UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA
DEPARTAMENTO DE ENGENHARIA DA COMPUTAÇÃO E AUTOMAÇÃO INDUSTRIAL

UM MÉTODO RECURSIVO APLICADO AO PROBLEMA DE LOCALIZAÇÃO EM VISÃO DE MÁQUINA

AUTOR: ANTONIO MARIA GARCIA TOMMASELLI
ORIENTADOR: CLÉSIO LUÍS TOZZI

Tese apresentada à Faculdade de Engenharia Elétrica da Universidade Estadual de Campinas, UNICAMP, como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Elétrica.

JUNHO, 1993

UNICAMP
BIBLIOTECA CENTRAL

Termo de aprovação

Este exemplar corresponde à redação final da tese defendida por Antonio Maria Garcia Tommaselli e aprovada pela Comissão Julgadora em 18 06 93.

Orientador

à *Giselle*

AGRADECIMENTOS

O autor deseja externar seus agradecimentos às pessoas e entidades abaixo relacionadas:

. ao Prof. Dr. Clésio Luís Tozzi pela orientação, pelas proveitosas discussões e pelo apoio no decorrer do trabalho;
. aos professores Lotufo, Wagner e Jino pela leitura do trabalho e pelas sugestões;
. aos colegas Castanho, Olga, Júlio e Berutti, pela amizade e pelo apoio no desenvolvimento do trabalho;
. aos demais professores, funcionários e alunos do DCA, que de uma forma ou de outra contribuíram para que este trabalho fosse concluído;
. à Universidade Estadual Paulista, UNESP, campus de Presidente Prudente, pela concessão de afastamento integral durante o período de realização dos créditos e de desenvolvimento da tese;
. aos companheiros do Departamento de Cartografia, da UNESP, campus de Presidente Prudente, que tornaram possível este trabalho;
. à Coordenadoria de Aperfeiçoamento do Pessoal de Ensino Superior, CAPES, pela concessão de bolsa;

SUMÁRIO

	Página
TÍTULO	i
TERMO DE APROVAÇÃO	ii
DEDICATÓRIA	iii
AGRADECIMENTOS	iv
SUMÁRIO	v
LISTA DE TABELAS	ix
LISTA DE FIGURAS	x
RESUMO	xii
ABSTRACT	xiii

CAPÍTULO I

1.	INTRODUÇÃO	1
1.1	CONSIDERAÇÕES PRELIMINARES	1
1.2	O PROBLEMA DE LOCALIZAÇÃO	3
1.3	OBJETIVOS DO TRABALHO	6
1.4	ESTRUTURA E CONTEÚDO DO TRABALHO	7

CAPÍTULO II

2.	TÉCNICAS EXISTENTES	8
2.1	CALIBRAÇÃO E LOCALIZAÇÃO	9
2.1.1	Método clássico de calibração	11
2.1.2	Outros métodos de localização e calibração usando pontos de apoio	17
2.1.3	Localização e calibração usando linhas e estruturas	18
2.1.4	Localização usando o modelo de Liu e Huang	22

2.2 RECONHECIMENTO 23

2.3 FILTRAGEM 26
2.3.1 Métodos usando Filtragem 26
2.3.2 Filtragem Kalman 28
2.3.3 Filtro em Batch 29
2.3.4 Filtro Kalman Padrão 30
2.3.5 Filtro Kalman Estendido Iterativo 32
2.3.6 Estatística de resíduos e divergência 34

2.4 RELAÇÕES ENTRE REFERENCIAIS 35

2.5 EXTRAÇÃO DE FEIÇÕES RETAS NA IMAGEM 38
2.5.1 Detecção de bordas 38
2.5.2 Vetorização 40

CAPÍTULO III

3. MODELO MATEMÁTICO E PROCESSO RECURSIVO 45

3.1 MODELO MATEMÁTICO 46

3.1.1 Relações entre as representações normal e paramétrica da reta no plano 46
3.1.2 Plano de interpretação 48
3.1.2.1 Plano de interpretação descrito no espaço imagem 49
3.1.2.2 Plano de interpretação descrito no espaço objeto 52
3.1.3 Modelo Funcional 52

3.2 FILTRAGEM KALMAN APLICADA AO MODELO SIMULTÂNEO 56
3.5.1 Sequência de procedimentos para o cálculo do IEKF 61

3.3 PROCESSO RECURSIVO ESTIMAÇÃO-EXTRAÇÃO DE FEIÇÕES ... 62
3.3.1 Conceito e estrutura do processo recursivo 62
3.3.2 Reconhecimento e estimativa predita 64

3.3.3 Modelo do objeto e transformação 65
3.3.4 Determinação de linhas visíveis e escolha de uma reta 65
3.3.5 Definição do polígono de busca 66
3.3.6 Geração das "scan-lines" 69
3.3.7 Busca na imagem - vetorização 72
3.3.7.1 Refinamento das coordenadas de máquina .. 72
3.3.7.2 Ajuste da reta pelo M.M.Q 73
3.3.7.3 Efeito dos erros nos parâmetros intrínsecos na avaliação da precisão da reta na imagem 75
3.3.8 Filtragem 78

3.4 CONCLUSÃO 78

CAPÍTULO IV

4. ANÁLISE EXPERIMENTAL DO MODELO 79

4.1 INTRODUÇÃO 79

4.2 ANÁLISE COM DADOS SIMULADOS 79
4.2.1 Metodologia da simulação 79
4.2.2 Características dos experimentos 82
4.2.3 Análise de um caso típico 83
4.2.4 Comparação dos resultados obtidos com o IEKF com os resultados do filtro em batch 86
4.2.5 Efeito da variação da distância câmara-objeto .. 87
4.2.6 Efeito da variação da distância focal da câmara . 89
4.2.7 Configurações deficientes 90
4.2.8 Variação da posição do objeto na imagem 91
4.2.9 Efeito da magnitude dos erros observacionais .. 93
4.2.10 Efeito da avaliação incorreta do erro observacional 95
4.2.11 Efeito da variação na estimativa predita inicial 98
4.2.12 Efeito de erros grosseiros nas observações ... 101
4.2.13 Efeito de erros sistemáticos não eliminados ... 102
4.2.14 Resultados para uma sequência de imagens 104

4.2.15 Sumário das conclusões obtidas com o estudo com dados simulados 108

4.3 RESULTADOS COM DADOS REAIS 109
4.3.1 Características dos experimentos 109
4.3.2 Resultados com o cubo estático 111
4.3.3 Resultados com o cubo em movimento 116
4.3.4 Avaliação da repetibilidade dos experimentos com uma única imagem 118
4.3.5 Resultados com uma sequência de imagens 122
4.3.6 Sumário das conclusões obtidas com experimentos com dados reais 125

4.4 REQUISITOS DE PRECISÃO EM APLICAÇÕES DE VISÃO DE MÁQUINA 126

CAPÍTULO V

5. CONCLUSÕES 128
5.1 SÍNTESE DO TRABALHO DESENVOLVIDO 128
5.2 CONCLUSÕES 129
5.3 RECOMENDAÇÕES 131

APÊNDICE 1

A1.1 TRANSFORMAÇÕES 133
A1.1.1 Translações 133
A1.1.2 Rotações 133
A1.2.3 Escala 134

A.1.2 EQUAÇÕES PROJETIVAS 135

5. REFERÊNCIAS BIBLIOGRÁFICAS 137

6. BIBLIOGRAFIA 142

LISTA DE TABELAS

Tabela		Página
4.2.1	Características dos experimentos	83
4.2.2	Parâmetros da reta e respectivas variâncias	84
4.2.3	Resultados obtidos com o IEKF	84
4.2.4	Comparação entre o filtro em batch e o IEKF	86
4.2.5	Características dos experimentos com a distância focal	89
4.2.6	Desvios padrão dos erros observacionais nos vértices das retas ($\sigma_x = \sigma_y$)	93
4.2.7	Desvios padrão dos erros observacionais nos vértices das retas ($\sigma_x \neq \sigma_y$)	94
4.2.8	Estimativas preditas para os 7 experimentos ...	98
4.3.1	Parâmetros intrínsecos da câmara CCD e seus respectivos desvios padrão estimados, obtidos com o método clássico, utilizando 6 imagens convergentes	109
4.3.2	Resultados com o IEKF utilizando a imagem 4.3.2..	112
4.3.3	Tempos de processamento, em segundos, para cada etapa do processo recursivo e área da janela de busca, em nº de pixels, para cada reta	114
4.3.4	Resultados com o cubo em movimento	117
4.3.5	Valores medidos na imagem (parâmetros das retas).	119
4.3.6	Valores preditos e filtrados, obtidos com o IEKF, para um cubo a 1,12 m da câmara	120
4.3.7	Resultados do processo recursivo para uma sequência de imagens	123
4.3.8	Diferenças entre os vetores estado para imagens vizinhas	124
4.4.1	Características de algumas aplicações envolvendo localização	126

LISTA DE FIGURAS

Figura		Página
1.1	Tempos de processamento em um processo de localização	4
1.2	Conceito de um processo recursivo, onde a etapa de estimação de parâmetros realimenta a etapa de extração de feições	5
2.1.1	Fluxo de procedimentos para o cálculo dos parâmetros de calibração pelo método clássico ..	16
2.2.1	Etapas no reconhecimento de objetos	24
2.3.1	Sequência da Filtragem Kalman padrão	31
2.3.2	Diagrama para o Filtro Kalman padrão com o particionamento do vetor das observações	32
2.3.3	Diagrama do Filtro Kalman Estendido Iterativo ..	34
2.4.1	Um problema típico em Visão Robótica e os referenciais utilizados	36
2.5.1	Operadores de Sobel	38
2.5.2	Transformada de Hough	41
2.5.3	Passos no método de agrupamento θ-ρ	43
3.1.1	Elementos das representações normal e paramétrica da reta	47
3.1.2	Plano de interpretação	49
3.1.3	Os vetores normais aos planos de interpretação no espaço imagem e espaço objeto	51
3.3.1	Fluxo do procedimento recursivo entre filtragem e extração de feições	63
3.3.2	Sequência de cálculos dos vértices do polígono de busca	68
3.3.3	Elementos do polígono de busca	69
3.3.4	"Scan lines" para o polígono de busca	70
3.3.5	Sequência para a geração de "scan-lines"	71

4.2.1	Gráficos correspondentes à convergência do IEKF .	85
4.2.2	Efeito da variação de distância nos experimentos com o IEKF	88
4.2.3	Efeito da variação da focal nos experimentos com o IEKF	89
4.2.4	Resultados obtidos com diferentes configurações .	91
4.2.5	Variação da posição do cubo na imagem	92
4.2.6	Efeito da magnitude dos erros observacionais ..	94
4.2.7	Efeito da magnitude dos erros observacionais para $\sigma_x \neq \sigma_y$	95
4.2.8	Efeito da estimação da variância das observações	96
4.2.9	Resíduos preditos padronizados para os 4 experimentos	97
4.2.10	Efeitos das estimativas preditas	99
4.2.11	Resíduos preditos padronizados	100
4.2.12	Efeito de erros grosseiros nas observações	101
4.2.13	Efeito de erros sistemáticos não eliminados ...	103
4.2.14	Sequência de imagens com o cubo a 88 cm da câmara.	104
4.2.15	Resultados para uma sequência de 4 imagens	105
4.2.16	Sequência de imagens com o cubo a 1,10 m da câmara	106
4.2.17	Resultados para uma sequência de 6 imagens	106
4.2.18	Resultados para uma sequência de 6 imagens, com erro de transição de estado menor	107
4.3.1	Referenciais usados nos experimentos reais	110
4.3.2	Imagem de bordas do cubo a 1,13 m	112
4.3.3	Resultados para a imagem 4.3.2	113
4.3.4	Redução progressiva da janela de busca	115
4.3.5	Imagem de bordas com o efeito de arrastamento causado pelo movimento do objeto	116
4.3.6	Convergência da estimativa para o cubo em movimento	118
4.3.7	Imagem de bordas do cubo a 1,13 m da câmara	119
4.3.8	Resultados para 3 imagens, tomadas na mesma posição	121
4.3.9	Sequência de imagens de bordas simulando o cubo em movimento	123
A1.1.1	Projeção perspectiva	136

RESUMO

Este trabalho procura apresentar contribuições à resolução dos problemas de localização da câmara e de objetos, utilizando como apoio linhas retas conhecidas no espaço objeto e cujas homólogas tenham sido extraídas na imagem.

Um modelo matemático do tipo explícito, e que permita a determinação simultânea dos seis parâmetros de orientação, é desenvolvido, sendo, então, tratado por Filtragem Kalman. A natureza recursiva do Filtro, possibilita a definição de um processo iterativo, em que as estimativas sequenciais dos parâmetros realimentam a etapa de extração de feições. Isto leva a uma gradativa redução do espaço de busca das feições na imagem, e à conseqüente diminuição do custo computacional.

Pressupõe-se a disponibilidade de estimativas preditas de boa qualidade, em consonância com o conceito de Visão de Verificação. Conclui-se que, dentro deste contexto, a introdução de métodos de estimação de parâmetros, além de emprestar confiabilidade ao processo, permite a redução do custo computacional.

Os procedimentos propostos foram implementados e testados com dados simulados e com dados reais. Os resultados obtidos são discutidos e analisados e as conclusões resultantes são apresentadas.

ABSTRACT

The aim of this work is to present contributions to the solution of the camera location and object location problems using known straight lines in the object space and extracted in the image.

An explicit and simultaneous mathematical model using straight lines is derived, aiming Kalman Filtering application. The recursive nature of the Filter can be used to define an iterative process in which the sequential estimates feedback the feature extraction step. This leads to a gradual reduction of the features search space and, thus, computation time also decreases.

The availability of good quality predicted estimates is assumed, according to the concept of Verification Vision. Besides the increase in reliability, estimation methods could also reduce the computational costs.

The proposed approach was tested with simulated and real data. The obtained results are discussed and conclusions are presented.

CAPÍTULO I

1. INTRODUÇÃO

1.1 CONSIDERAÇÕES PRELIMINARES

Um sistema de Visão de Máquina é constituido, basicamente, por uma câmara, um dispositivo de digitalização de imagens e de software para processamento destas imagens. A estrutura dos Sistemas de Visão sofre uma série de restrições, principalmente com relação aos ambientes em que operam e no tempo de resposta exigido. As primeiras teorias para implementação de sistemas de Visão Computacional, procuravam dividir o problema em módulos que tivessem alguma correlação com o sistema de Visão Humana (Marr, 1982). A complexidade destes sistemas inviabilizou sua aplicação a problemas práticos, com os recursos computacionais atuais.

A disciplina de Visão de Máquina diferencia-se da chamada Visão Computacional por dedicar-se à resolução de problemas práticos, sem adotar paradigmas gerais. Aloimonos (1988) ressalta o declínio nos esforços para a construção de sistemas de Visão completos, resultando em sistemas que executam um número limitado de tarefas em um ambiente restrito, mas com aplicações industriais reais.

Existem inúmeras aplicações de Visão Máquina, notadamente em automação da manufatura (Tommaselli e Tozzi, 1990). Tarefas como inspeção e medição de partes para controle de qualidade usando Visão já são usuais nas plantas industriais mais modernas, embora a maioria destes sistemas ainda sejam bidimensionais. Uma das aplicações importantes no contexto industrial é a Robótica. A informação Visual torna os robôs mais versáteis e flexíveis, aumentando sua produtividade. O sensor Visual permite localizar partes ou

objetos sem tocá-los e com alta precisão e confiabilidade, monitorar a posição do robô, verificar se estão operando adequadamente, prever e evitar colisões e detectar acidentes. *Confiabilidade, flexibilidade e resposta em tempo real* são os requisitos básicos para um sistema sensorial para Robótica. Outra aplicação bastante explorada é em veículos auto-dirigidos (AGV), onde o sistema de Visão pode ser usado para detectar obstáculos, para seguir uma trilha ou, ainda, como um sistema de posicionamento em relação ao mundo real.

