{x}_{0:k}) \frac{w(\mathbf{x}_{0:k})}{p(\mathbf{z}_{1:k})} q(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}) d\mathbf{x}_{0:k}. \quad (\text{A.8})$$

Assim, $p(\mathbf{z}_{1:k})$ pode ser expresso por

$$p(\mathbf{z}_{1:k}) = \int p(\mathbf{z}_{1:k}|\mathbf{x}_{0:k})p(\mathbf{x}_{0:k}) d\mathbf{x}_{0:k}, \quad (\text{A.9})$$

$$= \int w_k(\mathbf{x}_{0:k}) q(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}) d\mathbf{x}_{0:k}, \quad (\text{A.10})$$

obtendo-se

$$I(f_k) = \frac{\int f_k(\mathbf{x}_{0:k}) w_k(\mathbf{x}_{0:k}) q(\mathbf{x}_{0:k}|\mathbf{z}_{1:k})}{\int w_k(\mathbf{x}_{0:k}) q(\mathbf{x}_{0:k}|\mathbf{z}_{1:k}) d\mathbf{x}_{0:k}}, \quad (\text{A.11})$$

tal que

$$\hat{I}_N(f_k) = \frac{\frac{1}{N} \sum_{i=1}^N f_k(\mathbf{x}_{0:k}^{(i)}) w(\mathbf{x}_{0:k}^{(i)})}{\frac{1}{N} \sum_{j=1}^N w(\mathbf{x}_{0:k}^{(j)})}, \quad (\text{A.12})$$

$$= \sum_{i=1}^N f_k(\mathbf{x}_{0:k}^{(i)}) \tilde{w}_k^{(i)}, \quad (\text{A.13})$$

e

$$\tilde{w}_k^{(i)} = \frac{w(\mathbf{x}_{0:k}^{(i)})}{\sum_{j=1}^N w(\mathbf{x}_{0:k}^{(j)})}, \quad (\text{A.14})$$

são os *pesos por importância normalizados*.

Os pesos de importância $w_k^{*(i)}$ serão substituídos pela estimativa

$$\hat{w}_k^{*(i)} = N \tilde{w}_k^{(i)}. \quad (\text{A.15})$$

Este método de integração também pode ser interpretado como um método de amostragem onde a distribuição de probabilidade *a posteriori* $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$ é aproximada por

$$\hat{P}_N(d\mathbf{x}_{0:k} | \mathbf{y}_{1:k}) = \sum_{i=1}^N w_k^{(i)} \delta(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}), \quad (\text{A.16})$$

e a integral $\hat{I}_N(f_k)$ é a função $f_k(\mathbf{x}_{0:k})$ integrada em relação a medida empírica $\hat{P}_N(d\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$:

$$\hat{I}_N(f_k) = \int f_k(\mathbf{x}_{0:k}) \hat{P}_N(d\mathbf{x}_{0:k} | \mathbf{z}_{1:k}). \quad (\text{A.17})$$

A.0.0.2 Amostragem por Importância Sequencial

Uma vez que se tem como construir uma aproximação de $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k-1})$ e se deseja encontrar uma aproximação $\hat{P}_N(d\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$ de $p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$ com um novo conjunto de amostras sem modificar as trajetórias anteriormente simuladas. Emprega-se, então, a fatoração da função de importância $q(\mathbf{x}_{0:k} | \mathbf{z}_{1:k})$, tal que

$$q(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}) = q(\mathbf{x}_{0:k-1}) q(\mathbf{x}_k | \mathbf{x}_{0:k-1}) \quad (\text{A.18})$$

$$= q(\mathbf{x}_0) \prod_{j=1}^k q(\mathbf{x}_j | \mathbf{x}_j | \mathbf{x}_{0:j-1}, \mathbf{z}_{1:j}). \quad (\text{A.19})$$

Lembrando que a transição de estado é modelada como um processo Markoviano², pode-se escrever

$$p(\mathbf{x}_{0:k}) = p(\mathbf{x}_0) \prod_{j=1}^k p(\mathbf{x}_j | \mathbf{x}_{j-1}) \quad (\text{A.20})$$

e

$$p(\mathbf{z}_{1:k} | \mathbf{x}_{0:k}) = \prod_{j=1}^k p(\mathbf{z}_j | \mathbf{x}_j). \quad (\text{A.21})$$

Usando a Equação A.7 obtém-se

$$w_k^{(i)} = w_{k-1}^{(i)} \frac{p(\mathbf{z}_k | \mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)})}{q(\mathbf{x}_k^{(i)} | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})}. \quad (\text{A.22})$$

O Algoritmo 7 AMOSTRAGEMIMPORTANCIA mostrado a seguir, sumariza o processo de amostragem sequencial por importância.

Algoritmo 7 AMOSTRAGEMIMPORTANCIA

- 1: **para** $i = 1, \dots, N$ **faça**
 - 2: $\mathbf{x}_k^{(i)} \sim q(\mathbf{x}_k^{(i)} | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})$
 - 3: $\mathbf{x}_{0:k}^{(i)} \triangleq \{\mathbf{x}_{0:k-1}^{(i)}, \mathbf{x}_k^{(i)}\};$
 - 4: $w_k^{(i)} = w_{k-1}^{(i)} \frac{p(\mathbf{z}_k | \mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)})}{q(\mathbf{x}_k^{(i)} | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{z}_{1:k})};$
 - 5: $\tilde{w}_k^{(i)} = \frac{w_k^{(i)}}{\sum_{i=1}^N w_k^{(i)}};$
 - 6: **fim para**
-

Um problema grave está associado à amostragem por importância sequencial. Durante o decorrer do processo a variância dos pesos associados às partículas tende a crescer, levando, no limite, ao fenômeno da degeneração das partículas. Com a ocorrência da degeneração, apenas algumas poucas partículas terão pesos normalizados próximos a um, enquanto a grande maioria das partículas terão seus pesos muito próximos a zero.

Assim, devido a degeneração, o Algoritmo AMOSTRAGEMIMPORTANCIA pode tornar-se instável, devido ao elevado custo computacional para atualizar partículas cuja a contribuição é muito próxima do zero.

²Um processo estocástico é dito Markoviano se a probabilidade de transição de algum estado particular \mathbf{x}_{k+1} depende somente do estado anterior \mathbf{x}_k .

O *número efetivo de partículas* N_{ef} é uma maneira de medir a degeneração das partículas, dado por meio de [Frank (2003)]:

$$N_{ef} = \frac{N}{1 + Var(w^*(\mathbf{x}_{0:k}))}. \quad (\text{A.23})$$

O valor de N_{ef} não pode ser obtido diretamente, mas uma aproximação \hat{N}_{ef} pode ser encontrada por meio de

$$\hat{N}_{ef} = \frac{N}{\frac{1}{N} \sum_{i=1}^N (\hat{w}_k^{*(i)})^2} = \frac{1}{\sum_{i=1}^N (\tilde{w}_k^{(i)})^2}, \quad (\text{A.24})$$

onde $\tilde{w}_k^{(i)}$ é o peso normalizado.

Várias técnicas para redução de degeneração são relacionadas na literatura [Frank (2003)]. Na seção seguinte será apresentada a técnica de reamostragem por importância de amostragem.

A.0.0.3 Reamostragem por Importância de Amostragem

A técnica mais popular de reamostragem é a amostragem multidimensional ou *Reamostragem por Importância de Amostragem* (SIR, *Sampling Importance Resampling*), também conhecida como *Bootstrap Bayesiano* [Frank (2003); Frank et al. (2003)]. A idéia fundamental desta técnica é obter partículas amostrando N vezes a distribuição $\hat{P}_N(d\mathbf{x}_{0:k}|\mathbf{z}_{1:k})$. Em geral, a técnica é dividida em duas etapas. Na primeira etapa gera-se valores de uma distribuição auxiliar conhecida. Na segunda etapa utiliza-se um mecanismo de correção para que os valores sejam representativos (ao menos aproximadamente) da distribuição a posteriori. Na prática costuma-se tomar a priori como distribuição auxiliar. O Algoritmo REAMOSTRAGEM SIR sintetiza a técnica de re-amostragem.

Algoritmo 8 REAMOSTRAGEM SIR

- 1: **para** $i = 1, \dots, N$ **faça**
 - 2: $\mathbf{x}_k^{(i)} \sim p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)})$
 - 3: $w_k^{(i)} = p(\mathbf{z}_k|\mathbf{x}_k^{(i)})$;
 - 4: $\tilde{w}_k^{(i)} = \frac{w_k^{(i)}}{\sum_{i=1}^N w_k^{(i)}}$;
 - 5: REAMOSTRA()
 - 6: **fim para**
-

Em geral, o processo de re-amostragem é executado por um método que “escolhe” as partículas que serão usadas na próxima etapa de propagação do algoritmo. Várias técnicas

são empregadas com esse propósito [Hong e Djuric' (2004)]. Porém, é comum a utilização do método de *re-amostragem sistemática* [Bolic et al. (2004)], descrito a seguir.

A técnica de *re-amostragem sistemática* é baseada em um teorema da teoria da amostragem que afirma que em uma representação de misturas de distribuições, o estimador de Monte Carlo mais eficiente de alguma função de interesse, baseado em uma amostra de tamanho N , é aquele que usa $N_i \propto \beta_i \sigma_i^2$ amostras de seus componentes [Frank (2003)], onde β_i é o componente ponderado da mistura e σ_i^2 é a variância da função de interesse sobre seu i -ésimo componente. Para se aplicar o teorema ao filtro de partículas, basta escolher $N_i \propto \tilde{w}_k^{(i)}$ [Bolic et al. (2004); Frank (2003)].

Um conjunto U de N pontos é gerado em um intervalo $[0, 1]$. Os pontos se distanciam entre si por N^{-1} . A soma parcial dos pesos dos pontos é calculada por meio de $q_i \triangleq \sum_{j=1}^i \tilde{w}_k^{(j)}$, $i = 1, \dots, N$. Obtém-se, assim, o número de filhos N_i em U que estejam entre q_{i-1} e q_i . Em outras palavras, a re-amostragem sistemática calcula a soma cumulativa q_i dos pesos e compara com um número uniforme U_i para $i = 1, \dots, N$ [Bolic et al. (2004); Hong e Djuric' (2004)].

Apêndice B

O Rastreamento Visual de Objetos

O rastreamento visual objetiva obter informações sobre o estado de alvos que se movimentam em ambiente de monitoramento a partir de características visuais e espaço-temporais que descrevem o comportamento e as propriedades do objeto de interesse durante um período de tempo uma seqüência de imagens [Trucco e Verri (1998); Forsyth e Ponce (2002)], eventualmente tratando problemas de oclusão e auto-occlusão dos alvos [Ito e Sakane (2001)].

A forma mais rudimentar de rastrear visualmente objetos é por meio de agrupamentos de pixels [Forsyth e Ponce (2002)]. Em síntese, esta técnica consiste na associação de *blobs* por meio do cruzamento de quadros anteriores ou por meio do deslocamento de centróides que mantém uma proximidade entre os quadros, definido em termos de um limiar.

Idealmente, cada trajetória corresponde ao movimento do alvo na cena, mesmo que este torne-se temporariamente parcial ou totalmente ocluso. Abordagens mais recentes utilizam técnicas estocásticas de estimação para o casamento de características ao longo de uma seqüência de imagens. Uma dessas técnicas mais aplicadas consiste na utilização do filtro de Kalman [Bradshaw et al. (1997)].

No caso do rastreamento visual, o filtro de Kalman é aplicado como um algoritmo que estima a posição e a incerteza de uma característica que se movimenta no próximo quadro de uma seqüência de imagens [Trucco e Verri (1998); Forsyth e Ponce (2002)].

Como pontos positivos do Filtro de Kalman destacam-se a robustez ao ruído de medição aos erros de modelagem além de ser de fácil implementação e pouco sensível a ajustes externos, sendo uma técnica extremamente versátil. Seus pontos negativos são a carga computacional e a necessidade de um modelo linear do sistema para a sua implementação [Abidie Gonzalez (1992); Trucco e Verri (1998); Forsyth e Ponce (2002); Russel e Norving (2004)].

O filtro de Kalman calcula a melhor estimativa da posição da característica no instante t_k e suas incertezas [Browne Hwang (1997)]. Entretanto, apesar de sua vasta abrangência, o Filtro de Kalman apresenta uma séria limitação quando considerado o problema do rastreamento de múltiplos objetos e o da não linearidade dos modelos de transição e observação.

Para se contornar o problema da não linearidade do modelo, emprega-se frequentemente o *Filtro de Kalman Estendido* (FKE) [Rao e Dhawas (1995); Siouris et al. (1997); Browne Hwang (1997)]. Em um FKE, os estados são representados por meio de variáveis aleatórias Gaussianas. A cada iteração se aplica uma expansão por séries de Taylor para linearizar as funções de predição e de observação em torno do estado corrente.

Em geral, o algoritmos de rastreamento procuram por características visuais extraídas dos objetos de interesse em uma seqüência de imagens. Tais características são definidas e obtidas no espaço das imagens ou por meio do conhecimento prévio de parâmetros contidos na cena, como iluminação ou as relações geométricas entre a câmera e objetos. Características visuais serão discutidas em mais detalhes na seção seguinte.

B.1 Características Visuais

A obtenção da informação visual é determinada por meio das *características visuais*. Nesse contexto, as características visuais são definidas como [Corke (1996)]:

Definição B.1 (Característica visual) – *As medidas ou observações obtidas por uma câmera são denominadas de características visuais e são descritas como globais quando se obtém uma propriedade global da imagem, como o valor médio dos pixels, momentos ou o fluxo óptico. Por outro lado, são descritas como locais quando dizem respeito a subconjunto particular de pixels, como pontos, quinas ou segmentos de reta.*



As características visuais são extraídas das imagens dos objetos tomadas a partir de diferentes posições de câmera, de forma robusta e não ambígua. Um *parâmetro de uma característica visual* [Espiau et al. (1992); Corke (1996)] caracteriza-se como um valor real obtido a partir de uma ou mais características visuais. O comprimento de um segmento de reta, as coordenadas de um ponto, padrões de textura, cor, o valor dos semi-eixos de uma elipse e o raio de uma circunferência são exemplos de parâmetros de características visuais.

Dado um conjunto de n parâmetros de características visuais, defini-se um vetor $\mathbf{c} = [f_1 \dots f_n]^T$ como sendo o *vetor de características visuais*, onde $\mathbf{c} \in \mathcal{F} \subseteq \mathbb{R}^n$ no qual \mathcal{F} representa o *espaço dos parâmetros de características visuais*. Em geral, os parâmetros da característica visual são obtidos a partir de uma *imagem de referência* [Espiau et al. (1992)].

Em seqüência, conceitos relativos ao rastreamento visual de características será abordado em maiores detalhes.

B.1.1 A Cor dos Alvos como Característica Visual

Algumas técnicas de rastreamento visual empregam a cor do objeto de interesse para a associação nos quadros subseqüentes [Bruce et al. (2000)]. Em geral, estes métodos operam reduzindo a resolução de cor na imagem e adotando uma métrica para a distância nesse espaço de cor reduzido.

Segundo a teoria do tri-estímulo da cor [Trucco e Verri (1998); Rafael Gonzalez (2001); Forsyth e Ponce (2002)], a visão colorida é resultante da excitação de três tipos de cones fotoreceptores da retina. O espaço de cor RGB baseia-se nesta teoria, relacionando a cor a um primeiro estágio na retina do processo de visão do ser humano. A codificação da percepção da cor no espaço RGB não é eficiente. A descrição de uma cor é mais intuitiva quando se usam as características da matiz e da saturação do espaço HSI.

O espaço de cores conhecido como HSI (*Hue, Saturation, Intensity*) é uma forma alternativa ao espaço RGB de representação de cores. As cores são definidas pelos atributos de intensidade ou brilho, matiz e saturação. Esses atributos podem ser analisados e manipulados individualmente, ao contrário do sistema RGB, onde eles são interligados [Rafael Gonzalez (2001)].

O espaço HSI é representado graficamente como um cone e usa coordenadas cilíndricas polares para representar as cores. O vértice do cone do HSI representa o preto, enquanto o seu eixo coincide com o eixo acromático. A intensidade aumenta em sentido contrário ao do vértice. Uma seção circular do cone mostra a variação de matizes ao redor do perímetro. A saturação aumenta para fora do centro, passando de cinza para tons pastéis e destes para matizes espectrais puras [Rafael Gonzalez (2001); Forsyth e Ponce (2002)].

A matiz informa a cor predominante em um determinado pixel da imagem. Pode ser compreendido como a medida do comprimento de onda médio da luz que um pixel reflete ou emite, definindo a cor do objeto. No modelo HSI, os valores da matiz estão representados por uma seqüência radial ao redor dos círculos de saturação e do eixo de intensidade, um ângulo entre 0° e 360° [Rafael Gonzalez (2001); Forsyth e Ponce (2002)].

A saturação tem como função estimar a pureza da cor. Quanto mais saturada, isto é, quanto maior o valor de saturação, mais pura será a cor. A pureza também está relacionada com a quantidade de cor branca presente na cor. O valor da saturação expressa o intervalo de comprimentos de onda ao redor do comprimento de onda médio no qual a energia é refletida ou transmitida [Rafael Gonzalez (2001); Forsyth e Ponce (2002)]. No modelo HSI a saturação é representada como distância radial do ponto até o eixo central do cone .

A intensidade fornece a noção geral de quão iluminada está a imagem naquele momento. É a medida da energia total envolvida em todos os comprimentos de onda responsáveis pela sensação de brilho dessa energia incidente sobre o olho [Rafael Gonzalez (2001); Forsyth e Ponce (2002)]. Seu valor é dado no modelo HSI pela distância de um ponto até a origem ou ápice do cone.

O *histograma de cor* descreve a distribuição das cores no objeto, independente de sua forma e orientação, definindo, assim, um modelo de cor do objeto [Stricker (1994); Rafael Gonzalez (2001)]. O histograma $\mathbf{h}(\mathbf{I}) \in \times$ é um vetor onde cada elemento representa o número de pixels na imagem \mathbf{I} que é descrito por uma determinada cor.