Contudo, aplicações como Robótica e AGVs envolvem o processamento de informações do espaço tridimensional. Em Robótica, especificamente, existem dois problemas importantes e que serão tratados neste trabalho: a localização da câmara e a localização de objetos a partir de imagens.

Existem várias configurações possíveis para o sistema robô-câmara. A câmara pode ser fixada em uma posição adequada no espaço de trabalho, de modo a imagear a garra do robô e os objetos a serem manipulados. Neste caso, a posição da câmara é calculada apenas uma vez e o problema é a localização dinâmica da garra do robô e dos objetos. Uma das desvantagens desta configuração é a falta de flexibilidade, uma vez que a câmara deve ser posicionada de modo a imagear simultâneamente todo o espaço de trabalho possível do manipulador. Além disto, o rastreamento dos movimentos e a localização da garra do robô em tempo real são problemas bastante complexos. Esta configuração é denominada **"eye-off-hand"**.

Outra possibilidade é fixar a câmara à garra do manipulador, de maneira a "imagear" a área onde a garra atuará. Este tipo de configuração é conhecido como **"eye-in-hand"** e tem como vantagem a sua flexibilidade. É importante notar que a câmara estará se movimentando em conjunto com a garra, o que permite utilizar a Visão como forma de verificar a posição e orientação da garra, exigindo a localização da câmara cada vez que esta for movimentada. Além disto, é possível localizar os objetos a serem manipulados diretamente em relação à garra.

Existem vários métodos que permitem determinar a localização de um objeto no espaço a partir de observações visuais. O método de Estereo-Visão baseia-se na reconstrução da posição de pontos no espaço objeto a partir da intersecção de retas geradas a partir pontos homólogos em duas (ou mais) imagens, tomadas de pontos de vista distintos. Contudo, o método de Estereo-Visão apresenta algumas restrições devido ao custo computacional na etapa de determinação de correspondência entre pontos homólogos em imagens distintas ("stereo-matching").

O método mais frequentemente utilizado para a localização de objetos,

requer o uso de apenas uma câmara, pressupondo-se o conhecimento de alguns elementos da cena ("model-based Vision"). Este método monocular tem como grande vantagem a possibilidade de obtenção de resultados em tempo real pois o processamento se restringe a uma única imagem. Este método é objeto de estudo neste trabalho.

1.2 O PROBLEMA DE LOCALIZAÇÃO

Em problemas que envolvam aplicações tridimensionais, usualmente, a câmara deve ser localizada em relação ao sistema de coordenadas global, o que implica em encontrar uma transformação deste espaço para o espaço imagem. (Apêndice 1).

No problema de localização de objetos busca-se encontrar a transformação espaço imagem-objeto. Quando o referencial global coincidir com o referencial do objeto, esta transformação é a inversa da transformação definida no problema de localização da câmara. A maioria dos métodos de localização da câmara e de localização de objetos são semelhantes, podendo-se adotar uma abordagem única para os dois problemas.

Neste trabalho, será considerado o espaço objeto como referencial absoluto, buscando-se calcular a transformação do espaço objeto para o espaço imagem.

Nas abordagens tradicionais, esta transformação é calculada usando pontos com coordenadas no espaço objeto conhecidas, e cujos homólogos na imagem tenham sido medidos. Entretanto, localizar estes pontos na imagem é problemático, devido, não somente ao custo computacional envolvido, mas também à possibilidade de enganos. Estes enganos podem ser a seleção de pontos que não correspondam ao ponto pretendido, ou mesmo, à extração de um ponto não existente. Propostas mais recentes usam feições contínuas, como retas ou círculos e mesmo estruturas mais complexas, como formas e objetos poliédricos, para calcular a transformação objeto-imagem. Enquanto os pontos na imagem são apenas agrupamentos de pixels vizinhos na imagem, estas feições são elementos mais estruturados, ligados por algum tipo de relação funcional (por exemplo, os pixels que compõem uma linha reta na imagem)

As linhas retas, especificamente, possuem algumas vantagens importantes sobre os pontos: são fáceis de detectar em uma imagem, existem em abundância em ambientes industriais e podem ser extraídas da imagem com precisão subpixel, isto é, com erro menor que a dimensão do pixel (Liu e Huang, 1991).

Este problema de extração de feições em uma imagem é bastante complexo, quer pelo tempo de processamento, quer pela dificuldade em caracterizar adequadamente a correspondência entre uma feição no espaço objeto e sua homóloga no espaço imagem.

Na maioria dos métodos revisados, tanto usando pontos como outro tipo de feições (retas, linhas ou estruturas), os erros observacionais, resultantes do processo de extração das feições na imagem, não são tratados adequadamente. Em vários destes métodos o cálculo dos parâmetros de localização é feito em vários estágios, o que produz um efeito de propagação dos erros observacionais entre cada estágio, degradando a solução final.

Além destes problemas, na maioria dos trabalhos consultados, os autores consideram que a etapa de extração de feições na imagem e o estabelecimento de correspondência já tenha sido realizada. O custo computacional desta etapa, entretanto, é consideravelmente maior que o da etapa de cálculo dos parâmetros de localização. Lee et al (1990) relataram um experimento em que o tempo de processamento das etapas de extração de feições foi de cerca de 30 s, enquanto que o tempo de coleta da imagem foi de 1/30 s e o tempo de cálculo dos parâmetros de localização foi de apenas 1.07 ms. Isto significa o tempo de processamento para a extração de feições é de cerca de 28.000 vezes o tempo para o cálculo dos parâmetros. A figura 1.1 representa esquematicamente a relação entre estes tempos.

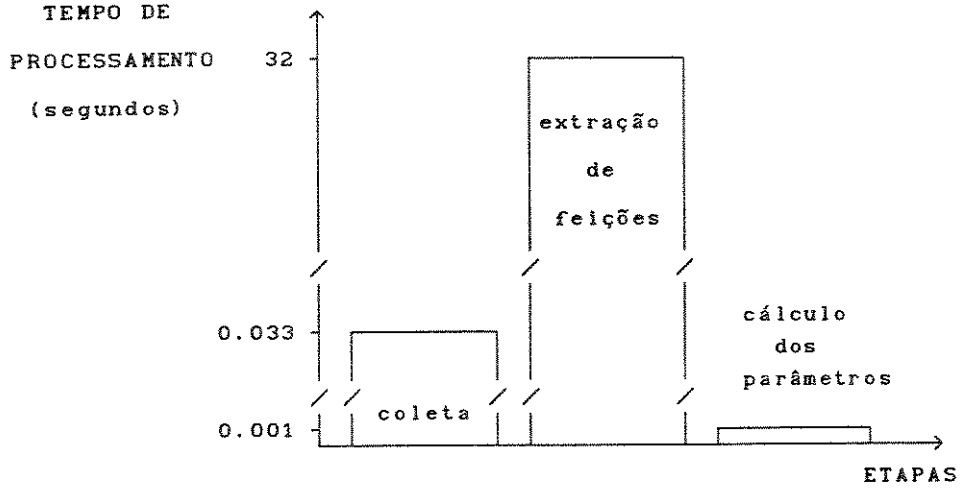


Figura 1.1 Tempos de processamento em um processo de localização.

A maioria dos autores propõe o uso de processamento paralelo para reduzir o custo computacional da etapa de extração de feições.

Outra maneira de reduzir o custo computacional na etapa de extração de feições, é restringir o espaço de busca destas feições a uma área provável de presença na imagem, com base em um valor aproximado dos parâmetros de localização da câmara. Esta estratégia de redução do custo computacional na etapa de extração de feições, reduz o trabalho de Visão a um processo de verificação da posição das feições em uma pequena janela, e à introdução de correções no estado do sistema, com base nas discrepâncias encontradas entre o valor predito e o valor medido. Este tipo de abordagem para a Visão de Máquina têm sido denominado de Visão de Verificação (Feddema e Lee, 1990 e Feddema et al, 1991).

Este conceito pode ser ainda melhorado, se forem empregados métodos recursivos de estimação de parâmetros, como a Filtragem Kalman. Apesar de aumentar o tempo de processamento na etapa de cálculo da localização, é importante observar que a natureza recursiva da Filtragem Kalman pode ser aproveitada para definir uma forma de realimentar o processo de extração de feições na imagem, reduzindo gradativamente o custo computacional desta etapa. A figura 1.2 ilustra esta proposta.

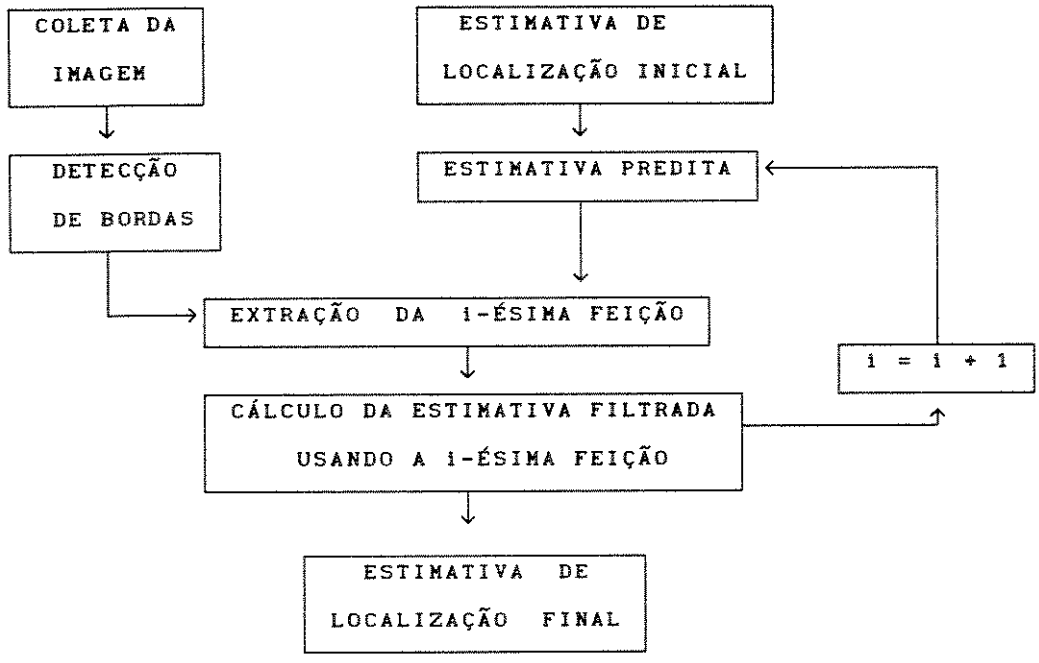


Figura 1.2 Conceito de um processo recursivo, onde a etapa de estimação de parâmetros realimenta a etapa de extração de feições.

Considerando os aspectos mencionados, pretende-se apresentar uma abordagem recursiva ao problema de localização da câmara, com o objetivo de sanar algumas das atuais restrições:

- os métodos atuais, em sua maioria, não tratam adequadamente os problemas de localização, no caso da presença de erros observacionais nas medidas extraídas da imagem;
- os resultados obtidos não são acompanhados de uma estimativa de precisão, não sendo possível avaliar se são adequados à aplicação proposta;
- a maioria dos modelos existentes para o tratamento de linhas retas calculam os parâmetros de localização em estágios distintos, o que causa propagação de erros entre estes estágios;
- o custo computacional da extração de feições na imagem é elevado, pois baseia-se na busca por toda a imagem ou, então, em uma região definida por uma única estimativa a priori. O custo computacional da extração de feições é muito superior ao custo computacional do cálculo dos parâmetros de localização;

1.3 OBJETIVOS DO TRABALHO

Considerando as restrições mencionadas, a proposta central deste trabalho pode ser definida como:

- definir um método de localização recursivo, baseado em filtragem, usando linhas retas, de tal maneira que as estimativas sequenciais possam realimentar o processo de extração de feições na imagem, reduzindo seu custo computacional;

Para tanto, podem ser definidos alguns objetivos do trabalho:

- desenvolver um modelo matemático usando linhas retas, que trate os seis parâmetros extrínsecos simultâneamente, em um único estágio;
- desenvolver um procedimento recursivo usando Filtragem Kalman, de modo a realimentar o processo de extração de feições com os resultados obtidos sequencialmente e, obter estimativas confiáveis. Este mesmo procedimento

deve permitir o processamento dinâmico de sequências de imagens;

- estudar o problema de localização usando o modelo desenvolvido com dados simulados, para determinar as expectativas de precisão, as configurações mais adequadas, e as situações potencialmente problemáticas;
- estudar o problema com dados reais, em situações parcialmente controladas, de modo a aferir a qualidade do método desenvolvido;

1.4 ESTRUTURA E CONTEÚDO DO TRABALHO

No primeiro capítulo foi estabelecido o problema a ser tratado, algumas necessidades existentes e os objetivos gerais do trabalho.

No segundo capítulo é revisada a base teórica relacionada aos temas necessários ao desenvolvimento do trabalho. Uma revisão dos métodos de calibração e localização é apresentada, com uma discussão dos problemas e restrições destes métodos. O problema de reconhecimento é tratado rapidamente, mostrando-se as principais características dos métodos baseados em modelos do objeto conhecidos. Ainda no segundo capítulo é discutido o problema de detecção de bordas e vetorização, além de serem mostradas as equações usadas na Filtragem Kalman.

No terceiro capítulo são apresentadas as principais contribuições deste trabalho. Inicialmente é desenvolvido o modelo matemático usando linhas retas. Em seguida é apresentada a sequência de procedimentos para o cálculo da Filtragem Kalman. Finalmente, é desenvolvido o conceito do processo recursivo, que permite a realimentação da etapa de extração de feições.

No quarto capítulo o modelo apresentado é submetido a testes com dados simulados, que permitem verificar seu funcionamento e comportamento ante várias situações e configurações. Os resultados são analisados e discutidos. São, então, apresentados alguns testes com dados reais, simulando um problema protótipo.

No quinto capítulo são apresentadas algumas conclusões e recomendações baseadas na experiência acumulada e nos problemas ainda não resolvidos e que demandam investigação adicional.

No Apêndice 1 são apresentadas as transformações empregadas no trabalho e as equações projetivas.

CAPÍTULO II

2. TÉCNICAS EXISTENTES

Este capítulo tem como objetivo apresentar uma revisão bibliográfica dos métodos de localização e calibração existentes, bem como das técnicas de filtragem e de extração de feições na imagem, que serão empregadas no processo recursivo, proposto neste trabalho.

Na **seção 2.1** são apresentados vários métodos de **localização e calibração**; na seção 2.1.1 é descrito o método clássico de calibração usando pontos de apoio, empregado para determinar os parâmetros intrínsecos e extrínsecos da câmara; na seção 2.1.2 são mencionados outros métodos de calibração e localização que também utilizam pontos de apoio; na seção 2.1.3 são revisados os métodos que se utilizam de retas, ou estruturas mais complexas, para o cálculo da localização. Entre estes, merece destaque o método de Liu e Huang (1988), apresentado na seçao 2.1.4.