Para se quantificar a semelhança entre histogramas de cores, faz-se necessária a definição de uma métrica no espaço vetorial formado pela distribuição de cores representadas nos histogramas. Assim, para dois histogramas de cor \mathbf{h}_1 e \mathbf{h}_2 dados, as principais normas para medir semelhanças entre eles são [Stricker (1994); Rafael Gonzalez (2001)]:

$$d_{L1}(\mathbf{h}_1, \mathbf{h}_2) = \|\mathbf{h}_1 - \mathbf{h}_2\|_{L1} \quad (\text{B.1})$$

$$= \sum_{i=1}^n |h_{1,i} - h_{2,i}|, \quad (\text{B.2})$$

$$d_{L2}(\mathbf{h}_1, \mathbf{h}_2) = \|\mathbf{h}_1 - \mathbf{h}_2\|_{L2} \quad (\text{B.3})$$

$$= \left(\sum_{i=1}^n |h_{1,i} - h_{2,i}|^2 \right)^{\frac{1}{2}}, \quad (\text{B.4})$$

$$d_{L3}(\mathbf{h}_1, \mathbf{h}_2) = \frac{\sum_{i=1}^n |h_{1,i} - h_{2,i}|}{\sum_{i=1}^n h_{2,i}}. \quad (\text{B.5})$$

Em alguns casos as medidas de similaridade do tipo d_{L1} por histogramas funciona bem quando os histogramas são esparsos. As medidas do tipo d_{L2} provocam o aparecimento de muitos falsos negativos, ou seja, não recupera todas as imagens cujos histogramas são perceptualmente similares. O uso da métrica d_{L3} , ocasiona o aparecimento de muitos falsos positivos [Stricker (1994)].

O uso da cor definida no espaço HSI e as métricas acima são combinadas de modo

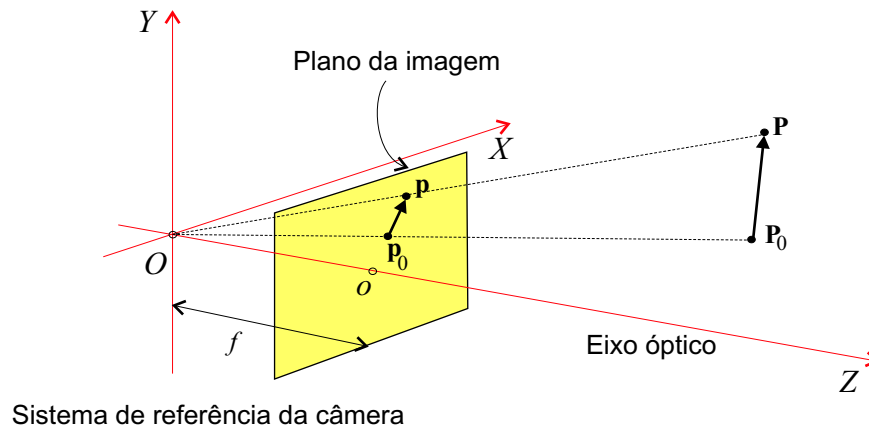


Figura B.1: Projeção do movimento de um ponto no plano de imagem.

a fornecerem diferentes formas de identificação da cor dos objetos. Na abordagem desenvolvida neste trabalho, um alvo tem suas características visuais representadas no espaço HSI e descritas por meio de um histograma de cor.

Um método usual para o rastreamento da cor [Novak e Shafer (1992)], consiste em armazenar o histograma em uma biblioteca que é posteriormente utilizada em um processo de retro-projeção. O objetivo é enfatizar grupos de pixels na cena que aparecem com mais frequência no modelo de cor do alvo, mas não tão frequentemente na cena. A imagem é segmentada por limiarização em grupos de pixels que indicam o alvo. Os objetos resultantes na imagem podem ser separados do fundo e, então, rastreados [Jacquot et al. (2005)].

B.2 Detecção do Movimento

A representação bidimensional do movimento de um objeto tridimensional é denominada de *campo de movimento* [Sonka et al. (1993); Trucco e Verri (1998); Forsyth e Ponce (2002)]. A cada ponto de um campo de movimento está associado um vetor velocidade, correspondendo a direção do movimento, velocidade e distância em uma seqüência de imagens. O campo de movimento pode ser definido como [Trucco e Verri (1998)]:

Definição B.2 (Campo de movimento) – *O campo de movimento é o campo vetorial de velocidade dos pontos da imagem, induzido pelo movimento relativo entre a câmera e a cena observada.*

■

A Figura B.1 mostra um sistema de referência da câmera. A origem desse sistema é o centro de projeção O . O eixo Z corresponde ao eixo óptico e f é a distância focal.

Seja \mathbf{P} um vetor do \mathbb{R}^3 definido por meio das coordenadas $\mathbf{P}[X, Y, Z]$ em relação ao sistema de referência da câmera. Seja \mathbf{p} a projeção de \mathbf{P} sob o plano de imagem, dada por

$$\mathbf{p} = f \frac{\mathbf{P}}{Z}. \quad (\text{B.6})$$

O movimento relativo entre \mathbf{P} e a câmera é descrito por meio de

$$\mathbf{V} = -\mathbf{T} - \boldsymbol{\omega} \times \mathbf{P}, \quad (\text{B.7})$$

onde \mathbf{T} é a componente de translação do movimento e $\boldsymbol{\omega}$ é a velocidade angular. Como o movimento é rígido, \mathbf{T} e $\boldsymbol{\omega}$ são os mesmos para qualquer \mathbf{P} [Espiau et al. (1992); Corke (1996)]. Separando a Equação B.7 em suas componentes, fica-se com

$$\begin{aligned} V_x &= -T_x - \omega_y Z + \omega_z Y, \\ V_y &= -T_y - \omega_z X + \omega_x Z, \\ V_z &= -T_z - \omega_x Y + \omega_y X. \end{aligned} \quad (\text{B.8})$$

Para se obter a relação entre a velocidade de \mathbf{P} e a velocidade correspondente a \mathbf{p} no plano da imagem, toma-se a derivada em relação ao tempo nos dois membros da Equação B.6. Assim, as *equações básicas do campo de movimento* são dadas por

$$\mathbf{v} = f \frac{Z\mathbf{V} - V_z\mathbf{P}}{Z^2}. \quad (\text{B.9})$$

Usando as Equações B.8 e B.9 a Equação B.9 é separada em suas componentes e reescrita da seguinte maneira:

$$\begin{aligned} v_x &= \frac{T_z x - T_x f}{Z} - \omega_y f + \omega_z y + \frac{\omega_x x y}{f} - \frac{\omega_y x^2}{f}, \\ v_y &= \frac{T_z y - T_y f}{Z} - \omega_x f + \omega_z x - \frac{\omega_y x y}{f} + \frac{\omega_x y^2}{f}. \end{aligned} \quad (\text{B.10})$$

O campo de movimento é a soma de duas componentes, uma que depende da translação e outra que depende da rotação. Em particular, as componentes de translação são dadas por

$$V_x = \frac{T_z x - T_x f}{Z}$$

e

$$V_y = \frac{T_z y - T_y f}{Z},$$

e as componentes de rotação, são

$$\Omega_x = -\omega_y f + \omega_z y + \frac{\omega_x x y}{f} - \frac{\omega_y x^2}{f}$$

$$\Omega_y = -\omega_x f + \omega_z x - \frac{\omega_y x y}{f} + \frac{\omega_x y^2}{f}$$

As equações a cima podem ser re-agrupadas e escritas na forma de um produto de matrizes da seguinte maneira [Espiau et al. (1992)]:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -\frac{1}{Z} & 0 & \frac{x}{z} & xy & -(1+x^2) & y \\ 0 & -\frac{1}{Z} & \frac{y}{z} & 1+y^2 & -xy & -x \end{bmatrix} \begin{bmatrix} v_x \\ v_y \\ v_z \\ \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}, \quad (\text{B.11})$$

onde $[\dot{x}\dot{y}]^T$ são as derivadas de \mathbf{p} em relação ao tempo, a matriz

$$\mathbf{J} = \begin{bmatrix} -\frac{1}{Z} & 0 & \frac{x}{z} & xy & -(1+x^2) & y \\ 0 & -\frac{1}{Z} & \frac{y}{z} & 1+y^2 & -xy & -x \end{bmatrix} \quad (\text{B.12})$$

é denominada de *Jacobiano da imagem* [Espiau et al. (1992); Corke (1996)] e os vetores

$$\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} \quad (\text{B.13})$$

e

$$\begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}, \quad (\text{B.14})$$

são respectivamente as velocidades linear e angular associadas ao deslocamento do vetor \mathbf{P} .

As equações a cima descrevem propriedades importantes sobre o campo de movimento, dentre as quais destacam-se:

- A parte do campo de movimento relacionada com a velocidade angular não necessita da informação de profundidade Z .
- Se for considerada somente a componente de translação com $T_z \neq 0$, o campo de movimento é radial e todos os vetores de movimento divergem (ou convergem) para um vetor \mathbf{p}_0 , com coordenadas dadas por

$$\begin{aligned} V_x &= -f \frac{T_x}{Z}, \\ V_y &= -f \frac{T_y}{Z}. \end{aligned}$$

Se $T_z = 0$, então o campo de movimento é paralelo.

- O comprimento dos vetores do campo de movimento é inversamente proporcional a profundidade Z ; se $T_z \neq 0$ o comprimento dos vetores também é inversamente proporcional a distância entre \mathbf{p} e \mathbf{p}_0 .
- As coordenadas do \mathbf{p}_0 definem o *ponto de fuga da direção de translação*.
- Se os vetores do campo de movimento estão em sentido contrário a \mathbf{p}_0 , então \mathbf{p}_0 é chamado de *foco de expansão*.
- Se os vetores do campo de movimento estão direcionados em sentido a \mathbf{p}_0 , então \mathbf{p}_0 é chamado de *foco de contração*.
- O movimento relativo a dois pontos coincidentes não dependem da componente de rotação.

As principais técnicas para a detecção do movimento serão vistas em maiores detalhes na seção seguinte.

B.2.0.1 Detecção do Campo de Movimento por Subtração de Fundo

Segundo Collins et al. (2000), as três principais abordagens para a detecção do campo de movimento, são:

- a diferenciação temporal,
- a subtração de fundo (ou subtração de referência) e a
- determinação do fluxo óptico.

A *diferenciação temporal* envolve a diferença entre dois quadros consecutivos e a criação de uma imagem binária. Os valores binários são obtidos na imagem diferença entre quadros consecutivos determinados a partir de um limiar pré-definido. Com a imagem binária criada, valores como a área total do movimento, coordenadas do centróide e o deslocamento das coordenadas do centróide podem ser calculados. Essas características podem ser obtidas segundo as equações abaixo, representadas no sistema de referência da imagem binária $B_{m \times n}$ em cada instante de tempo:

Área

$$A_t = \sum_{i,j} \mathbf{B}_t(i, j) \quad (\text{B.15})$$

Coordenadas do centróide

$$\bar{x}_t = \frac{1}{A_t} \sum_{i,j} x \mathbf{B}_t(i, j) \quad (\text{B.16})$$

$$\bar{y}_t = \frac{1}{A_t} \sum_{i,j} y \mathbf{B}_t(i, j) \quad (\text{B.17})$$

Deslocamento das coordenadas dos centróides

$$u_t = \bar{x}_t - \bar{x}_{t-1} \quad (\text{B.18})$$

$$v_t = \bar{y}_t - \bar{y}_{t-1} \quad (\text{B.19})$$

A diferenciação temporal é uma técnica aplicável em ambientes dinâmicos, porém a extração de características não é geralmente muito boa, o que pode levar a erros grosseiros na identificação de pontos do alvo.

A *subtração de referência* ou *subtração de fundo* é a forma de detecção mais simples e a mais utilizadas no rastreamento visual de objetos [Schalkoff (1989); Trucco e Verri (1998)]. Assume-se que o objeto de interesse esteja mais próximo da câmera do que os demais elementos presentes na cena. Uma câmera convenientemente posicionada no ambiente captura uma imagem \mathbf{B} da cena na qual não existe a presença de alvos; a imagem \mathbf{B} torna-se a imagem de referência, também denominada de *imagem de fundo* [Trucco e Verri (1998)].

Existem duas abordagens para a obtenção de B [Sonka et al. (1993)]: a primeira abordagem utiliza informações e estatísticas obtidas de diferentes imagens no decorrer do tempo. A outra abordagem utiliza os dois quadros mais recentes para montar a imagem diferença \mathbf{D} , como será explicado a seguir.

Qualquer mudança na cena pode ser detectada por meio de uma comparação entre a imagem de referência \mathbf{B} e qualquer nova imagem capturada. Supondo que uma nova imagem da cena \mathbf{S} contém um alvo de interesse, a imagem diferença $\mathbf{D} = |\mathbf{B} - \mathbf{S}|$ fornece o valor absoluto entre pixels correspondentes em \mathbf{B} e \mathbf{S} .

Cada pixel d_{ij} de \mathbf{D} maior um limiar τ , é modificado devido a presença do alvo na cena. Esses pixels são denominados de *pixels alvo* e são a saída da fase de detecção.

O limiar τ é escolhido de forma tal que compense uma relação entre o ruído e os eventuais pixels alvo da imagem, em outras palavras, valores de τ muito altos desprezam o ruído e valores muito baixos podem inviabilizar a detecção do alvo.

A imagem diferença pode ser calculada com base em características mais complexas, como a média da intensidade de brilho em uma vizinhança de pixels, características locais de textura, etc. Um pixel d_{ij} da imagem \mathbf{D} , terá valor 1 devido a uma das seguintes razões:

1. b_{ij} é um pixel sobre um objeto em movimento e s_{ij} é um pixel do fundo estático da cena (ou vice-versa).
2. b_{ij} é um pixel sobre um objeto em movimento e s_{ij} é um pixel sobre outro objeto em movimento.
3. b_{ij} é um pixel sobre um objeto em movimento e s_{ij} é um pixel sobre outra parte do mesmo objeto em movimento.
4. Ruído, perturbações no posicionamento da câmera estática, etc.

A maior vantagem dessa técnica é não ser necessário nenhum modelo geométrico ou fotométrico da cena, o que favorece a sua utilização em diferentes ambientes sem a necessidade de usar procedimentos complexos de calibração.

O cálculo do *fluxo óptico* pode ser usado para detectar objetos móveis independentes e de maneira mais robusta, porém seu cálculo é computacionalmente complexo, o que pode ser inviável em tempo real sem a presença de um hardware especializado.

A determinação analítica do fluxo óptico é afetada pela presença de ruído no brilho da imagem, que causa descontinuidades no campo de movimento, pela ocorrência de *oclusões* entre diferentes objetos em movimento, particularmente entre os alvos em movimento e o fundo estático, e pelo denominado *problema de abertura* que ocorre devido a impossibilidade de recuperar a direção do movimento através de uma abertura que é menor que o tamanho do alvo que apresenta características visuais (como textura ou arestas) insuficientes para perceber a componente transversal do seu movimento [Wu (2001)].

Cada método de cálculo do fluxo óptico é adequado para determinadas aplicações. Alguns algoritmos foram desenvolvidos para obter estimativas muito precisas da velocidade na imagem, às custas de um longo tempo de processamento; outros métodos encontram uma estimativa da velocidade com precisão e custo computacional baixo, permitindo a aplicação dos mesmos em sistemas de tempo real. Barron et al. (1994) fornece uma análise detalhada das diversas técnicas de determinação do fluxo óptico.

Uma vez detectado o movimento de um alvo, um sistema típico de rastreamento usando subtração de referência, por exemplo, irá procurar agrupar os pixels que descrevem o movimento. O objetivo do agrupamento é atribuir um rótulo para cada pixel que seja um potencial pixel do alvo.

Assim, pixels que pertençam ao mesmo alvo físico deverão possuir o mesmo rótulo. A forma mais simples de agrupamento é conhecida como *conexão de componentes* na qual a cada vizinhança de pixels adjacentes que compartilham as mesmas características é atribuído um mesmo rótulo [Sonka et al. (1993)].

As técnicas de conexão de componentes mais comuns são as denominadas de *conectividade-quatro* (*four-connectedness*) e *conectividade-oito* (*eight-connectedness*) [Rafael Gonzalez (2001)].

Na conectividade-quatro cada pixel em particular é considerado adjacente somente aos pixels que estejam diretamente a cima, a baixo, a esquerda e a direita deste compartilhem a mesma característica em comum.

Na conectividade-oito todos os pixels em volta de um pixel em particular que compartilham as arestas e quinas desse pixel possuem as mesmas características em comum são considerados adjacentes.

O agrupamento de pixels gera regiões sem forma definida denominadas de "*blobs*". Porém, os *blobs* podem ocorrer em agrupamentos de pixels que não pertençam necessariamente ao alvo. Assim, alguns testes podem ser realizados objetivando garantir a que regiões

da imagem correspondam, de fato, aos alvos. Os procedimentos típicos variam desde a simples remoção de alvos que não possuam um número mínimo de pixels até técnicas mais sofisticadas que detectam *blobs* formados pela reflexão, sombra ou outras mudanças na iluminação [Sonka et al. (1993); Trucco e Verri (1998)].

Apêndice C

Aspectos Geométricos do Posicionamento de Sensores

A questão fundamental no problema de posicionamento de sensores é saber como posicionar as câmeras de tal forma que se obtenham as melhores observações sobre os alvos. Este problema é essencialmente geométrico e alguns conceitos importantes serão apresentados em seguida.

Um *polígono* P é uma seqüência ordenada de pontos $p_1, p_2, \dots, p_n \in \mathbb{R}^2$, $n \geq 3$, chamados de *vértices* de P , junto a conjunto de segmentos de reta que ligam p_i a p_{i+1} , $i = 1, \dots, n - 1$ e p_n a p_1 , chamadas de *arestas* de P .

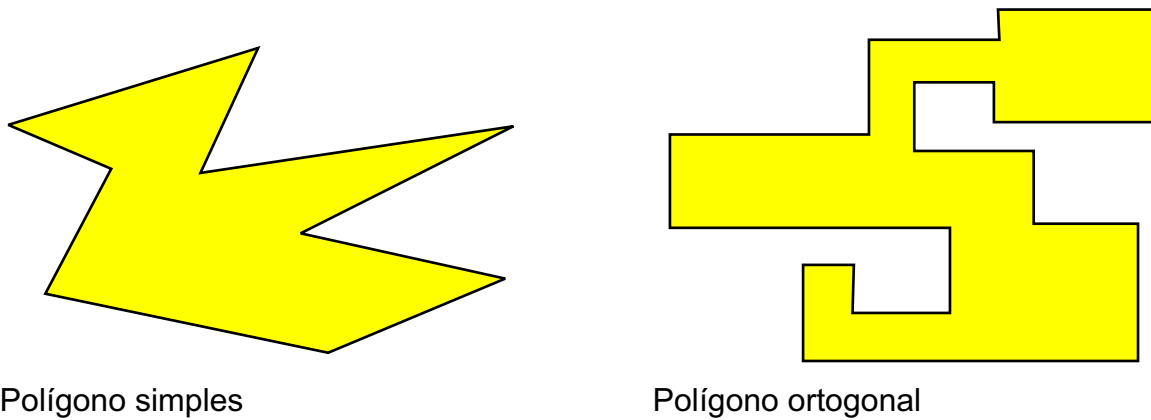
O polígono P é dito *simples* se quaisquer duas arestas não consecutivas estas não se interceptam em nenhum ponto. Um polígono simples divide o plano em duas regiões, uma não contornada chamada de *região exterior* e uma contornada, o interior. Todo polígono simples¹ é dito *ortogonal* se todas suas arestas sejam paralelas aos seus eixos x ou y . Exemplos de polígonos estão mostrados na Figura C.1.