Na **seção 2.2** é discutido o problema de **reconhecimento** de objetos tridimensionais, com uma abordagem na qual se admite que os modelos dos objetos a serem reconhecidos estão armazenados em uma base de dados ("model-based Vision").

Na **seção 2.3** é apresentada uma revisão dos métodos de **Filtragem Kalman**, que se constitui no método de tratamento do modelo estocástico a ser empregado neste trabalho.

Na **seção 2.4** é apresentado o conjunto das **relações entre referenciais** encontrados em uma aplicação protótipo, no caso, uma configuração "eye-in-hand" monitorando uma esteira transportadora de objetos.

Finalmente, na **seção 2.5** são descritos métodos de extração de feições em uma imagem digital; estes métodos permitem extrair de uma imagem em tons de cinza as feições necessárias para a solução do problema de localização.

2.1 CALIBRAÇÃO E LOCALIZAÇÃO

Localizar um objeto em relação a um sistema de coordenadas qualquer, significa encontrar três translações e três rotações entre este referencial e o referencial em que o objeto é descrito.

No caso da **localização de objetos**, muitas abordagens estabelecem o referencial da câmara como absoluto e procuram calcular a transformação da câmara para o objeto.

O problema de **localização da câmara**, por sua vez, consiste em determinar a transformação do referencial global para o referencial da câmara; em Fotogrametria, este mesmo problema é denominado de **Resseção Espacial.**

No problema de **localização da câmara** são calculados apenas as três rotações e as três translações entre o referencial objeto e o referencial imagem, denominados **parâmetros extrínsecos ou parâmetros de orientação exterior**. Alguns métodos levam em conta também os parâmetros que descrevem a geometria interna da câmara, conhecidos como **parâmetros intrínsecos ou parâmetros de orientação interior** (para câmaras CCD, os mais importantes são a distância focal, o centro da imagem ou ponto principal, a(s) constante(s) de distorção ótica e o fator de escala em x). O método em que os parâmetros extrínsecos são calculados em conjunto com os parâmetros intrínsecos é denominado de **Calibração da Câmara.**

Mais recentemente, alguns autores, ao desenvolverem métodos para a localização da câmara, também denominaram este problema de Calibração da Câmara, criando uma certa confusão de nomenclaturas. Neste trabalho, estabelece-se que **Calibração** refere-se especificamente à obtenção dos parâmetros intrínsecos e extrínsecos; no problema de **localização**, parte-se do pressuposto que os parâmetros intrínsecos são conhecidos a priori, buscando-se a determinação apenas dos parâmetros extrínsecos. Fica estabelecido, ainda, que o referencial absoluto a ser adotado é o referencial do espaço objeto.

Os métodos de localização e calibração existentes podem ser classificados:

. Com relação ao conjunto de parâmetros a serem determinados:

a) **Calibração** : São calculados tanto os parâmetros intrínsecos quanto os parâmetros extrínsecos. Em alguns métodos existe um agrupamento

destes parâmetros para tornar o modelo linear, obtendo-se novos parâmetros sem significado físico. Neste caso, seria mais rigoroso classificar tais métodos como "parametrização", ao invés de calibração. É o caso da DLT, proposta por Abdel-Aziz e Karara (1971);

b) **Localização** : São calculados apenas os parâmetros extrínsecos. Pode ser ainda subdividida em **localização da câmara** ou **localização de objetos**;

. **Com relação ao conjunto de feições usadas como apoio:**

a) **Pontos** : São aqueles métodos em que pontos no espaço imagem e seus homólogos no espaço objeto são usados para permitir a determinação da localização ou calibração;

b) **Linhas:** São métodos propostos mais recentemente, com o objetivo de minimizar alguns problemas encontrados com o uso de pontos. Estas linhas podem ser cônicas, mas a maioria dos métodos utiliza linhas retas, quer pela facilidade em detectá-las na imagem, quer pela simplicidade nos modelos resultantes;

c) **Estruturas:** São métodos que utilizam conjuntos de linhas, como por exemplo, cubos, paralelepípedos, quadrados, formas retangulares, etc.... Embora os modelos obtidos sejam frequentemente mais simples que os anteriores, este tipo de configuração restringe a aplicação, pois o método só é aplicável onde existir um conjunto de feições adequado;

. **Com relação ao número de estágios para o cálculo dos parâmetros:**

a) **Simultâneo:** Todos os parâmetros são calculados simultâneamente, através da resolução de sistemas de equações; quando o modelo é não linear são necessárias iterações;

b) **Multi-estágios:** O cálculo dos parâmetros é feito através de estágios consecutivos. Existem métodos em que são calculados grupos de parâmetros a cada estágio ou métodos em que um parâmetro é calculado em cada estágio. Embora estes métodos multi-estágios ofereçam a possibilidade de redução do custo computacional existe um efeito de propagação dos erros entre cada estágio, o que não ocorre nos métodos simultâneos; com isto, estes métodos multi-estágios são mais sensíveis ao ruído (Weng et al, 1992);

2.1.1 MÉTODO CLÁSSICO DE CALIBRAÇÃO.

Este procedimento de calibração é baseado na equação de colinearidade, (Apêndice 1) acrescida dos parâmetros intrínsecos. Este problema é não linear, sendo usado o Método dos Mínimos Quadrados para tratar os erros observacionais e encontrar uma solução única. São usados pontos de apoio cujas coordenadas no espaço objeto são conhecidas e cujas coordenadas no espaço imagem são medidas. Para maiores detalhes ver: Tommaselli e Tozzi (1991) e Faig e Moniwa (1973).

Neste método, denominado clássico, algumas restrições devem ser observadas para a obtenção de resultados confiáveis:

- os pontos de apoio não devem ser coplanares ou então as imagens devem ser oblíquas em relação ao plano médio destes pontos;
- devem ser tomadas várias imagens do mesmo conjunto de pontos, preferencialmente rotacionadas de 90° em relação ao eixo ótico;
- todas as imagens são, então, tratadas simultâneamente, permitindo obter os parâmetros extrínsecos de todas as imagens e o conjunto de parâmetros intrínsecos comuns.

Embora este método clássico seja bastante preciso, é impraticável a sua aplicação em problemas que demandem tempo real, pois:

- o procedimento exige várias imagens;
- são necessárias várias iterações até que o processo de estimação de parâmetros atinja uma convergência satisfatória;
- a extração dos pontos imagem e sua correspondência com os pontos no espaço objeto também é dispendiosa computacionalmente.

O modelo matemático completo, que relaciona coordenadas dos pontos de controle no espaço objeto e suas homólogas no referencial de máquina (frame buffer), considerando apenas os erros sistemáticos de grande magnitude, é dado pelas equações (2.1.1).

$$x_f - c_x + \delta x_r + \delta g_x = f \frac{(X - Xc)\, m_{11} + (Y - Yc)\, m_{12} + (Z - Zc)\, m_{13}}{(X - Xc)\, m_{31} + (Y - Yc)\, m_{32} + (Z - Zc)\, m_{33}}$$

$$y_f - c_y + \delta y_r + \delta g_y = f \frac{(X - Xc)\, m_{21} + (Y - Yc)\, m_{22} + (Z - Zc)\, m_{23}}{(X - Xc)\, m_{31} + (Y - Yc)\, m_{32} + (Z - Zc)\, m_{33}} \qquad (2.1.1)$$

onde:

. x_f e y_f são as coordenadas do ponto no referencial de máquina, e X, Y, Z são as coordenadas do seu homólogo no espaço objeto;

. c_x e c_y são as coordenadas do centro da imagem (ponto principal);

. δx_r e δy_r são os erros devido à distorção radial simétrica, que são modelados analíticamente pelas equações (2.1.2):

$$\delta x_r = (k_1 r^2 + k_2 r^4 + k_3 r^6).x$$
$$\delta y_r = (k_1 r^2 + k_2 r^4 + k_3 r^6).y \qquad (2.1.2)$$

. $r^2 = (x_f - c_x)^2 + (y_f - c_y)^2$ (2.1.3)

. δg_x e δg_y são os erros devido ao fator de escala em x, sendo modelados pelas equações (2.1.4):

$$\delta g_x = (x_f - c_x).ds_x$$
$$\delta g_y = 0 \qquad (2.1.4)$$

. O parâmetro f é a distância focal da câmara;

. m_{ij} são os elementos da matriz de rotação (Apêndice 1);

. Xc, Yc, Zc são as coordenadas do centro perspectivo da câmara (C.P.);

Os parâmetros δ_i , c_x e c_y são correções às coordenadas de máquina (x_f, y_f), transformando-as em coordenadas no espaço imagem, que correspondem ao valor calculado por meio das equações projetivas.

Quando não se conhecem os parâmetros dos modelos dos erros sistemáticos, deve-se calculá-los simultâneamente com os parâmetros extrínsecos. Este procedimento é conhecido como **parametrização dos erros sistemáticos**, ou **auto-calibração** ou simplesmente, **calibração**. A grande vantagem deste procedimento é a confiabilidade que se obtém, pois todos os parâmetros são calculados ao mesmo tempo.

Para simplificar o modelo de calibração, considera-se apenas o termo k_1 para as distorções óticas. Explicitando as coordenadas de máquina em função dos parâmetros intrínsecos e extrínsecos, o modelo para uma câmara CCD

torna-se:

$$x_f = c_x - (x_f - c_x)(k_1 r^2) - (x_f - c_x)ds_x + f\frac{m}{q}$$
$$y_f = c_y - (y_f - c_y)(k_1 r^2) + f\frac{n}{q} \qquad (2.1.5)$$

Se as coordenadas de máquina forem utilizadas em seu estado bruto, ou seja, coordenadas do pixel, todos os valores serão calculados tendo como unidade o pixel. Para evitar este problema, basta transformar previamente as coordenadas do "frame buffer" para uma escala aproximada, usando-se o tamanho aproximado do pixel. Esta transformação preliminar engloba, ainda, a translação para o centro aproximado da imagem, que é o centro do "framme buffer". As diferenças para o centro real serão obtidas por cx e cy.

Supondo que x_f e y_f são quantidades sujeitas à flutuações probabilísticas (observações), e associando à elas resíduos v_x e v_y obtem-se o modelo do tipo:

$$L_a = F(X_a) \quad \text{com} \quad L_a = L_b + v \qquad (2.1.6)$$

onde:

$$X_a^T = [c_x,\ c_y,\ k_1,\ ds_x,\ f,\ \kappa,\ \phi,\ \omega,\ X_c,\ Y_c,\ Z_c]$$

O vetor L_b é composto pelas coordenadas de máquina dos pontos de controle (observações):

$$L_b^T = [\ x_f^1,\ y_f^1,\ x_f^2,\ y_f^2, \ldots,\ x_f^i,\ y_f^i, \ldots .\ ,x_f^m,\ y_f^m\ ]$$

Cada ponto de controle gera duas equações e, portanto, o vetor L_b terá ordem 2m, onde m é o número de pontos.

O cálculo do vetor dos parâmetros X_a, é o objetivo do procedimento de calibração.

O vetor X_0 é constituído por valores aproximados para os parâmetros. Estes valores podem ser arbitrados ou obtidos por cálculos aproximados, ou ainda de calibrações anteriores. Quanto mais precisas as aproximações, menor será o número de iterações necessárias para a convergência da solução.

A matriz **A**, das derivadas parciais será do tipo:

$$\mathbf{A} = \begin{bmatrix} \frac{\partial Fx^1}{\partial c_x} & \frac{\partial Fx^1}{\partial c_y} & \frac{\partial Fx^1}{\partial k_1} & \frac{\partial Fx^1}{\partial ds} & \frac{\partial Fx^1}{\partial f} & \frac{\partial Fx^1}{\partial \kappa} & \frac{\partial Fx^1}{\partial \phi} & \frac{\partial Fx^1}{\partial \omega} & \frac{\partial Fx^1}{\partial Xc} & \frac{\partial Fx^1}{\partial Yc} & \frac{\partial Fx^1}{\partial Zc} \\ \frac{\partial Fy^1}{\partial c_x} & \frac{\partial Fy^1}{\partial c_y} & \frac{\partial Fy^1}{\partial k_1} & \frac{\partial Fy^1}{\partial ds} & \frac{\partial Fy^1}{\partial f} & \frac{\partial Fy^1}{\partial \kappa} & \frac{\partial Fy^1}{\partial \phi} & \frac{\partial Fy^1}{\partial \omega} & \frac{\partial Fy^1}{\partial Xc} & \frac{\partial Fy^1}{\partial Yc} & \frac{\partial Fy^1}{\partial Zc} \\ \cdots\cdots & & \cdots\cdots\cdots & & & & & & & \cdots\cdots & \\ \frac{\partial Fx^m}{\partial c_x} & \frac{\partial Fx^m}{\partial c_y} & \frac{\partial Fx^m}{\partial k_1} & \frac{\partial Fx^m}{\partial ds} & \frac{\partial Fx^m}{\partial f} & \frac{\partial Fx^m}{\partial \kappa} & \frac{\partial Fx^m}{\partial \phi} & \frac{\partial Fx^m}{\partial \omega} & \frac{\partial Fx^m}{\partial Xc} & \frac{\partial Fx^m}{\partial Yc} & \frac{\partial Fx^m}{\partial Zc} \\ \frac{\partial Fy^m}{\partial c_x} & \frac{\partial Fy^m}{\partial c_y} & \frac{\partial Fy^m}{\partial k_1} & \frac{\partial Fy^m}{\partial ds} & \frac{\partial Fy^m}{\partial f} & \frac{\partial Fy^m}{\partial \kappa} & \frac{\partial Fy^m}{\partial \phi} & \frac{\partial Fy^m}{\partial \omega} & \frac{\partial Fy^m}{\partial Xc} & \frac{\partial Fy^m}{\partial Yc} & \frac{\partial Fy^m}{\partial Zc} \end{bmatrix}$$

onde:

$$\frac{\partial Fx}{\partial c_x} = 1 + k_1 . r^2 + ds_x \qquad \frac{\partial Fx}{\partial c_y} = 0$$

$$\frac{\partial Fx}{\partial k_1} = -(x_f - c_x) . r^2 \qquad \frac{\partial Fx}{\partial ds_x} = -(x_f - c_x)$$

$$\frac{\partial Fx}{\partial f} = \frac{m}{q}$$

$$\frac{\partial Fy}{\partial c_x} = 0 \qquad \frac{\partial Fy}{\partial c_y} = 1 + k_1 . r^2$$

$$\frac{\partial Fy}{\partial k_1} = -(y_f - c_y) . r^2 \qquad \frac{\partial Fy}{\partial ds_x} = 0$$

$$\frac{\partial Fy}{\partial f} = \frac{n}{q}$$

$$\frac{\partial Fx}{\partial \kappa} = \frac{f}{q}(\Delta X . m_{21} + \Delta Y . m_{22} + \Delta Z . m_{23}) = f . \frac{n}{q}$$

$$\frac{\partial Fx}{\partial \phi} = -\frac{f}{q . q} [q (\Delta X . \text{sen}\phi . \cos\kappa - \Delta Y . \text{sen}\omega . \cos\phi . \cos\kappa + \Delta Z . \cos\omega . \cos\phi . \cos\kappa) + m (\Delta X . \cos\phi + \Delta Y . \text{sen}\omega . \text{sen}\phi - \Delta Z . \cos\omega . \text{sen}\phi)]$$

$$\frac{\partial Fx}{\partial \omega} = -\frac{f}{q . q} [q (\Delta Y . m_{13} - \Delta Z . m_{12}) - m(\Delta Y . m_{33} - \Delta Z . m_{32})]$$

$$\frac{\partial Fx}{\partial Xc} = -\frac{f}{q . q} (q . m_{11} - m . m_{31}) \qquad \frac{\partial Fx}{\partial Yc} = -\frac{f}{q . q} (q . m_{12} - m . m_{32})$$

$$\frac{\partial Fx}{\partial Zc} = -\frac{f}{q . q} (q . m_{13} - m . m_{33})$$

$$\frac{\partial Fy}{\partial \kappa} = -\frac{f}{q}\,(\ \Delta X.m_{11} + \Delta Y.m_{12} + \Delta Z.m_{13}) = -f.\frac{m}{q}$$

$$\frac{\partial Fy}{\partial \phi} = -\frac{f}{q.q}\,[-q\,(\ \Delta X.\mathrm{sen}\phi.\mathrm{sen}\kappa - \Delta Y.\mathrm{sen}\omega.\cos\phi.\mathrm{sen}\kappa + \Delta Z.\cos\omega.\cos\phi.\mathrm{sen}\kappa) + n\,(\ \Delta X.\cos\phi + \Delta Y.\mathrm{sen}\omega.\mathrm{sen}\phi - \Delta Z.\cos\omega.\mathrm{sen}\phi)]$$

$$\frac{\partial Fy}{\partial \omega} = -\frac{f}{q.q}\,[\ q\,(\ \Delta Y.m_{23}. - \Delta Z.m_{22}\ ) - n\,(\ \Delta Y.m_{33} - \Delta Z.m_{32})]$$

$$\frac{\partial Fy}{\partial Xc} = -\frac{f}{q.q}\,(\ q.m_{21} - n.m_{31}) \qquad \frac{\partial Fy}{\partial Yc} = -\frac{f}{q.q}\,(\ q.m_{22} - n.m_{32})$$

$$\frac{\partial Fy}{\partial Zc} = -\frac{f}{q.q}\,(\ q.m_{23} - n.m_{33})$$

O procedimento de calibração pelo método clássico introduz algumas situações críticas, oriundas da existência de correlação entre os parâmetros. Para quebrar esta correlação e aumentar a precisão do processo, é útil processar simultaneamente várias imagens. Se a focal não for alterada, um conjunto de imagens terá em comum os mesmos parâmetros intrínsecos e apenas os parâmetros extrínsecos diferentes. O sistema toma a seguinte forma, para um arranjo multi-imagem:

O vetor dos parâmetros ajustados:

$$X_a^T = [c_x,\ c_y,\ k_1,\ ds_x,\ f,\ \kappa^1,\ \phi^1,\ \omega^1,\ X_c^1,\ Y_c^1,\ Z_c^1, \kappa^2,\ \phi^2,\ \omega^2,\ X_c^2,\ Y_c^2,\ Z_c^2$$

$$\ldots\ldots\ldots\ \kappa^n,\ \phi^n,\ \omega^n,\ X_c^n,\ Y_c^n,\ Z_c^n\ ]$$

O vetor das observações L_b terá ordem $2mxn$, onde m é o número de pontos e n é o número de imagens:

$$L_b^T = [x_1^1,\ y_1^1,\ x_2^1,\ y_2^1,\ \ldots,\ x_m^1,\ y_m^1,\ x_1^2,\ y_1^2,\ \ldots x_o^2,\ y_o^2,\ \ldots.\ x_p^n,\ y_p^n]$$

A matriz peso será calculada a partir da variância das observações. Desprezando eventuais correlações entre as coordenadas x e y, a matriz peso das observações será considerada diagonal, e os elementos desta diagonal serão dados por:

$$p_{x_i} = \frac{\sigma_0^2}{\sigma_x^2} \qquad \text{e} \qquad p_{y_i} = \frac{\sigma_0^2}{\sigma_y^2} \tag{2.1.7}$$

onde:

- σ_0^2 é a variância da observação de peso unitário;
- σ_x^2 e σ_y^2 são as variâncias das observações;

O cálculo dos parâmetros será realizado num processo iterativo que pode ser resumido através das etapas mostradas na figura 2.1.1.

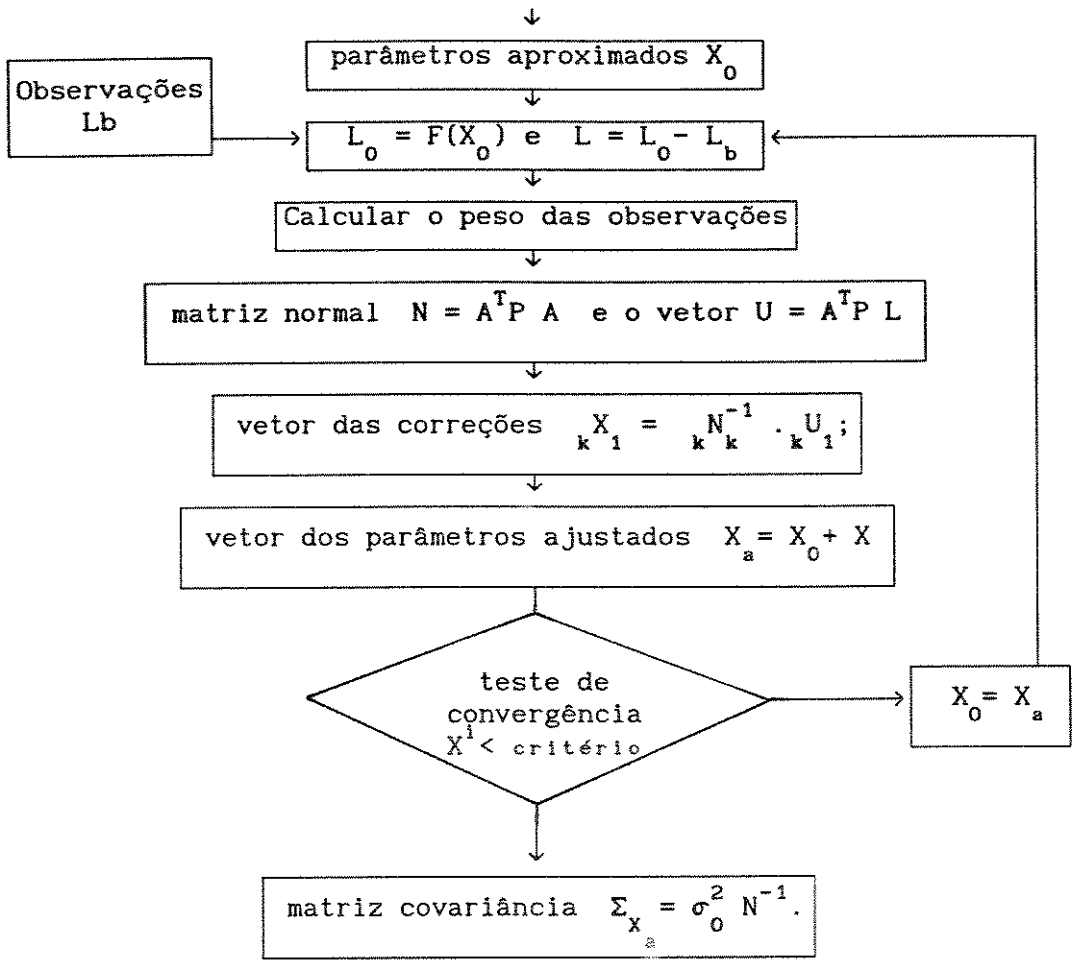


Figura 2.1.1 Fluxo de procedimentos para o cálculo dos parâmetros de calibração pelo método clássico.

Neste método, obtem-se um sistema de equações normais de ordem k (k = 5 + 6x(nº de imagens)). O número de iterações dependerá do número de pontos, da configuração geométrica e dos parâmetros aproximados.

Um estudo mais aprofundado da eficiência deste método, bem como resultados experimentais, pode ser encontrado em Tommaselli e Tozzi (1991). Nesse trabalho, procurou-se mostrar como os parâmetros intrínsecos variam de

acordo com a distância focal. Esta hipótese, confirmada experimentalmente, é útil para uma câmara do tipo auto-foco, pois seria necessário calibrar apenas a focal em tempo real, sendo que os demais parâmetros poderiam ser interpolados de tabelas previamente calculadas.

2.1.2 Outros métodos de localização e calibração usando pontos de apoio.

Como alternativa ao método clássico, apresentado em 2.1.1, algumas abordagens procuram adotar um modelo linear, evitando a linearização por Taylor e a consequente necessidade de iterações. Um destes procedimentos é a DLT, Transformação Linear Direta, desenvolvida por Abdel-Aziz e Karara (1971). Este método de calibração, apesar de bastante eficiente, não fornece explicitamente os parâmetros extrínsecos.

Outra abordagem linear de calibração, desenvolvida mais recentemente por Tsai (1985), utiliza a restrição de alinhamento radial e um conjunto de pontos coplanares para separar o cálculo dos parâmetros de calibração em estágios, que resultam em equações lineares. No primeiro estágio calcula-se a matriz de rotação e as translações em x e y; no segundo estágio calcula-se a translação em z, a distância focal e a distorção ótica. O fator de escala em x e o centro da imagem não são calculados, sendo necessário recorrer a outros métodos (Tsai e Lenz, 1989).

Um modelo robusto de localização foi apresentado por Fischler e Bolles (1981), com o objetivo de levar em conta, além dos erros randômicos, também os erros grosseiros que são frequentemente cometidos no processo de classificação dos pontos de apoio na imagem. Este paradigma, denominado RANSAC (Random Sample Consensus) foi aplicado ao problema de localização da câmara. Foram desenvolvidas expressões analíticas para o cálculo dos parâmetros, como uma alternativa ao procedimento iterativo convencional, ou aos modelos lineares. Os autores desenvolveram um dos primeiros estudos onde menciona-se o aspecto de multiplicidade de soluções nos problemas de localização. Além da questão de multiplicidade de soluções, outros problemas que não foram discutidos pelos autores devem ser considerados, como a propagação de erros observacionais entre o cálculo de cada um dos parâmetros. Uma das grandes vantagens deste método é que não são necessárias aproximações iniciais.

Outras soluções analíticas para o problema de localização da câmara e localização de objetos foram propostas, tendo como objetivo básico desenvolver

expressões fechadas para o cálculo dos parâmetros como, por exemplo, as abordagens de Holt e Netravali (1991), Haralick et al (1989), Wang (1992), Zeng e Wang (1992) e Linnainmaa et al (1988). Expressões fechadas ("closed form") são aquelas que permitem o cálculo da localização sem a necessidade de valores aproximados iniciais e iterações.

Um dos grandes gargalos dos processos mencionados é o problema de localização dos pontos de apoio na imagem, que consiste em calcular as coordenadas imagem do ponto e determinar a correspondência entre o ponto na imagem e seu homólogo no espaço objeto. Fischler e Bolles (1981) referem-se a este problema como problema de classificação. Eventuais erros de classificação são considerados erros grosseiros e seus efeitos na solução do problema de localização são desastrosos.

A localização dos pontos na imagem é feita através da técnica de "template matching" ou, então, utilizando junções de arestas. A localização dos pontos usando junções é mais simples e precisa, pois baseia-se nas técnicas de detecção de bordas e de vetorização. A localização de junções, entretanto, possui alguns problemas como a existência de múltiplas intersecções entre as retas detectadas na imagem.

Para evitar estes problemas de detecção de pontos e junções algumas propostas mais recentes utilizam diretamente as linhas retas ou curvilíneas como apoio para a calibração e localização da câmara.

2.1.3 Localização e Calibração usando linhas e estruturas.

A utilização de feições lineares nos problemas de calibração e localização, particularmente as linhas retas, possui uma série de vantagens em relação aos pontos de apoio:

- ambientes modificados pelo homem e, em particular, ambientes industriais, contém uma ampla gama de arestas, que serão imageadas como bordas;
- as linhas retas são fáceis de detectar em uma imagem, usando procedimentos de detecção de bordas e métodos de vetorização;
- o problema de correspondência é menos dependente de valores iniciais, pois as linhas, geralmente, agrupam-se em estruturas mais complexas, que podem ser reconhecidas (por exemplo, um conjunto de retas paralelas);
- os parâmetros que descrevem linhas, especialmente as retas, podem ser

obtidos com precisão sub-pixel, como será verificado no decorrer do trabalho;

Os primeiros trabalhos na área de Visão de Máquina caracterizaram-se pelo estudo de cenas compostas de objetos poliédricos, definidos por um conjunto de arestas retas, cujas imagens são, também, retas. A característica destes trabalhos iniciais foi a tentativa de reconhecer e localizar objetos estabelecendo hipóteses de localização e testando estas hipóteses. De modo geral, esta mesma filosofia está presente ainda hoje nos métodos de reconhecimento. Entretanto, com relação à localização, métodos mais eficientes e precisos são necessários.

Os trabalhos iniciais em Visão de Máquina, baseavam o reconhecimento e a localização em junções determinadas pelas intersecções das arestas dos poliedros (Roberts, 1982), isto é, pontos de apoio.

Para evitar o uso de pontos e junções alguns métodos têm sido propostos usando outros tipos de feições como apoio como, por exemplo, linhas retas ou curvilíneas.

Halarick (1980) demonstrou que linhas retas são projetadas, inclusive em perspectiva, como linhas retas, estabelecendo equações projetivas específicas para linhas retas.

Masry (1980) introduziu o uso de feições digitalizadas como uma fonte de apoio para o cálculo da resseção espacial, a partir de mapas digitalizados da área fotografada.

Lugnani (1980) desenvolveu e testou modelos e metodologias para o caso da resseção espacial, além de propor modelos para a transformação de similaridade. A abordagem proposta supõe que não há correspondência entre os pontos discretos que definem uma feição no espaço objeto, e os pontos observados na mesma feição na fotografia (espaço imagem). Existe, então, apenas uma correspondência feição a feição.

Barnard (1983) apresentou um método para o cálculo da orientação espacial de planos através da localização de linhas e pontos de fuga e da projeção de ângulos. Barnard partiu do princípio que as formas imageadas eram constituídas de linhas retas e planos. Embora, neste trabalho, Barnard estivesse enfatizando um problema de reconstrução das formas do objeto, alguns conceitos desenvolvidos são aplicáveis ao problema de localização. É o caso do conceito de plano de interpretação, que será discutido no Capítulo 3.

Liu e Huang (1988-a e 1988-b) desenvolveram um algoritmo para a estimativa do movimento de corpos rígidos usando a correspondência de linhas

retas. O procedimento foi posteriormente aplicado ao problema de localização da câmara (Liu, Huang e Faugeras, 1990), permitindo o cálculo dos parâmetros de orientação em dois estágios (detalhes na seção 2.1.4).

Tommaselli e Lugnani (1988) propuseram um modelo matemático específico para linhas retas, baseado na equivalência entre os parâmetros do plano formado pela reta no espaço imagem e o centro perspectivo, e a reta no espaço objeto e o mesmo C.P.. Neste modelo não há necessidade de correspondência entre um ponto na reta no espaço objeto e seu homólogo no espaço imagem. O procedimento foi testado com o problema de resseção espacial (localização) e formação de modelos (estereo).

Mulawa e Mikhail (1988) apresentaram o desenvolvimento de equações para o tratamento de linhas retas e seções cônicas (círculos, elipses, hipérboles e parábolas). Os modelos desenvolvidos foram analisados em relação à procedimentos como a resseção espacial. O modelo para linhas retas fundamenta-se na constatação de que os vetores posição de um ponto na reta no espaço objeto e de um ponto na imagem pertencente à reta homóloga, e o vetor direção da reta são coplanares. Os modelos obtidos são do tipo implícito.