Dados dois pontos p e q de um polígono P , diz-se que p é *visível* por q se o segmento de reta que une p a q está totalmente contido em P .

Uma coleção de pontos H de P *ilumina* ou *guarda*² P se todo ponto u de P é visível por um ponto p de H . O termo *ilumina* significa que cada elemento de H é uma fonte de luz que emite luz em todas as direções. A região iluminada por um ponto p define o *polígono de visibilidade* do ponto p (ver Figura C.2).

¹para simplificar a notação, todo polígono simples será referenciado somente por polígono daqui para frente.

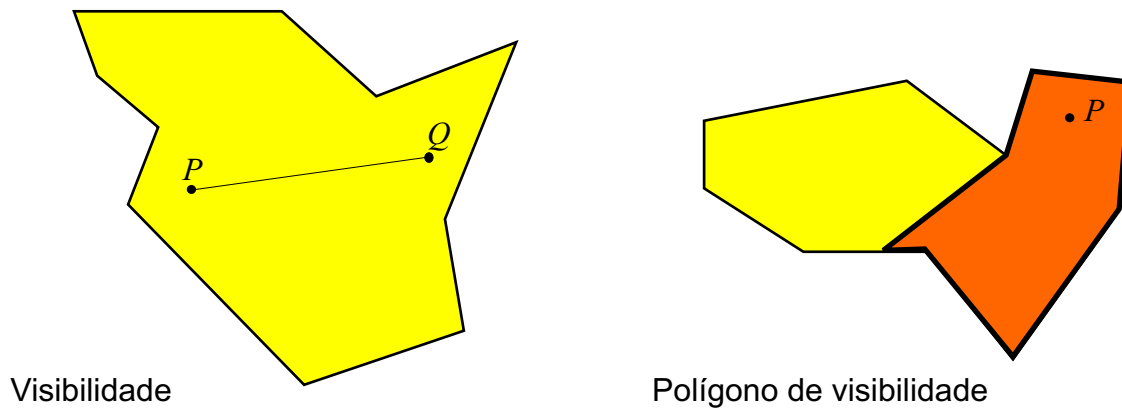
²Os termos guardar ou iluminar são equivalentes e usados de acordo com o resultado ou prova obtida pelos autores.



Polígono simples

Polígono ortogonal

Figura C.1: Exemplo de um polígono simples (à esquerda da figura) e um polígono ortogonal (à direita da figura).



Visibilidade

Polígono de visibilidade

Figura C.2: Polígono de visibilidade de um ponto.

Uma *triangulação* T de um polígono P , é a partição de P em um conjunto de triângulos com pares interiores disjuntos de tal maneira que as arestas destes triângulos sejam arestas ou diagonais de P que juntam pares de vértices. Uma triangulação de um polígono P possui exatamente $n - 2$ triângulos [Preparata e Shamos (1985); de Berg et al. (1985)].

Um *grafo* $G(V, E)$ consiste em um conjunto de elementos V chamados de vértices de $G(V, E)$, junto a um conjunto E denominado de arestas de $G(V, E)$. Dois vértices u e v de $G(V, E)$ são ditos *adjacentes* se o par $\{u, v\}$ é um elemento de E .

Um *caminho* de $G(V, E)$ é uma seqüência de vértices distintos v_1, \dots, v_k tal que v_i e v_{i+1} , $i = 1 \dots, k - 1$, $k \geq 2$, são adjacentes em $G(V, E)$.

Um *grafo de visibilidade* $GV(P)$ de um polígono P é um grafo no qual seu conjunto de vértices é o conjunto de vértices de P e dois vértices u e v são adjacentes em $VG(P)$

se eles são visíveis em P . A Figura C.3 ilustra um polígono e seu respectivo grafo de visibilidade.

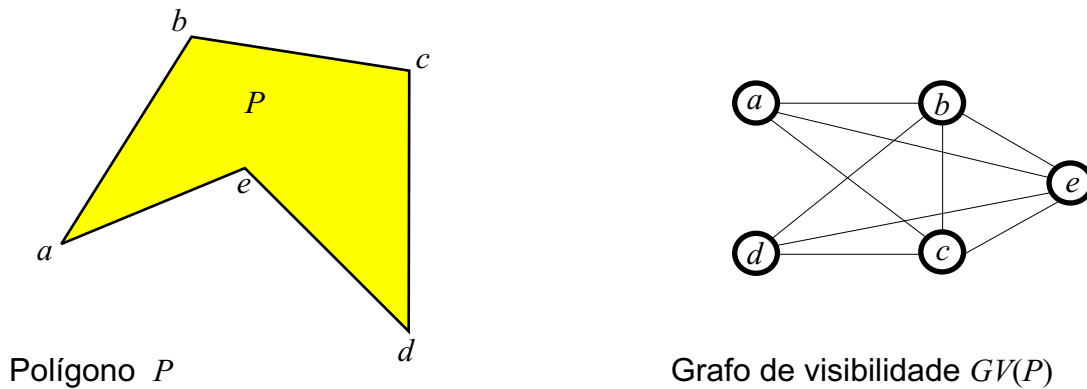


Figura C.3: Um polígono simples e seu grafo de visibilidade.

A abordagem geométrica para o problema de planejamento de sensores está diretamente relacionada com o problema denominado de "O Problema da Galeria de Arte" (PGA). Este problema foi introduzido em 1975 no trabalho de Chavatal (1975). Na sua formulação inicial, o PGA caracterizava-se como uma resposta a um problema geométrico de determinar o número mínimo de guardas suficientes para guardar uma galeria de arte de forma poligonal com n paredes. Nesta contextualização, o problema pode ser compreendido como um problema de posicionamento de sensores e está relacionado com o problema de decomposição poligonal, como será mostrado em seguida por meio do Teorema C.1.

Teorema C.1 *O número suficiente e ocasionalmente necessário de guardas para vigiar uma galeria de arte com n vértices é $\lfloor \frac{n}{3} \rfloor$ guardas.*

Prova. Considere um polígono qualquer P com n vértices. Uma triangulação T de P é obtida acrescentando-se $n - 2$ diagonais no interior de P . Colore-se os vértices de P usando três cores $\{1, 2, 3\}$, de tal maneira que dois vértices que sejam unidos por uma aresta de P ou por uma diagonal de T recebam cores diferentes, particionando o conjunto de pontos em três classes cromáticas C_1, C_2 e C_3 . Cada classe cromática guarda o interior de P . O resultado é obtido pelo número de guardas da menor classe cromática. ■

Para ver que $\lfloor \frac{n}{3} \rfloor$, considera-se o polígono pente P_p mostrado na Figura C.5 com $n = 3m$ vértices. É fácil verificar que para guardar P_p necessita-se de pelo menos m guardas.

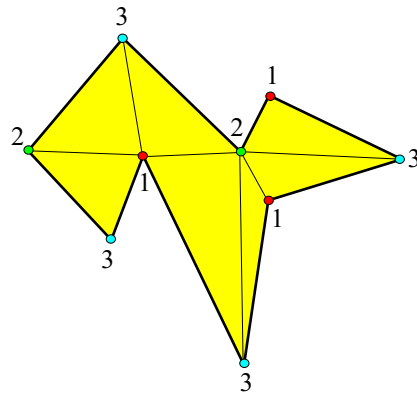


Figura C.4: A triangulação e a colorização dos vértices de um polígono.

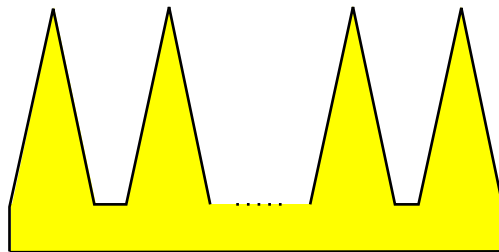


Figura C.5: Em um polígono pente com $n = 3m$ vértices se verifica a condição de necessidade do número de guardas.

Diferentes variações do problema da galeria de arte incluem polígonos com buracos, guardas que podem ver em cone ou com visão omnidirecional, galerias de arte tridimensionais e guardas com capacidade de mobilidade [Yamashita et al. (1997); ORourke (1983, 1987)]. Porém, duas variações do PGA apresentam importantes conseqüências para o trabalho aqui desenvolvido. São, respectivamente, o PGA com guardas móveis e guardas de aresta e o PGA com guardas cooperativos.

Um *guarda de aresta* é um guarda que tem a habilidade de se movimentar pelas arestas de um polígono [ORourke (1987)]. Um ponto q pode ser considerado como sob vigilância se for visível de alguma posição no caminho de um guarda. Um *guarda móvel* consegue se mover ao longo de um caminho fechado contido em um polígono P .