Dhome et al (1989) criaram um método para determinar a localização de objetos, baseado na interpretação de um tripleto de linhas retas deste objeto em uma única imagem. A formulação desenvolvida segue os mesmos princípios de outros modelos, definindo-se um "plano de interpretação" que contém uma reta no objeto, sua correspondente na imagem e o centro perspectivo. Os modelos obtidos são análogos aos de Liu, Huang e Faugeras (1990).

Dhome et al (1989), considerando a resolução do sistema de equações 3x3 muito dispendiosa, partiram para a obtenção de uma solução analítica direta. Consideraram, então, uma situação particular, com somente três retas no objeto, estabelecendo um sistema de coordenadas do modelo do objeto, de tal maneira que a primeira linha fosse colinear com o eixo **X**, e que a segunda linha fosse paralela ao plano **XoY**, realizando operação análoga no espaço imagem. Com isto restaram apenas duas rotações incógnitas, para as quais os autores desenvolveram uma expressão do oitavo grau, que pode ser resolvida por métodos iterativos. Para calcular as translações recorreram a um sistema de 3 equações lineares, semelhantes às desenvolvidas por Liu et al (1990). Embora bastante eficiente, este algoritmo baseia-se no uso de apenas três linhas retas e não tem nenhum tipo de avaliação dos erros finais nos parâmetros obtidos.

Wang e Tsai (1990) desenvolveram um método de calibração da câmara, baseado no conceito de linhas e pontos de fuga e usando um paralelepípedo

regular como objeto de calibração. O modelo foi desenvolvido a partir da equação de colinearidade, encontrando-se expressões que relacionam os pontos de fuga com os parâmetros intrínsecos e extrínsecos. Os parâmetros intrínsecos calibrados são a distância focal e o centro da imagem. Não foram considerados a constante de distorção ótica nem o fator de escala em x. As linhas que conectam os três pontos de fuga definem um triângulo, cujo ortocentro é o centro da imagem. Os ângulos de orientação da câmara são calculados a partir das inclinações das linhas de fuga. A calibração da distância focal é feita a partir da área do triângulo definido pelos três pontos de fuga. Um trabalho semelhante pode ser encontrado em Chen et al (1989), embora, neste caso, sejam calculados apenas os parâmetros extrínsecos.

Usando uma estratégia semelhante, Caprile e Torre (1990) apresentaram um método para calibrar simultaneamente duas ou mais câmaras, a partir de um conjunto adequado de feições no espaço, como, por exemplo, um conjunto de linhas retas. O método é baseado no conceito de pontos de fuga, sendo dividido em dois passos: no primeiro passo são calculados alguns dos parâmetros intrínsecos da câmara (distância focal e centro da imagem) usando um cubo como objeto de calibração; no segundo passo são calculados os parâmetros extrínsecos a partir de um conjunto de linhas coplanares.

Echigo (1990) propôs um método para a calibração da câmara a partir de três conjuntos de linhas retas. Inicialmente calculam-se as rotações, independentemente da distância focal e das translações. Depois, são calculadas as translações e a distância focal. Chen e Jiang (1991) também desenvolveram um método de calibração baseado no conceito de pontos de fuga.

Lee et al (1990) empregaram formas retangulares com dimensões conhecidas para determinar a localização da câmara . Este tipo de estrutura (composta por feições retas) é abundante em ambientes industriais e pode ser exemplificado por formas como portas, mesas, janelas, corredores, etc... .Para desenvolver as expressões para o cálculo dos seis parâmetros extrínsecos, os autores partiram de um retângulo, de tal maneira que os eixos do referencial do espaço objeto fossem paralelos aos lados do retângulo. Este procedimento é bastante simples, e utiliza expressões analíticas, cujo tempo de processamento é pequeno. Os autores desenvolveram procedimentos que permitem, de modo semelhante ao apresentado, o uso de formas retangulares parciais, para o caso de oclusão de algum canto. Para aumentar a precisão e confiabilidade do processo podem ser usadas várias formas retangulares, obtendo-se várias estimativas para os parâmetros, calculando-se, então, um valor médio.

Um método análogo, usando a projeção perspectiva de um retângulo, pode

ser encontrado em Haralick (1989). Haralick mostrou que existe informação suficiente em um retângulo de tamanho desconhecido para calcular os ângulos de orientação da câmara em relação ao plano deste retângulo. Mostrou ainda, que, se o tamanho do retângulo for conhecido, é possível calcular as coordenadas tridimensionais de seus vértices. Posteriormente, Penna (1991) generalizou o método desenvolvido por Haralick, utilizando-se de quadriláteros arbitrários, mas com formato e tamanho conhecidos.

Han e Rhee (1992) propuseram uma técnica usando um círculo como objeto de calibração. Chen e Tsai (1991) usaram objetos com uma face poligonal e faces laterais perpendiculares como objeto de calibração.

Além dos métodos revisados, específicos para calibração e localização, algumas abordagens propostas para o cálculo de movimentos de corpos rígidos ("motion") também podem ser adaptados para aquelas finalidades. Foi o caso de Liu et al (1990) que adaptaram um modelo de estimação de "motion", proposto anteriormente (Liu e Huang (1988a e 1988b)). Nesta linha, vários trabalhos que baseiam-se na correspondência de linhas retas vêm sendo apresentados, podendo-se citar: Mitiche e Habelrih (1989), Spetsakis e Aloimonos (1990), Salari e Jong (1990), Wang et al (1991), Weng et al (1992) e Goldgof et al (1992). Em Mitiche et al (1989) encontra-se uma discussão sobre alguns métodos de estimação de movimento e estrutura de objetos, com ênfase no número mínimo de imagens e de linhas.

2.1.4 Localização usando o modelo de Liu e Huang.

O modelo de Liu e Huang (1988-a e 1988-b), posteriormente adaptado para o problema de localização (Liu et al, 1990) baseia-se no uso de linhas retas, que são conhecidas no espaço objeto e cujas homólogas são medidas no espaço imagem. Os autores partem da hipótese de que as correspondências entre retas homólogas são estabelecidas previamente. No modelo de Liu e Huang é utilizada uma sequência de transformações em que as rotações são aplicadas antes das translações.

Uma das principais características desta abordagem é a separação do cálculo dos parâmetros de localização em dois estágios.

No primeiro estágio são calculadas as rotações, com a resolução de um sistema de ordem 3. Este sistema é estabelecido a partir das equações (2.1.8). Nestas equações, n é o vetor das direções da reta no espaço objeto e **N** é o

vetor normal ao plano projetante no espaço imagem. O vetor **N**, rotacionado, e portanto no referencial objeto, é perpendicular à linha reta no espaço objeto; o produto escalar destes vetores é nulo:

$$\mathbf{n}^T.R.\mathbf{N} = 0 \qquad (2.1.8)$$

O vetor **n** é composto pelos cossenos diretores da reta no espaço, enquanto que o vetor **N** pode ser calculado usando a equação da reta medida na imagem. A matriz R é a matriz de rotação e seus elementos são as incógnitas desta equação. Com um algorítmo linear (considerando os m_{ij} como incógnitas) são necessárias, no mínimo oito retas para resolver este sistema e, com o algorítmo não linear (considerando os 3 ângulos de rotação como incógnitas), apenas três, porém em um processo iterativo. Usando o modelo (2.1.8) as observações serão os parâmetros da reta imagem e cada reta gerará uma equação.

No segundo estágio são calculadas as translações, através de um modelo linear com 3 incógnitas (equação (2.1.9)).

$$\mathbf{N}.\mathbf{T} = \mathbf{N}.(R^{-1}\mathbf{p}o) \qquad (2.1.9)$$

onde **T** é o vetor das translações e **p**o é um ponto qualquer na reta no espaço objeto. A matriz de rotação R foi calculada no estágio anterior e o vetor **N** também é conhecido (calculado a partir da reta na imagem). Esta segunda etapa é mais simples, pois o modelo é linear.

A principal vantagem do algorítmo de Liu e Huang é o desacoplamento do cálculo dos parâmetros em duas etapas, embora o modelo obtido seja do tipo implícito.

2.2 Reconhecimento

O problema de reconhecimento de objetos tridimensionais, a partir de imagens bidimensionais, têm recebido considerável atenção desde os primeiros trabalhos de Visão. Métodos bastante simples podem ser utilizados para classificar objetos e reconhecê-los, usando características como cor, número de vértices, tamanho aproximado, etc.... Quando o número de objetos possíveis for muito grande, ou possuirem características semelhantes, estes critérios simples podem não ser suficientes para reconhecer inequivocamente um objeto. A comparação de primitivas extraídas na imagem do objeto com o modelo armazenado

é, então, a alternativa mais utilizada ("model-based Vision"). Contudo, um dos grandes obstáculos ao reconhecimento é o fato do ponto de vista, isto é, a posição e orientação da câmara, serem desconhecidos.

Este tipo de abordagem de reconhecimento consiste, basicamente, em processar a(s) imagem(ns) de um objeto para extrair algumas primitivas (denominadas de primitivas de indexação) e verificar qual dos modelos pré-armazenados corresponde ao objeto imageado. O trabalho de Roberts (1982), considerado um dos pioneiros em Visão de Máquina, procurava solucionar o problema de reconhecimento em conjunto com o problema de localização. Roberts usou como primitivas as junções de objetos poliédricos.

De modo geral, as várias abordagens existentes para o reconhecimento de objetos adotam estratégias baseadas na extração de primitivas de indexação (Dickinson et al, 1992) e na definição de hipóteses com base nestas primitivas. Estas hipóteses referem-se aos objetos possíveis, dentre aqueles relacionados na base de dados, e às possíveis correspondências entre as primitivas na imagem e no modelo do objeto. Com base nestas correspondências hipotéticas são calculadas as possíveis localizações do(s) objeto(s). As hipóteses são, então, testadas e refutadas com base na projeção de outras primitivas do modelo para a imagem, verificando-se a existência desta primitiva na imagem real. A figura 2.2.1 ilustra esta sequência de procedimentos.

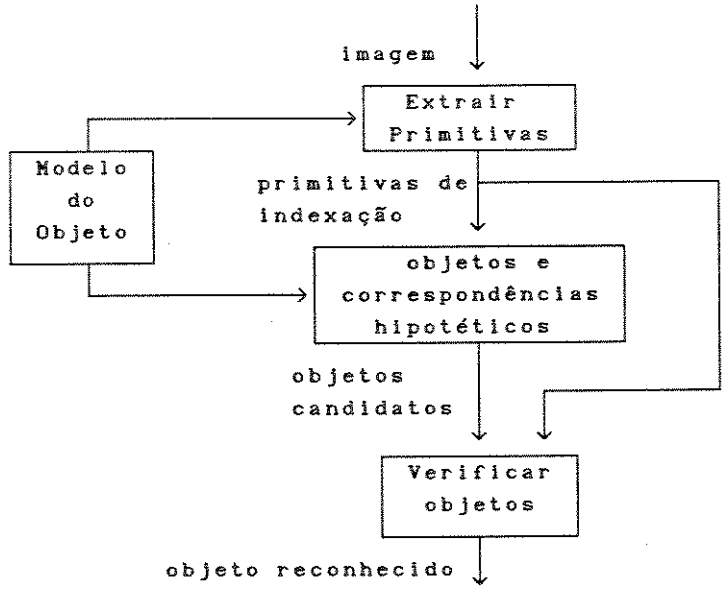


Figura 2.2.1 Etapas no reconhecimento de objetos.
(Adaptado de Dickinson et al (1992))

As primitivas de indexação podem ser desde junções simples, retas ou até estruturas tridimensionais, como cilindros, cones, etc... Estruturas mais complexas, embora exigindo esforço na tarefa de agrupá-las na imagem, reduzem sensivelmente a tarefa de busca e a quantidade de hipóteses. Ao contrário, estruturas simples, do tipo reta, permitem inúmeras correspondências hipotéticas com os modelos de objetos da base de dados.

O problema de reconhecimento pode ser reduzido, portanto, à definição de algumas correspondências iniciais, que serão posteriormente validadas ou refutadas.

Estas correspondências devem ser baseadas em feições (primitivas) invariantes com o ponto de vista (Lowe, 1987), e que sejam suficientemente significantes, de modo a representarem características do objeto. Entre as primitivas mencionadas por Lowe (1987) estão retas paralelas, retas colineares e retas com extremos próximos. Assume-se que as retas paralelas no objeto sofrerão apenas uma pequena convergência devido à projeção perspectiva na imagem. Bergevin e Levine (1992) apresentam um estudo sobre feições invariantes com o ponto de vista, delineando um sistema de reconhecimento baseado em características mais grosseiras dos objetos.

O problema de reconhecimento pode ser, ainda, definido através de busca em grafos. Neste caso, as feições na imagem são estruturadas como um subgrafo do modelo real do objeto. Wong (1992) descreve um método onde os nós do grafo correspondem aos vértices do objeto e as arestas correspondem às bordas do objeto.

Embora estes métodos sejam concebidos para reconhecer um objeto dentre vários possíveis, a determinação de correspondência pode ser considerada como um problema particular do reconhecimento, pois envolve as hipóteses de localização para apenas um objeto. Estas hipóteses podem ser reduzidas, ainda, usando heurísticas adicionais, como, por exemplo, as posições de estabilidade do objeto. Estas posições podem ser especificadas pelo usuário ao definir a base de dados, ou então calculadas com base em algumas condições de estabilidade (Kriegman, 1992).

No presente trabalho, supõe-se que o objeto já foi reconhecido e que uma localização aproximada tenha sido determinada. Uma maneira de acelerar o reconhecimento e localização iniciais é trabalhar em um nível de resolução mais grosseiro, deixando para o processo recursivo, a ser descrito, a tarefa de refinar a estimativa.

2.3 FILTRAGEM

2.3.1 MÉTODOS USANDO FILTRAGEM

A adoção de técnicas de filtragem, além de permitir a estimação de parâmetros levando em conta as observações passadas sem ter que processar todo o conjunto novamente, possui a vantagem de permitir o processamento sequencial, isto é fornecer estimativas do vetor estado a cada observação introduzida. Estas estimativas podem ser empregadas para realimentar a etapa de extração das feições, diminuindo recursivamente o espaço de busca, e o tempo de processamento.

Vários estudos tem demonstrado que a etapa de extração das feições na imagem, e o estabelecimento de correspondência destas com elementos no espaço objeto, é extremamente onerosa computacionalmente (Lee et al, 1990). De fato, se a estratégia de busca de uma determinada feição basear-se na busca por toda a imagem pode-se prever um tempo de processamento excessivo.

Se a informação a priori sobre a localização da câmara for usada, pode-se delimitar uma área provável de presença da feição na imagem. É possível, então, definir um processo de vetorização somente para esta janela e economizar operações desnecessárias. Com um processo de filtragem, a cada nova observação introduzida a estimativa de localização é melhorada e, consequentemente, o espaço de busca diminuido.

Isto implica em afirmar que a adoção de um procedimento recursivo de filtragem, além de aumentar a confiabilidade do processo de localização e de permitir uma avaliação de sua precisão, permite uma redução do custo computacional na etapa de extração de feições na imagem.

Esta característica recursiva está presente na Filtragem Kalman. No MMQ tradicional a ordem do sistema a resolver é igual ao número de parâmetros incógnitos. Isto não acontece na Filtragem Kalman, onde pode-se tratar as observações não correlacionadas isoladamente, eliminando qualquer necessidade de inversão de matrizes. Esta característica da Filtragem Kalman permite estabelecer um método de localização que trate todos os parâmetros simultâneamente a custo computacional equivalente ao de um método de vários estágios.