ORourke (1987) apresenta uma prova de que $\lfloor \frac{n}{4} \rfloor$ são sempre suficientes e ocasionalmente necessários para guardar qualquer polígono com n vértices.

Seja Q um conjunto de pontos contidos em um polígono P . O grafo de visibilidade de Q , $GV(Q)$, é o grafo no qual o conjunto de vértices é formado pelos elementos de Q que são mutuamente visíveis. O conjunto de guardas é chamado de *cooperativo* se eles guardam todo P e $GV(Q)$ é conexo.

O principal resultado para o PGA com guardas cooperativos, mostrado ainda em ORourke (1987), é que o número mínimo de guardas cooperativos para um polígono simples é um problema NP-difícil.

Referências Bibliográficas

- A. Doucet, S. G.e Andrieu, C. (2000). On sequential monte carlo sampling methods for bayesian filtering. *Statistics and Computing*, 10(3):197–208.
- Abidi, M. A.e Gonzalez, R. C. (1992). *Data Fusion in Robotics and Machine Intelligence*. Academic Press Inc.
- Bajcsy, R. (1988). Active perception. In IEEE, editor, *IEEE Proceedings*, volume 76, páginas 966–1005.
- Balch, T.e Parker, L. E. (2001). *Robot Teams*. A.K. Peters, 1a edição.
- Bar-Shalom, Y. (1978). Tracking methods in a multitarget environment. *IEEE Transactions on Automatic Control*, 23(4):618–626.
- Bar-Shalom, Y. (1990). *Multitarget-Multisensor Tracking: Advanced Applications*, volume I. Artech House.
- Bar-Shalom, Y. (1992). *Multitarget-Multisensor Tracking: Applications and Advances*, volume III. Artech House.
- Bar-Shalom, Y., Li, X. R.,e Kirubarajan, T. (2001). *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons.
- Barron, J., Fleet, D., Beauchemin, S.,e Burkitt, T. (1994). Performance of optical flow techniques. *Computer Vision and Pattern Recognition Journal*, 92:236–242.
- Berriss, W., Price, W.,e Bober, M. (2003). Real-time visual analysis and search algorithms for intelligent video surveillance. In *International Conference on Visual Information Engineering*, páginas 226–229.
- Bjorling-Sachs, I. (1998). Edge guards in rectilinear polygons. *Computational Geometry*, 11(2):111–123.

- Black, J. (2004). Applications of video surveillance technology in airports. In *38th Annual 2004 International Carnahan Conference on Security Technology*, páginas 42–46.
- Black, J., Ellis, T., e Makris, D. (2003). Wide area surveillance with a multi camera network. *Intelligent Distributed Surveillance Systems*, páginas 21–25.
- Blackman, S. (2004). Multiple hypothesis tracking for multiple target tracking. *IEEE Aerospace and Electronic Systems Magazine*, 19(1):5–18.
- Bolic, M., Djuric, P. M., e Hong, S. (2004). Resampling algorithms for particle filters: A computational complexity perspective. *EURASIP Journal of Applied Signal Processing*, 15(1):2267–2277.
- Bose, P., Shermer, T., Toussaint, G., e Zhu, B. (1997). Guarding polyhedral terrains. *Computational Geometry*, 7(3):173–185.
- Boult, T. (2003). Geo-spatial active visual surveillance on wireless networks. In *Applied Imagery Pattern Recognition Workshop*, páginas 244–249.
- Boult, T., Micheals, R., Gao, X., e Eckmann, M. (2001). Into the woods: Visual surveillance of noncooperative and camouflaged targets in complex outdoor settings. In *Proceedings of the IEEE*, páginas 1382–1401. IEEE – Institute of Electrical and Electronics Engineers.
- Bowden, R. e Kaew-Tra-Kul-Pong, P. (2005). Towards automated wide area visual surveillance: tracking objects between spatially-separated, uncalibrated views. In *IEE Proceedings on Vision and Image Signal Processing*, volume 152, páginas 555–566.
- Bradshaw, K., Reid, I., e Murray, D. (1997). The active recovery of 3d motion trajectories and their use in prediction. In *IEEE Transaction on Pattern Analysis and Machine Intelligence*, volume 19, páginas 219–234.
- Brown, R. G. e Hwang, P. Y. C. (1997). *Introduction to Random Signals and Applied Kalman Filtering with Matlab Exercises and Solutions*. John Wiley & Sons, 3a. edição.
- Bruce, J., Balch, T., e Veloso, M. (2000). Fast and cheap color image segmentation for interactive robots. In *Workshop on Interactive Robotics and Entertainment*, volume 2, páginas 122–133.
- Chavatal, V. (1975). A combinatorial theorem in plane geometry. *Journal of Computorial Theory*, B(18):39–41.

- Chen, H.-T., Lin, H.-H., e Liu, T.-L. (2001). Multi-object tracking using dynamical graph matching. In *Computer Vision and Pattern Recognition (CVPR'01)*, volume 2, páginas 210–217.
- Chen, S. Y. e Li, Y. (2004). Automatic sensor placement for model-based robot vision. *IEEE Transaction on Systems, Man, and Cybernetics—Part B: Cybernetics*, 34(1):393–408.
- Chen, W.-P., Hou, J., e Sha, L. (2004). Dynamic clustering for acoustic target tracking in wireless sensor networks. *IEEE Transactions on Mobile Computing*, 3(3):258–271.
- Chiu, P. e Lin, F. (2004). A simulated annealing algorithm to support the sensor placement for target location. In *Canadian Conference on Electrical and Computer Engineering*, volume 2, páginas 867–870.
- Collins, R., Lipton, A., e Kanade, T. (1999). A system for video surveillance and monitoring. In *Eighth International Topical Meeting on Robotics and Remote Systems*, volume 1, páginas 282–308. ANS – American Nuclear Society.
- Collins, R., Lipton, A., Kanade, T., Fujiyoshi, H., Duggins, D., Tsin, Y., Tolliver, D., Enomoto, N., e Hasegawa, O. (2000). A system for video surveillance and monitoring. Technical Report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University.
- Corke, P. I. (1996). *Visual Control of Robots: high-performance visual servoing*. Research Studies Press Ttd, 1a edição.
- Cormick, J. M. e Isard, M. (2001). Bramble: A bayesian multiple-blob tracker. In *8th International Conference on Computer Vision (ICCV 2001)*, páginas 34–41.
- Cucchiara, R., Grana, C., Prati, A., e Vezzani, R. (2005). Computer vision system for in-house video surveillance. In *IEE Proceedings Publication on Vision, Image and Signal Processing*, volume 152, páginas 242–249.
- Daeipour, E., Blair, W., e Bar-Shalom, Y. (1997). Bias compensation and tracking with monopulse radars in the presence of multi path. *IEEE Transactions on Aerospace and Electronic Systems*, 33(3):863–882.
- de Berg, M., van Kreveld, M., Overmars, M., e Schwarzkopf, O. (1985). *Computational Geometry: Algorithms and Applications*. Springer-Verlag, 2a. edição.

- Dietl, M., Gutmann, J.-S., e Nebel, B. (2001). Cooperative sensing in dynamic environments. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'01)*, volume 1, páginas 88–97.
- Dockstader, S. L. e Tekalp, A. M. (2001). Multiple camera fusion for multi-object tracking. In *IEEE Workshop on Multi-Object Tracking*, páginas 95–102.
- Dudek, G. e Jenkin, M. (2000). *Computational Principles of Mobile Robotics*. Cambridge University Press.
- Elfes, A. (1990). Occupancy grids: A stochastic spatial representation for active robot perception. In *Sixth Conference on Uncertainty in AI*, páginas 60–70.
- Ellis, T. (2002). Multi-camera video surveillance. In *36th Annual 2002 International Carnahan Conference on*, páginas 228–233.
- Espiiau, B., Chaumette, F., e Rives, P. (1992). A new approach to visual servoing. *IEEE Transactions on Robotics and Automation*, 8(3):313–326.
- Foresti, G. L. e Micheloni, C. (2003). A robust feature tracker for active surveillance of outdoor scenes. *Electronic Letters on Computer Vision and Image Analysis*, 1(1):21–34.
- Forsyth, D. A. e Ponce, J. (2002). *Computer Vision: A Modern Approach*. Prentice Hall, 1a. edição.
- Frank, O. (2003). *Multiple Target Tracking*. Tese de doutorado, Swiss Federal Institute of Technology, Department of Electrical Engineering, Swiss Federal Institute of Technology Zurich (ETH), Switzerland.
- Frank, O., J. Nieto, J. G., e Scheduling, S. (2003). Multiple target tracking using sequential monte carlo methods and statistical data association. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 3, páginas 2718–2723.
- Gasser, G., Bird, N., Masoud, O., e Papanikolopoulos, N. (2004). Human activities monitoring at bus stops. In *Proceedings of the IEEE International Conference on Robotics Automation*, volume 1, páginas 90–96.
- Gordon, N., Salmond, D., e Smith, A. (1993). A novel approach to nonlinear/non-gaussian baesian state stimation. volume 140, páginas 107–133.