A aplicação de Filtragem permite, ainda, considerar o aspecto dinâmico de várias aplicações de Visão, como é o caso da localização de objetos. Nestas

aplicações o objeto ou a câmara podem ser móveis (ou ambos), mas estes movimentos são conhecidos aproximadamente. É o caso de um objeto movendo-se em uma esteira, observado por uma câmara. O resultado da localização no instante t_k pode ser usado como uma restrição no instante seguinte, considerando-se que o deslocamento neste intervalo é aproximadamente conhecido.

Este tipo de abordagem dinâmica para o problema de localização não é encontrada frequentemente na literatura. É mais comum, entretanto, no tratamento de problemas relacionados ao movimento de corpos rígidos ("motion").

Usando um modelo simplificado e adotando algumas restrições, Broida e Chellapa (1986) utilizaram Filtragem Kalman para calcular recursivamente os parâmetros de movimento de um corpo rígido, a partir de uma sequência de imagens de dois pontos. Um aspecto utilizado nesta abordagem é com relação ao número de pontos; enquanto as demais técnicas tentam minimizar o efeito do ruído adotando um grande número de pontos (o que demanda tempo computacional para localizar os pontos na imagem), os autores optaram por usar poucos pontos em muitas imagens, melhorando recursivamente a estimativa dos parâmetros de movimento, usando Filtragem Kalman.

Uma prática semelhante foi desenvolvida por Liang et al (1989) para o problema de calibração, tanto da câmara quanto do robô. Neste trabalho os autores trataram a calibração de parâmetros intrínsecos e extrínsecos como um problema de filtragem. Com base em uma sequência de imagens, cada imagem contendo um único ponto de apoio, os parâmetros são recursivamente melhorados usando a Filtragem Kalman. A câmara mantém-se fixa, por um certo período de tempo, enquanto coleta imagens de um robô em cuja garra está materializado o ponto de apoio. As coordenadas do ponto de apoio são calculadas pelo modelo cinemático deste robô. Uma das grandes vantagens deste método é que os parâmetros são continuamente calibrados e qualquer variação, como a mudança da focal, por exemplo, pode ser detectada e filtrada durante a operação do sistema. Este procedimento pode ser considerado robusto, na medida em que está preparado para detectar e se adaptar a mudanças inesperadas nos parâmetros da câmara.

Mais recentemente, Weng et al (1992) discutindo a questão de unicidade e de confiabilidade das soluções no problema de "motion and structure" usando linhas retas, empregaram técnicas de otimização para melhorar os resultados obtidos a partir de expressões fechadas.

2.3.2 FILTRAGEM KALMAN

Existem vários métodos de estimativa de parâmetros a partir de observações, desde o Método dos Mínimos Quadrados (MMQ), concebido por Gauss e Legendre, até a mais recente Filtragem Kalman, introduzida por Kalman na década de 60. O MMQ não assume nenhuma distribuição de probabilidade a priori para as observações, e supõe que os parâmetros não possuam variação temporal, isto é, que o sistema seja estático.

A Filtragem Kalman é um método de estimação Bayeseano, não tendencioso e de variância mínima. Este método pode ser desenvolvido tanto a partir do MMQ (supondo-se **n** e **w** com média zero), como a partir do teorema de Bayes, considerando a probabilidade condicional conhecida a priori (Jazwinski, 1970). A Filtragem Kalman é adequada para o tratamento de sistemas dinâmicos, pois considera a variação temporal do vetor estado. Além disto, alguns dos algoritmos foram desenvolvidos de forma recursiva, reduzindo o esforço computacional.

Seja um sistema estocástico discreto dinâmico descrito pela equação:

$$\mathbf{x}_{k+1} = \phi\ (\mathbf{x}_k,\ t_{k+1},\ t_k) + \Gamma\ (\mathbf{x}_k,\ t_k)\ \mathbf{w}_{k+1} \qquad (2.3.1)$$

onde: $\mathbf{x}_k$ é o vetor estado no instante t_k;

ϕ é uma função n-vetorial de transição de estado;

Γ é uma matriz (n×r) e;

$\mathbf{w}_k$ é um vetor (r) com uma sequência Gaussiana branca de média zero e matriz covariância Q_k, usualmente chamado de ruído de transição de estado.

O modelo (2.3.1) é chamado modelo do sistema ou, ainda, modelo dinâmico.

Seja o vetor $\mathbf{z}_k$ das observações:

$$\mathbf{z}_k = h(\mathbf{x}_k,\ t_k) + \mathbf{n}_k \qquad k = 1,\ 2,\ldots \qquad (2.3.2)$$

onde: $\mathbf{z}_k$ é o vetor das observações;

$\mathbf{n}_k$ é o vetor dos ruidos nas observações, com média zero e matriz covariância R_k.

O modelo (2.3.2) é o chamado modelo de medidas.

Seja, agora, a sequência de observações:

$$Z_l = [\, z_1,\ z_2, \ldots\ldots,\ z_l \,]$$

Dado Z_l o problema de **estimação** discreta consiste em calcular uma estimativa de x_k baseada em Z_l. Se $k = l$, o problema é chamado de **filtragem** discreta. Se $k > l$, o problema é chamado de **predição** discreta. Se $k < l$, o problema é chamado de **"smoothing discreto"**. Os problemas a serem tratados neste trabalho referem-se a problemas de Filtragem discreta, onde tanto o modelo do sistema quanto o modelo de medidas são discretos.

Quando o modelo do sistema e de medida são lineares então:

$$\phi\ (x_k,\ t_{k+1},\ t_k) = F(t_{k+1},\ t_k).x_k \qquad (2.3.3)$$

$$h(x_k,\ t_k) = H(t_k).x_k \qquad (2.3.4)$$

Existem várias formulações diferentes para o problema de filtragem.

2.3.3 Filtro em Batch (Filtro inverso linear)

Esta abordagem, para o caso estático, coincide com o MMQ. Trata-se de um tratamento simultâneo de todas as observações disponíveis no instante t_k.

Dados o sistema dinâmico e as observações referentes a um modelo de medidas linear, a estimativa do estado no instante t_k é dada por:

$$\hat{x}_{k;k} = (H_k^T R_k^{-1} H_k + P_{k;k-1}^{-1})^{-1} (H_k^T R_k^{-1} z_k + P_{k;k-1}^{-1} \hat{x}_{k;k-1}) \qquad (2.3.5)$$

A inversa da matriz covariância para o instante t_k:

$$P_{k;k}^{-1} = P_{k;k-1} + H_k^T R_k^{-1} H_k \qquad (2.3.6)$$

onde: $\hat{x}_{k;k-1}$ representa o estado predito para o instante t_k com base nas medidas realizadas no instante t_{k-1};

$P_{k;k-1}$ representa a matriz covariância predita para o instante t_k com base nas medidas realizadas em t_{k-1}.

É interessante notar que a matriz covariância predita $P_{k;k-1}$ tem o efeito de uma restrição na estimativa de $\hat{x}_{k;k}$.

Para atualizar a estimativa, isto é, estabelecer uma predição para o próximo instante:

$$\hat{x}_{k+1:k} = F_k \hat{x}_{k:k} \tag{2.3.7}$$

$$P_{k+1:k} = F_k P_{k:k} F_k^T + Q_k \tag{2.3.8}$$

É importante notar que a predição da matriz covariância é independente das observações. A obtenção da estimativa para o estado envolve a inversão de uma matriz (nxn). Este Filtro é chamado de Filtro inverso. No próximo ítem será apresentada a formulação que permite reduzir este problema para uma forma que exija a inversão de uma matriz (mxm).

2.3.4 Filtro Kalman Padrão (Kalman-Bucy)

A estimação do estado usando o Filtro Kalman padrão, para o caso discreto, com modelos lineares, é baseada nas expressões (2.3.9) e (2.3.10).

$$\hat{x}_{k:k} = \hat{x}_{k:k-1} + K_k (z_k - H_k \hat{x}_{k:k-1}) \tag{2.3.9}$$

onde K_k é a matriz ganho dada pela expressão:

$$K_k = P_{k:k-1} H_k^T (H_k P_{k:k-1} H_k^T + R_k)^{-1} \tag{2.3.10}$$

Analisando a equação (2.3.9) verifica-se que o ganho K representa um fator multiplicativo aos resíduos, gerando uma correção à estimativa predita. O termo $(z_k - H_k \hat{x}_{k:k-1})$ é o vetor dos resíduos ou, "inovações" (Liang et al, 1989). O termo $(H_k P_{k:k-1} H_k^T + R_k)$ é a matriz covariância destes resíduos.

A matriz covariância do estado filtrado pode ser calculada através da expressão (2.3.11).

$$P_{k:k} = (I - K_k H_k) P_{k:k-1} \tag{2.3.11}$$

Esta expressão, embora simples, está sujeita a instabilidades numéricas. Jazwinski (1970), apresenta uma expressão para calcular a matriz covariância que garante que a matriz obtida será positivo definida, sendo melhor condicionada para cálculos computacionais.

$$P_{k:k} = (I - K_k H_k) P_{k:k-1} (I - K_k H_k)^T + K_k R_k K_k^T \qquad (2.3.12)$$

A transição do estado para o próximo instante é feita de acordo com as expressões (2.3.7) e (2.3.8).

A figura 2.3.1 mostra a sequência da Filtragem Kalman padrão.

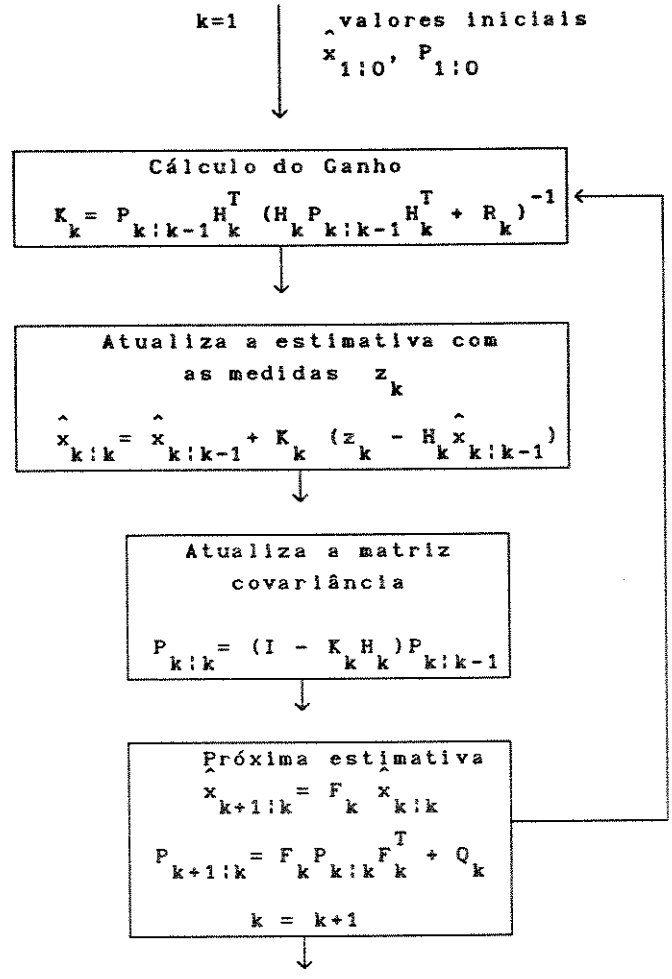


Figura 2.3.1 Sequência da Filtragem Kalman padrão.

Para evitar a inversão de uma matriz (mxm) pode-se usar o procedimento acima para estimar recursivamente o estado no instante t_k, processando uma observação de cada vez. Isto evita a inversão de matrizes, pois o termo $(H_k P_{k:k-1} H_k^T + R_k)$ torna-se um escalar. Isto é feito supondo-se as observações em um instante t_k não correlacionadas. Processa-se z_k^1, z_k^2, ..., z_k^l, uma de cada vez usando o procedimento anterior l vezes e fazendo $F = I$ e $\Gamma = Q = 0$, ou seja, o modelo dinâmico é sempre o mesmo no instante t_k, embora o estado e

sua matriz covariância sejam filtrados recursivamente a partir das i observações.

A figura 2.3.2 mostra um diagrama que resume a sequência empregada no Filtro Kalman Padrão, com o particionamento do vetor das observações.

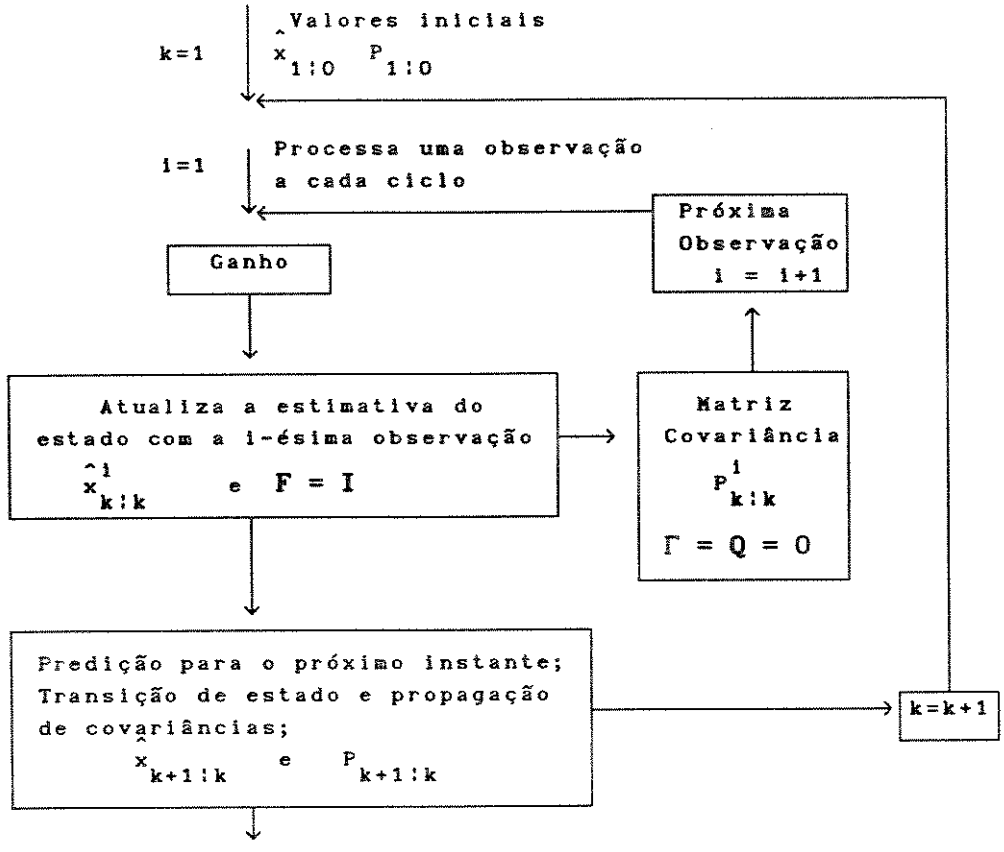


Figura 2.3.2 Diagrama para o Filtro Kalman Padrão com o particionamento do vetor das observações.

2.3.5 Filtro Kalman Estendido Iterativo - IEKF (Modelo não-linear)

Os procedimentos apresentados nos ítens anteriores são válidos para modelos lineares. Quando os modelos, especialmente o modelo de medida, forem não lineares, deve-se adaptar os procedimentos de filtragem, levando-se em conta a linearização do modelo e a consequente necessidade de iterações.