- Guibas, L. J., Latombe, J.-C., LaValle, S. M., Lin, D., e Motwani, R. (1997). Visibility-based pursuit-evasion in a polygonal environment. *Lecture Notes in Computer Science*, páginas 17–30.
- Guibas, L. J., Latombe, J.-C., LaValle, S. M., Lin, D., e Motwani, R. (1999). A visibility-based pursuit-evasion problem. *International Journal of Computational Geometry and Applications*, 9(5):471–494.
- Gutman, P. O. e Velger, M. (1990). Tracking targets using adaptive kalman filtering. *IEEE Transactions on Aerospace and Electronic Systems*, 26(5):691–699.
- Haag, M., Frank, T., Kollnig, H., e Nagel, H.-H. (2001). Influence of an explicitly modelled 3d scene on the tracking of partially occluded vehicles. *Computer Vision and Image Understanding*, 65(2):206–225.
- Han, K. e Veloso, M. (1997). Physical model based multi-objects tracking and prediction in robosoccer. In *Working Notes of the AAAI 1997 Fall Symposium on Model-directed Autonomous Systems*, volume 1, páginas 34–42.
- Harville, M. (2002). Stereo person tracking with adaptive plan-view statistical templates. In *ECCV Workshop on Statistical Methods in Video Processing*,, páginas 62–72.
- Hlin, P. (2001). An addition to art galleries with interior walls. *Discrete and Computational Geometry*, 25(2):311–334.
- Hong, S. e Djuric', M. B. P. M. (2004). An efficient fixed-point implementation of residual resampling scheme for high-speed particle filters. *IEEE Signal Processing Letters*, 11(5):482–485.
- Hsieh, T. T. (2004). Using sensor networks for highway and traffic applications. *IEEE Potentials*, 23(2):13–16.
- Huang, T. e Russel, S. (1997). Object identification in a bayesian context. volume 1, páginas 1276–1282.
- Hue, C.; Le Cadre, J.-P. P. P. (2002). Sequential monte carlo methods for multiple target tracking and data fusion. *IEEE Transactions on Signal Processing*, 50(2):309–325.
- Isard, M. e Blake, A. (1998). Condensation - conditional density propagation for visual tracking. *International Journal of Computer Vision*, 1(29):5–28.

- Ishiguro, H. (1997). Distributed vision system: A perceptual information infrastructure for robot navigation. In *IJCAI 97 – International Joint Conference on Artificial Inteligence*, páginas 36–43.
- Ito, K. e Sakane, S. (2001). Robust view-based visual tracking with detection of occlusions. In *Proceedings of the IEEE International Conference on Robotics Automation*, volume 2, páginas 1207–1213.
- Jacquot, A., Sturm, P., e Ruch, O. (2005). Adaptative tracking of non rigid objects based on color histograms and automatic parameter selection. In *Proceedings of IEEE Workshop on Motion and Video Computing, Breckenridge, Colorado*, páginas 103–109.
- James M. Ferryman, Stephen J. Maybank, A. D. W. (2000). Visual surveillance for moving vehicles. *International Journal of Computer Vision*, 37(2):187–197.
- Javed, O., Khan, S., Rasheed, Z., e Shah, M. (2000). Camera handoff: Tracking in multiple uncalibrated stationary cameras. In *IEEE Workshop on Human Motion*, páginas 113–118.
- Javed, O., Khan, S., Rasheed, Z., e Shah, M. (2003). Tracking across multiple cameras with disjoint views. In *9th IEEE International Conference on Computer Vision*, volume 2, páginas 952–958.
- Jennings, C. (2002). Robust finger tracking with multiple cameras. In *Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real-Time Systems*, volume 1, páginas 100–107.
- Júnior, B. M. e de Oliveira Anido, R. (2004). Distributed real-time soccer tracking. In *2nd international workshop on Video surveillance & sensor networks*, volume 1, páginas 65–71.
- Jung, B. e Sukhatme, G. S. (2002a). A region-based approach for cooperative multi-target tracking in a structured environment. In *2002 IEEE/RSJ International Conference on Intelligent Robots and Systems*, páginas 2764–2769.
- Jung, B. e Sukhatme, G. S. (2002b). Tracking targets using multiple robots: The effect of environment occlusion. *Autonomous Robots Journal*, 13(3):191–205.
- Jung, B. e Sukhatme, G. S. (2004). A generalized region-based approach for multi-target tracking in outdoor environments. In *Proceedings of the IEEE International Conference on Robotics Automation*, volume 1, páginas 205–210.

- Jung, Y.-K. e Ho, Y.-S. (2001). A feature-based vehicle tracking system in congested traffic video sequences. *Lecture Notes in Computer Science*, 2195(1):190–192.
- Kang, J., Cohen, I., e Medioni, G. (2003). Soccer player tracking across uncalibrated camera streams. In *5th IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance VS-PETS03*, volume 1, páginas 85–95.
- Kemmotsu, K. e Kanade, T. (1994). Sensor placement design for object pose determination with three light-stripe range finders. In *Proceedings of the IEEE International Conference on Robotics Automation*, volume 2, páginas 1357–1364.
- Kettner, V. e Zabih, R. (1999). Bayesian multi-camera surveillance. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, páginas 250–259.
- Kogut, G. e Trivedi, M. (2001). Maintaining the identity of multiple vehicles as they travel through a video network. In *IEEE Workshop on Multi-Object Tracking*, páginas 29–34.
- Kotecha, J., Ramachandran, V., e Sayeed, A. (2005). Distributed multitarget classification in wireless sensor networks. *IEEE Journal on Selected Areas in Communications*, 23(4):703–713.
- Lane, D. M., Chantler, M. J., e Dai, D. (1998). Robust tracking of multiples objects in sector-scan sonar image sequences using optical flow motion estimation. *IEEE Journal of Oceanic Engeneering*, 23(1):31–46.
- LaValle, S. M., Gonzales-Banos, H. H., Becker, C., e Latombe, J. (1997). Motion strategies for maintaining visibility of a moving target. In *Proceedings of the IEEE International Conference on Robotics Automation*, páginas 731–736.
- LaValle, S. M. e Hutchinson, S. A. (1998). Optimal motion planning for multiple robots having independent goals. *IEEE Transactions on Robotics and Automation*, 14(6):912–925.
- LaValle, S. M., Simov, B., e Slutzki, G. (2002). An algorithm for searching a polygonal region with a flashlight. *International Journal of Computational Geometry and Applications*, 12(1-2):87–113.
- Leonard, J. J., Durrant-Whyte, e Hugh, F. (1991). Mobile robot localization by tracking geometric beacon. *IEEE Transactions on Robotics and Automation*, 7(3):376–383.

- Liaw, B.-C., Huang, N. F., e Lee, R. C. T. (1993). The minimum cooperative guards problem on k-spiral polygons. In *Proceedings of the Fifth Canadian Conference on Computational Geometry*, páginas 97–102.
- Lin, F. e Chiu, P. (2005). A near-optimal sensor placement algorithm to achieve complete coverage-discrimination in sensor networks. *IEEE Communications Letters*, 9(1):43–45.
- Lou, J., Yang, H., Hu, W., e Tan, T. (2002). usual vehicle tracking using an improved ekf. In *The 5th Asian Conference on Computer Vision (ACCV2002)*, páginas 23–25.
- MacCormick, J. e Blake, A. (1999). A probabilistic exclusion principle for tracking multiple objects. In *Proceedings of the IEEE Seventh International Conference on Computer Vision*, volume 5, páginas 646–654.
- Makris, D., Ellis, T., e Black., J. (2004). Bridging the gaps between cameras. In *CVPR*, volume 2, páginas II–205– II–210.
- Matsuyama, T. (1998). Cooperative distributed vision – dynamic integration of visual perception, action and communication. In *Proceedings of Image Understanding Workshop*, páginas 365–384.
- McKenna, S. J., Jabri, S., Duric, Z., Rosenfeld, A., e Wechsler, H. (2000). Tracking groups of people. *Computer Vision and Image Understanding*, 80(1):42–53.
- McKenna, T. (2003). Video surveillance and human activity recognition for anti-terrorism and force protection. In *IEEE Conference on Advanced Video and Signal Based Surveillance*, páginas xi+376.
- Meier, E. B. e Ade, F. (1999). Using the condensation algorithm to implement tracking for mobile robots. In *Third European Workshop on Advanced Mobile Robots (Eurobot99)*, páginas 73–80.
- Menegatti, E. e Pagello, E. (2001). Cooperative distributed vision for mobile robots. In *Proc. Workshop on AI, Vision and Pattern Recognition - 7th Conference of Italian Association for Artificial Intelligence*, páginas em CD-ROM.
- Menegatti, E., Scarpa, A., Massarin, D., Ros, E., e Pagello, E. (2003). Omnidirectional distributed vision system for a team of heterogeneous robots. In *Proc. of IEEE Workshop on Omnidirectional Vision (Omnivis'03), in Computer Vision and Pattern Recognition (CVPR 2003)*, volume 1, páginas 432–438.