Jazwinski (1970) apresenta uma extensão do Filtro Kalman padrão para problemas não lineares, partindo de uma linearização por Taylor do modelo de medidas, mas sem realizar iterações. É o Filtro Kalman Estendido. Considerando este filtro, pode-se adotar um "iterador", ou seja, linearizar o modelo novamente, mas em relação à nova estimativa obtida com o Filtro. Este é o

Filtro Kalman Estendido Iterativo. As expressões (2.3.13), (2.3.14) mostram como estimar o vetor estado através de um processo iterativo, usando as medidas disponíveis em um instante t_k.

$$\hat{\eta}_{i+1} = \hat{x}_{k;k-1} + K_{k;\eta i}(z_k - h(\eta i, t_k) - M_{k;\eta i}[\hat{x}_{k;k-1} - \hat{\eta}_i]) \qquad (2.3.13)$$

onde **K**, o ganho Kalman, é dado por:

$$K_{k;\eta i} = P_{k;k-1} M^T_{k;\eta i} (M_{k;\eta i} P_{k;k-1} M^T_{k;\eta i} + R_k)^{-1} \qquad (2.3.14)$$

A matriz **M** é a matriz das derivadas parciais da função h em relação aos elementos do vetor estado, no ponto de linearização $\bar{x}_k$:

$$M_{k;\bar{x}_k} = \left[\frac{\partial h_i(\bar{x}_k, t_k)}{\partial x_j} \right] \qquad (2.3.15)$$

O vetor η_i é o iterador, isto é, uma estimativa para $\hat{x}_{k;k}$ na iteração atual. Na primeira iteração $\eta_1 = \hat{x}_{k;k-1}$, ou seja, lineariza-se o modelo no ponto representado pela estimativa predita. O resultado da primeira iteração η_2, obtido através da equação (2.3.13), representa uma estimativa do vetor estado com base na j-ésima observação no instante t_k. Esta estimativa do estado, melhor que a estimativa predita, é usada na segunda iteração, ou seja, o modelo é linearizado em η_2, e assim sucessivamente, até que haja convergência. Esta convergência é verificada testando dois valores consecutivos η_i e η_{i+1} e verificando se a diferença é menor que o critério pré-estabelecido. Com isto, o resultado da última iteração η_i é assumido como a estimativa do estado no instante t_k, $\hat{x}_{k;k}$, através da j-ésima observação.

Após a convergência das iterações pode-se calcular a matriz covariância, usando a expressão (2.3.16), análoga à (2.3.12).

$$P_{k;k} = (I - K_k M_k) P_{k;k-1} (I - K_k M_k)^T + K_k R_k K^T_k \qquad (2.3.16)$$

Quando o modelo do sistema também for não linear pode-se adotar um procedimento de linearização por Taylor. É importante ressaltar que o modelo do sistema é responsável pela transição do estado entre dois instantes, e de sua respectiva matriz covariância.

Com relação ao particionamento do vetor das observações, são válidas as

considerações apresentadas na seção 2.3.2, ou seja, quando as observações em um dado instante t_k forem não correlacionadas é possível calcular recursivamente o vetor estado neste instante, processando uma observação por vez. Isto significa que existirão três "loops" no processamento das observações. O "loop" interno refere-se às iterações em uma determinada observação; o segundo "loop" refere-se ao processamento de cada observação em um instante e o terceiro "loop" diz respeito ao processamento dinâmico, envolvendo o modelo do sistema. A figura 2.3.3 apresenta um diagrama representando este processo iterativo.

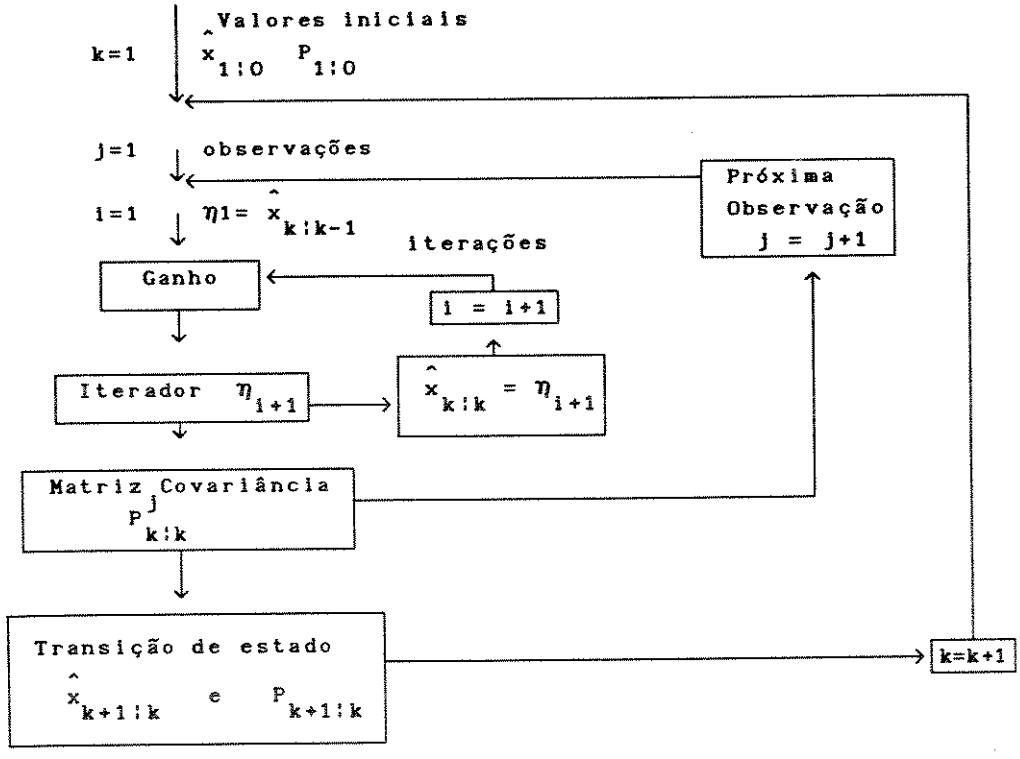


Figura 2.3.3 Diagrama do Filtro Kalman Estendido Iterativo

2.3.6 Estatística de resíduos e divergência.

A verificação da qualidade da estimativa, e mesmo a verificação de possíveis erros grosseiros nas observações, podem ser feitas analisando os resíduos preditos e sua matriz covariância, e verificando se estes satisfazem as propriedades estatísticas esperadas.

Os resíduos preditos, ou mais precisamente, como ressalta Jazwinski (1970), os erros residuais nas medidas preditas, podem ser obtidos através da equação (2.3.17):

$$r_{k:k-1} = (z_k - H_k \hat{x}_{k:k-1}) \qquad (2.3.17)$$

e sua matriz covariância por:

$$Y_{k:k-1} = (H_k P_{k:k-1} H_k^T + R_k) \qquad (2.3.18)$$

O resíduo predito padronizado pode, então, ser definido como o resíduo dividido pelo seu desvio padrão.

Uma maneira simples de avaliar o desempenho do filtro é verificar se os resíduos estão dentro de 1σ ou 2σ. Existem testes mais completos, como aqueles propostos por Teunissen e Salzmann (1989).

O fenômeno da divergência no filtro pode ser detectado analisando os resíduos. A divergência ocorre quando o erro no estado torna-se muito grande e a matriz covariância do erro no estado torna-se inesperadamente pequena. Nesta situação, os resíduos preditos também são muito maiores que seus desvios padrão, definidos pela matriz covariância dos resíduos (2.3.18).

Na realidade, as matrizes $P_{k:k-1}$ e $P_{k:k}$ são aproximações lineares da matriz covariância do erro no estado. Quando ocorre o fenômeno de divergência esta aproximação também converge para um valor incorreto.

Veress e Youcai (1987) estudaram uma maneira de calcular com maior exatidão a matriz covariância do erro no estado em métodos sequenciais, embora isto tenha implicado em um custo computacional adicional.

Além dos erros grosseiros de medida a avaliação incorreta da matriz covariância dos erros nas medidas, estimativas preditas tendenciosas, modelos de sistema inadequados, são causas de divergência do filtro.

2.4 RELAÇÕES ENTRE REFERENCIAIS

Uma aplicação típica de Visão Robótica em um ambiente industrial é a localização de objetos por meio de uma câmara montada sobre a garra do robô. Este objeto pode estar estático ou movendo-se sobre alguma superfície, normalmente uma esteira rolante. Na figura 2.4.1 apresenta-se um esquema desta situação, com a representação de todos os referenciais utilizados.

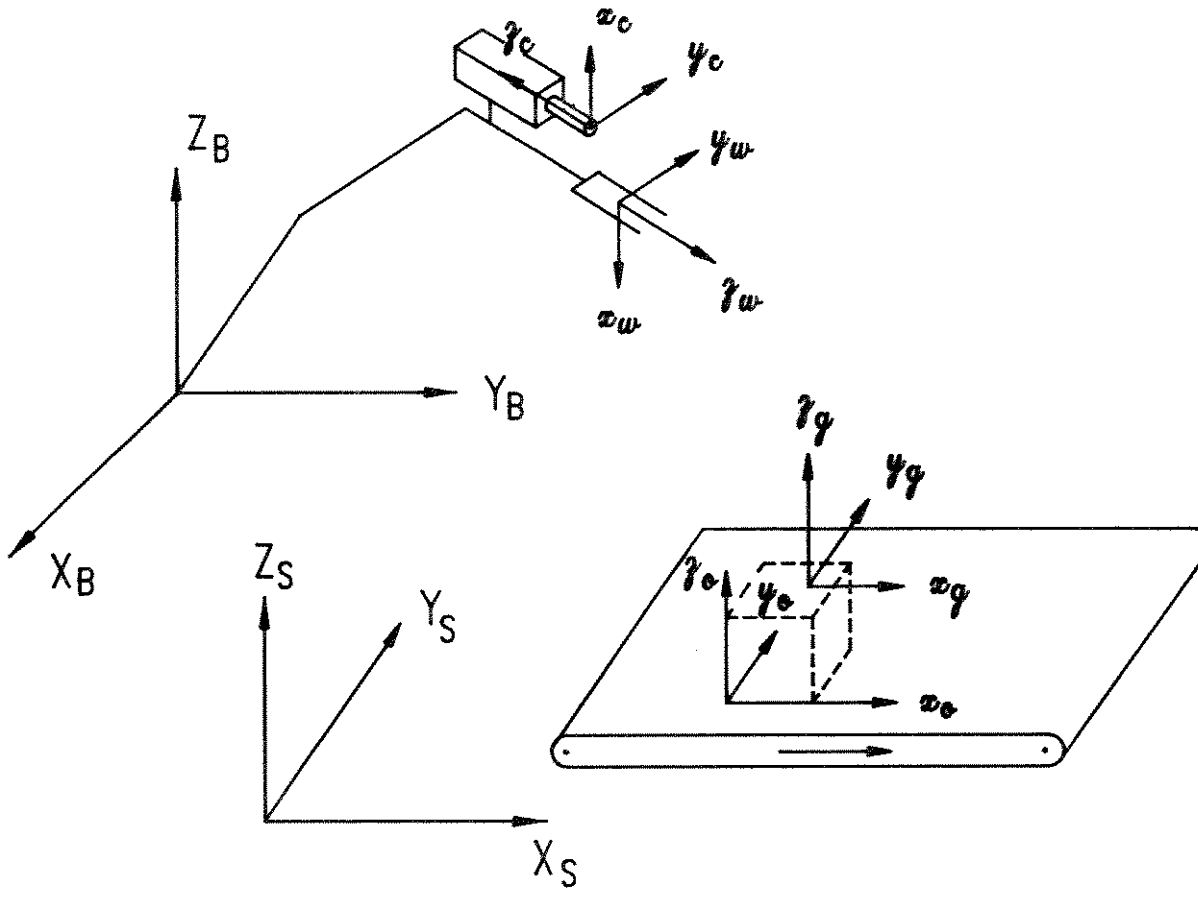


Figura 2.4.1 Um problema típico em Visão Robótica e os referenciais utilizados.

As seguintes transformações podem ser definidas, considerando-se as notações apresentadas por Paul (1981);

. ${}^{B}T_{S}$ descreve o referencial da esteira com relação ao referencial da base do manipulador;

. ${}^{S}T_{O}$ descreve a posição e orientação do referencial do objeto em relação ao referencial da esteira. O modelo do objeto é dado no referencial do objeto;

. ${}^{O}T_{G}$ descreve a posição e orientação com que a garra deve abordar o objeto. Este referencial deve ser especificado em relação ao mesmo referencial do modelo do objeto;

. ${}^{B}T_{W}$ descreve o referencial da garra em relação à base do manipulador. Esta transformação é conhecida como o modelo cinemático do robô e é obtida através de sucessivas transformações entre as juntas deste manipulador;

- ${}^{W}T_{C}$ descreve o referencial da câmara em relação ao referencial da garra. Obtem-se esta transformação através de um processo de calibração do conjunto câmara-robô (Tsai e Lenz, 1989);
- ${}^{C}T_{G}$ descreve o referencial ("goal frame") em que o objeto deve ser abordado em relação à câmara;

Supondo-se que os parâmetros intrínsecos da câmara tenham sido pré-calibrados, podem ser definidos, por meio de equações de transformação, os dois problemas básicos a serem discutidos neste trabalho.

No problema de **localização da câmara** supõe-se que as transformações ${}^{B}T_{S}$, ${}^{S}T_{O}$, ${}^{O}T_{G}$, ${}^{W}T_{C}$ são conhecidas a priori. Isto implica em afirmar que o objeto (ou outro conjunto de feições) foi posicionado com precisão em relação à base do manipulador. Pode-se calcular a transformação ${}^{C}T_{G}$ usando os métodos de calibração ou de localização e, então, calcular ${}^{B}T_{W}$ através da seguinte equação de transformação:

$$ {}^{B}T_{W} = {}^{W}T_{C} \cdot {}^{C}T_{G} \cdot {}^{O}T_{G}^{-1} \cdot {}^{S}T_{O}^{-1} \cdot {}^{B}T_{S}^{-1} \qquad (2.4.1) $$

Esta equação permite a estimação dos parâmetros do modelo cinemático do manipulador usando o método de localização da câmara. Esta estimativa pode ser usada para verificar a posição e orientação reais da garra.

No problema de **localização de objetos** supõe-se que as transformações ${}^{B}T_{S}$, ${}^{O}T_{G}$, ${}^{B}T_{W}$ e ${}^{W}T_{C}$ sejam conhecidas. A transformação ${}^{B}T_{W}$ é obtida usando o modelo cinemático do robô ou através de uma localização anterior. A transformação ${}^{C}T_{G}$ é calculada usando Visão, permitindo obter uma estimativa para a localização do objeto através da seguinte equação de transformação:

$$ {}^{S}T_{O} = {}^{O}T_{G} \cdot {}^{C}T_{G}^{-1} \cdot {}^{W}T_{C}^{-1} \cdot {}^{B}T_{W}^{-1} \cdot {}^{B}T_{S} \qquad (2.4.2) $$

Em cada instante t_k o objeto ocupará uma nova posição que pode ser predita conhecendo-se as equações que descrevem o movimento. Neste exemplo simples pode-se assumir o movimento da esteira em uma única direção, sendo que a velocidade é aproximadamente conhecida.