- Michael, T. S. e Pinciu, V. (2003). Art gallery theorems for guarded guards. *Computational Geometry*, 26(3):247–258.
- Michelsoni, C., Foresti, G., e Snidaro, L. (2005). A network of co-operative cameras for visual surveillance. In *IEE Proceedings on Vision, Image and Signal Processing*, páginas 205–212.
- Murray, D. e Basu, A. (1994). Motion tracking with an active camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(5):449–459.
- Navarro-Serment, L., Dolan, J., e Khosla, P. (2004). Optimal sensor placement for cooperative distributed vision. In *ICRA*, volume 1, páginas 939–944.
- Novak, C. L. e Shafer, S. (1992). Estimating scene properties from color histograms. Technical Report CMU-CS-92-212, Computer Science Department, Carnegie Mellon University, Pittsburgh, PA.
- ORourke, J. (1983). Galleries need fewer mobile guards: A variation on chvatal’s theorem. *Geometriae Dedicata*, 14(3):273–283.
- ORourke, J. (1987). *Art Gallery Theorems and Algorithms*. Oxford University Press.
- Park, S.-T. e Lee, J. G. (1998). Design of a practical tracking algorithm with radar measurements. *IEEE Transactions on Aerospace and Electronic Systems*, 34(4):1337–1344.
- Parker, L. E. (1996). On the design of behavior-based multi-robot teams. *Advanced Robotics*, 10(6):547–578.
- Parker, L. E. (2002). Distributed algorithms for multi-robot observation of multiple moving targets. *Autonomous Robots*, 12(3):231–255.
- Patrick Pérez, J. V. e Blake, A. (2004). Fusion for visual tracking with particles. In *Proceedings of IEEE*, volume 3, páginas 95–105.
- Penny, D. (1999). Multi-sensor management for passive target tracking in an anti-submarine warfare scenario. In *IEE Colloquium on Target Tracking: Algorithms and Applications*, volume 3, páginas 1–5.
- Plankers, R. e Fua, P. (2001). Tracking and modeling people in video sequences. *Computer Vision and Image Understanding*, 81(3):285–302.

- Preparata, F. e Shamos, M. (1985). *Computational Geometry: An Introduction*. Springer-Verlag, 1a edição.
- Rafael Gonzalez, R. W. (2001). *Digital Image Processing*. Prentice Hall, 2a edição.
- Rao, K. D. e Dhawas, J. (1995). Parallel implementation of radar tracking extended kalman filters on transputer networks. *IEEE Transactions on Aerospace and Electronic Systems*, 31(2):857–862.
- Russel, S. e Norving, P. (2004). *Inteligência Artificial*. Elsevier, 2a edição.
- Sachs, S., Rajko, S., e LaValle, S. M. (2004). Visibility-based pursuit-evasion in an unknown planar environment. *The International Journal of Robotics Research*, 23(1):3–26.
- Saito, H., Baba, S., Kimura, M., Vedula, S., e Kanade, T. (1999). Appearance-based virtual view generation of temporally-varying events from multi-camera images in the 3d room. In *Second International Conference on 3-D Digital Imaging and Modeling*, páginas 516–525.
- Schalkoff, R. J. (1989). *Digital Image Processing and Computer Vision*. John Wiley & Sons, 1a. edição.
- Schulz, D., Burgard, W., Fox, D., e Cremers, A. B. (2001). Tracking multiple moving targets with a mobile robot using particle filters and statistical data association. *International Journal of Computational Geometry and Applications*, 5:120–135.
- Siebel, N. T. (2003). *Design and Implementation of People Tracking Algorithms for Visual Surveillance Applications*. Tese de doutorado, Department of Computer Science, The University of Reading.
- Siebel, N. T. e Maybank, S. (2002). Fusion of multiple tracking algorithms for robust people tracking. In *7th European Conference on Computer Vision (ECCV 2002)*, volume IV, páginas 373–387.
- Simov, B., LaValle, S. M., e Slutzki, G. (2002). A complete pursuit-evasion algorithm for two pursuers using beam detection. In *Proceedings of the IEEE International Conference on Robotics Automation*, páginas 618–623.
- Siouris, G., Chen, G., e Wang, J. (1997). Tracking an incoming ballistic missile using an extended interval kalman filter. *IEEE Transactions on Aerospace and Electronic Systems*, 33(1):232–240.

- SIVAM, P. (Consulta realizada em 24.05.2005). Sistema de vigilância da amazônia. Disponível na WWW em <http://www.sivam.gov.br/>.
- Sogo, T., Ishiguro, H., e Ishida, T. (1999). Mobile robot navigation by distributed vision agent. In *PRIMA*, páginas 96–110.
- Sogo, T., Ishiguro, H., e Ishida, T. (2001). Mobile robot navigation by distributed vision system. *New Generation Computing*, 19(2):121–137.
- Song, T., Ahn, J. Y., e Um, T. Y. (1990). A passive tracking filter for missile capture. *IEEE Transactions on Aerospace and Electronic Systems*, 26(5):867–875.
- Sonka, M., Hlavac, V., e Boyle, R. (1993). *Image Processing and Machine Vision*. Chapman & Hall Computing, 1a. edição.
- Spletzer, J. R. e Taylor, C. J. (2003). Dynamic sensor planning and control for optimally tracking targets. *International Journal of Robotics Research*, 22(1):7–20.
- Ssu, K.-F., Ou, C.-H., e Jiau, H. (2005). Localization with mobile anchor points in wireless sensor networks. *IEEE Transactions on Vehicular Technology*, 54(3):1187–1197.
- Stricker, M. A. (1994). Bounds for the discrimination power of color indexing techniques. In *SPIE Proceedings – Storage and Retrieval for Image and Video Databases II*, volume 2185, páginas 15–24.
- Stroupe, A. e Balch, T. (2003). Value-based observation with robot teams (vbort) for dynamic targets. In *Proceedings of the International Conference on Advanced Robotics*, volume 1, páginas 55–65.
- Sukhatme, G. S., Montgomery, F., e Vaughan, R. T. (2001). *Robot Teams: From Diversity to Polymorphism*, chapter Experiments with Cooperative Aerial-Ground Robots, páginas 345–368.
- Suzuki, I. e Yamashita, M. (1992). Search for a mobile intruder in a polygonal region. *SIAM Journal of Computing*, 21(5):863–888.
- Tao, H., Sawhney, H. S., e Kumar, R. (1999). A sampling algorithm for tracking multiple objects. In *Workshop on Vision Algorithms*, páginas 53–68.
- Tarabanis, K., Allen, P., e Tsai, R. (1995). A survey of sensor planning in computer vision. *IEEE Transactions on Robotics and Automation*, 11(1):86–104.

- Thucker, N., A.J.Lacely, e Courtney, P. (2002). An empirical desing methodology for construction of vision systems. Technical Report 2002-005, University of Manchester, Medical School.
- Timothy W. McLain, R. W. B. e Kelsey, J. M. (2002). Experimental demonstration of multiple robot cooperative target intercept. In *Guindance, Navigation, and Control Conference and Exhibit*, volume 2, páginas 140–149.
- Toyama, K. e Blake, A. (2001). Probabilistic tracking in metric space. In *8th. Internatinal Conference on Computer Vision - ICCV 2001*, volume 2, páginas 50–57.
- Trucco, E. e Verri, A. (1998). *Introductory Techniques for 3D Computer Vision*. Prentice Hall.
- Tóth, C. D. (2000). Art gallery problem with guards whose range of vision is 180o. *Computational Geometry*, 17(3-4):121–134.
- Tóth, C. D. (2002a). Art galleries with guards of uniform range of vision. *Computational Geometry*, 21(3):185–192.
- Tóth, C. D. (2002b). Illumination in the presence of opaque line segments in the plane. *Computational Geometry*, 21(3):193–204.
- Tóth, C. D. (2003). Guarding disjoint triangles and claws in the plane. *Computational Geometry*, 25(1-2):51–65.
- Ukita, N. e Matsuyama, T. (2005). Real-time cooperative multi-target tracking by communicating active vision agents. *Computer Vision and Image Understanding*, 97(2):137–179.
- Vercauteren, T., Guo, D., e Wang, X. (2005). Joint multiple target tracking and classification in collaborative sensor networks. *IEEE Journal on Selected Areas in Communications*, 23(4):714–723.
- Vidal, R. e Sastry, S. (2002). Vision-based detection of autonomous vehicles for pursuit-evasion games. In *IFAC 15th. Triennial World Congress*, volume 1, páginas 30–45.
- Wong, T., Tsuchiya, T., e Kikuno, T. (2004). A self-organizing technique for sensor placement in wireless micro-sensor networks. In *18th International Conference on Advanced Information Networking and Applications*, volume 1, páginas 78–83.

- Woo, D. e Capson, D. (2000). 3d visual tracking using a network of low-cost pan/tilt cameras. In *Canadian Conference on Electrical and Computer Engineering*, volume 2, páginas 884–889.
- Wu, Y. (2001). Optical flow and motion analysis. Technical report, Northwestern University.
- Wu, Y. e Huang, T. S. (2004). Robust visual tracking by integrating multiple cues based on co-inference learning. *International Journal on Computer Vision*, 58(1):55–71.
- Y. Aloimonos, I. W. e Bandopadhyay, A. (1988). Active vision. *International Journal of Computer Vision*, 2:333–356.
- Yamashita, M., Umemoto, H., Suzuki, I., e Kameda, T. (1997). Searching for mobile intruders in a polygonal region by a group of mobile searchers (extended abstract). In *Symposium on Computational Geometry*, páginas 448–450.
- Yeddanapudi, M., Bar-Shalom, Y., Pattipati, K., e Deb, S. (1995). Ballistic missile track initiation from satellite observations. *IEEE Transactions on Aerospace and Electronic Systems*, 31(3):1054–1071.
- Zia Khan, T. B. e Dellaert, F. (1999). Efficient particle filter-based tracking of multiple interacting targets using an mrf-based motion model. Technical report, College of Computing, Georgia Institute of Technology Atlanta, Georgia.