2.5 EXTRAÇÃO DE FEIÇÕES RETAS NA IMAGEM

2.5.1 DETECÇÃO DE BORDAS

O uso de linhas, e especificamente as retas, é possível ao adotar a premissa de que as arestas de objetos provocarão um efeito de descontinuidade na imagem. Normalmente, esta descontinuidade ocorre no nível de cinza, embora outras características, como por exemplo a textura, possam ser usadas para detectar tais descontinuidades.

A detecção de bordas recebe considerável tratamento na literatura, considerando-se que esta técnica é a base de uma grande parte dos procedimentos em Visão, como reconhecimento e localização.

Existem vários métodos para a detecção de descontinuidades no nível de cinza, que podem ser classificados em: métodos que usam aproximações do gradiente; métodos que usam "templates" em várias direções e; métodos que ajustam intensidades locais a modelos paramétricos (Ballard e Brown, 1982). Na primeira categoria estão os operadores de Roberts e de Sobel, que são ainda os mais usados, devido ao seu pequeno custo computacional.

Neste trabalho, devido à necessidade de respostas em tempo real, foi escolhido o detector de bordas baseado nos operadores de Sobel. A detecção de bordas, neste caso, consiste em aplicar os operadores a cada pixel na imagem e sua vizinhança, calculando-se, então, a magnitude do gradiente e sua direção. A comparação da magnitude do gradiente com um limiar pré-estabelecido permite definir as bordas em uma imagem.

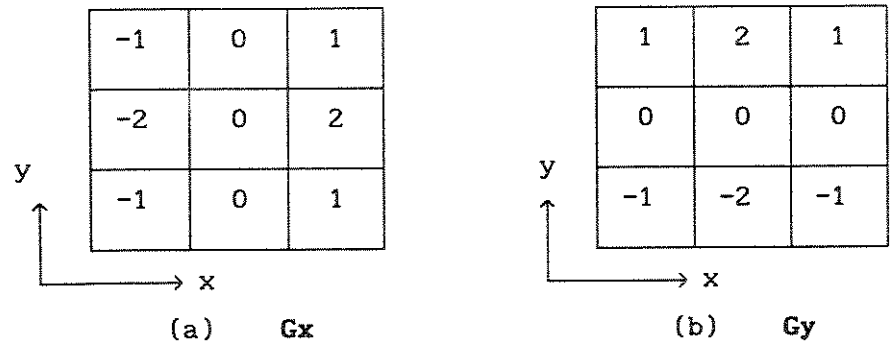


Figura 2.5.1 Operadores de Sobel;
(a) Gradiente em x;
(b) Gradiente em y;

Os operadores de Sobel são apresentados na figura 2.5.1., com uma ligeira

adaptação com relação ao sentido adotado para os eixos das ordenadas e abscissas.

A magnitude do gradiente é calculada através da expressão (2.5.1) e a direção do gradiente através da (2.5.2).

$$G(x,y) = [Gx^2 + Gy^2]^{1/2} \qquad (2.5.1)$$

$$\alpha(x,y) = \text{arctg}(Gy/Gx) \qquad (2.5.2)$$

A expressão (2.5.2) permite calcular a direção do gradiente no ponto central da janela 3x3. O sentido do gradiente é do nível de cinza mais escuro para o nível de cinza mais claro. Na implementação prática de um sistema de localização, esta informação pode ser usada como uma heurística no estabelecimento de correspondência entre as linhas retas. Neste caso, em algumas situações particulares regras adicionais devem ser estabelecidas.

Quando:

$$Gy = 0 \quad \begin{cases} Gx > 0 \Rightarrow \alpha(x,y) = 0; \\ Gx < 0 \Rightarrow \alpha(x,y) = 180; \end{cases}$$

$$Gx = 0 \quad \begin{cases} Gy > 0 \Rightarrow \alpha(x,y) = 90; \\ Gy < 0 \Rightarrow \alpha(x,y) = 270; \end{cases}$$

Para aumentar a precisão da informação direcional, pode-se aumentar o tamanho da janela no cálculo do ângulo α. Várias expressões existentes permitem calcular este ângulo a partir dos valores do gradiente em uma janela 3x3 (equivaleria a uma janela 5x5 na imagem). Kittler et al (1987) estudaram 3 métodos de realçamento da informação direcional, a partir dos valores de gradiente gerados pelos operadores de Sobel. Concluiram que o melhor método é aquele que despreza a contribuição do pixel central da janela. A direção do gradiente pode, então, ser calculada pela expressão (2.5.3).

$$\alpha(i,j) = \frac{\sum_{k=-1}^{1} \sum_{l=-1}^{1} Gy(i+k,\ j+l) - Gy(i,j)}{\sum_{k=-1}^{1} \sum_{l=-1}^{1} Gx(i+k,\ j+l) - Gx(i,j)} \qquad (2.5.3)$$

Na implementação desta técnica é importante reduzir o espaço de armazenamento da informação direcional; para isto, recorre-se a uma normalização, estabelecendo-se uma variação de 0 a 255 para o ângulo, o que possibilita a sua representação em 8 bits.

A informação direcional é extremamente importante, pois permite a aplicação de métodos de vetorização mais eficientes, e que baseiam-se nesta informação a priori. A precisão desta estimativa varia com uma série de fatores, mas os parâmetros dominantes são a amplitude da escala de cinza utilizada e a resolução espacial da imagem.

Em aplicações industriais é comum a existência de sistemas de Visão de Máquina que trabalham com imagens binárias. Neste caso, para evitar a geração de informação direcional grosseira, deve-se aumentar o tamanho da janela de cálculo do ângulo. À medida que a amplitude da escala de cinza utilizada aumenta, há um aumento proporcional na qualidade da informação direcional. Estas considerações são válidas para imagens focalizadas e com bordas nítidas. A desfocalização da imagem produziria efeitos negativos tanto sobre a definição das bordas, quanto sobre a qualidade da informação direcional.

O produto final desta etapa é uma imagem que associa a cada pixel uma magnitude e uma direção para o gradiente local. A moldura da imagem, isto é, a primeira e última colunas e a primeira e última linhas, não é considerada. Normalmente, atribuem-se valores nulos para o gradiente nesta porção da imagem, ou, então, repetem-se os valores dos pixels vizinhos.

2.5.2 Vetorização.

A imagem de bordas obtida com o uso dos operadores de Sobel representa elementos de descontinuidade, porém, desconectados. O objetivo da vetorização é conectar os elementos de borda, transformando vários elementos individuais em uma feição contínua. Vários métodos vem sendo propostos para esta finalidade, variando em grau de complexidade e fazendo maior ou menor uso de conhecimento a priori para realizar a tarefa. Algumas limitações para o caso concreto de Visão de Máquina são relevantes, especialmente a necessidade de tempo real. Métodos baseados em sucessivas iterações sobre a imagem, com o objetivo de realizar a eliminação de ruídos ou o afinamento tornam-se indesejáveis.

Uma técnica muito utilizada para a segmentação de linhas retas é a

Transformada de Hough. A transformada de Hough procura transformar pontos de borda isolados em uma representação paramétrica no espaço de Hough.

Na figura 2.5.2 (a) a reta $y = a.x + b$ é definida por uma sucessão de pontos, que são pixels isolados, representando pontos de máximo local do gradiente. Reescrevendo a equação da reta como $b = y - ax$, define-se um espaço que associa, a cada ponto no espaço xy, uma reta no espaço ab (figura 2.5.2 (b)). Isto significa que, uma reta no espaço xy, gerará uma família de retas no espaço ab, e que o ponto de intersecção destas retas no espaço Hough determinará os parâmetros da reta original na imagem.

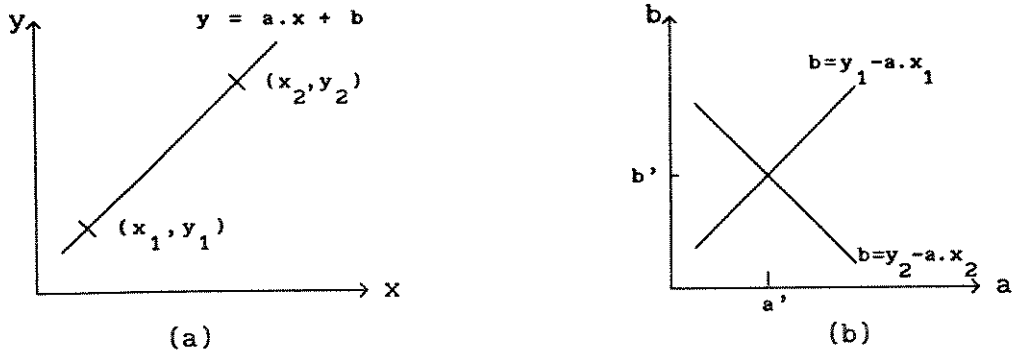


Figura 2.5.2 Transformada de Hough:
(a) espaço original xy;
(b) espaço ab - Hough;

Para implementar computacionalmente este procedimento, basta dividir o espaço paramétrico em células discretas de acumulação. A sequência de cálculos pode ser resumida da seguinte maneira (Ballard e Brown, 1982):

1. definir o espaço de acumulação, através dos intervalos possíveis de variação de a (a_{max} e a_{min}) e b (b_{max} e b_{min}), bem como o número de células, o que definirá a resolução dos parâmetros;
2. formar uma matriz de acumuladores e inicializá-la com valores nulos;
3. para cada ponto na imagem de bordas que esteja acima de um determinado limiar, incrementar todos os pontos no acumulador ao longo de uma linha. Isto significa variar **a** e calcular **b** de acordo com a equação $b = y - a.x$. Ao final deste passo todos as células pertencentes a esta linha foram incrementadas em uma unidade;
4. Realizar este processo de acumulação para toda a imagem de gradiente. Buscar, então, pelos pontos de máximo no espaço paramétrico, que corresponderão às retas esperadas. O valor armazenado no acumulador indicará o número de pixels contidos na reta;

Esta abordagem, entretanto, apresenta problemas quando a reta for vertical, pois tanto **a** quanto **b** tendem ao infinito. Para contornar este problema basta recorrer a uma parametrização diferente para a reta. A equação (2.5.4) apresenta a equação normal da reta, que é usada numa abordagem mais genérica. Ao usar esta equação para descrever a reta, adota-se um espaço do tipo θ-ρ, e as linhas descritas neste espaço são agora senóides, ao invés de retas.

$$\rho = x.\cos\theta + y.\mathrm{sen}\theta \qquad (2.5.4)$$

Embora a transformada de Hough seja bastante simples do ponto de vista teórico, sua implementação merece uma série de estratégias adicionais. O custo computacional da transformada de Hough é muito alto, especialmente quando deseja-se uma alta precisão na definição dos parâmetros da reta, pois é necessário aumentar proporcionalmente o número de células. Além disto, a busca pelos pontos de máximo requer uma exaustiva iteração sobre o espaço paramétrico, a fim de evitar pontos de máximo falsos, gerados por intersecções que não correspondam às retas reais. Outro problema com a transformada de Hough é a possibilidade de perda de segmentos de bordas curtos.

Uma maneira de reduzir o custo computacional da transformada de Hough é utilizar a informação direcional, gerada na detecção de bordas. Esta informação pode ser usada para reduzir o espaço de acumulação para um determinado ponto, somente no entorno do ângulo definido pelo gradiente.

Outra maneira de extrair linhas de maneira eficiente, proposta por Dudani e Luk (1978), baseia-se na classificação dos pixels de acordo com sua informação direcional. Neste processo, o espaço paramétrico é definido sem as células de acumulação mencionadas anteriormente.

Para cada pixel, além da informação posicional (x,y) e da magnitude do gradiente, é possível obter a informação direcional local que permite calcular um ρ aproximado, podendo-se definir, então, uma lista de elementos de borda, onde cada pixel possui os elementos: (x, y, θ, ρ).

O primeiro passo no processo é determinar a distribuição da informação direcional. Isto é feito calculando-se um histograma para os ângulos θ obtidos durante a detecção de bordas. Realiza-se, então, uma determinação de máximos locais neste histograma, para estabelecer as classes θ. Para cada classe θ obtém-se, então, o histograma em ρ, permitindo a definição das classes ρ.

Pode-se, então, agrupar os elementos de borda de acordo com a informação

direcional e, em seguida, para cada grupo direcional, reagrupar os elementos de acordo com o ρ. Finalmente, agrupar em cada classe θ-ρ os elementos contíguos, de modo que, no final do processo, todos os pixels pertencentes a uma mesma linha estejam agrupados em uma classe.

Uma vez agrupados todos os pixels correspondentes a uma mesma classe, ajusta-se uma reta por estes pixels usando o Método dos Mínimos Quadrados, obtendo uma representação altamente precisa para a reta. Os passos para este processo são resumidos na figura 2.5.3.

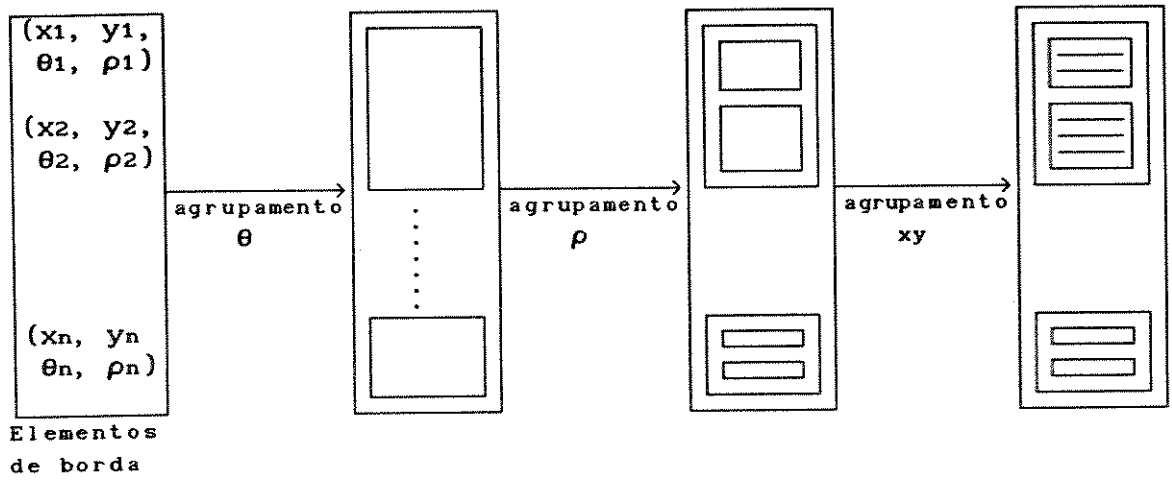


Figura 2.5.3 Passos no método de agrupamento θ-ρ

Este método é extremamente eficiente do ponto de vista computacional, além de permitir obter uma descrição das retas com alta precisão, devido ao ajuste final da reta pelo Método dos Mínimos Quadrados. Entretanto, está sujeito a uma pré-filtragem da imagem, para eliminar elementos de borda espúrios, além de requerer uma informação direcional de boa qualidade.

Para imagens tomadas com iluminação controlada é possível usar diretamente a imagem de bordas, obtida usando os operadores de Sobel, dispensando a filtragem inicial. Para reduzir o número de pixels a serem agrupados em uma determinada classe adota-se um procedimento de selecionar, para os pixels em uma mesma linha de varredura ("scan line") e de uma mesma classe, apenas o pixel que apresentar a maior magnitude do gradiente. Isto permite armazenar somente os pixels de máximo gradiente da reta.

Outra hipótese, que permitiria maior precisão, embora com demanda computacional adicional, seria usar todos os elementos de bordas, mas adotando um peso proporcional à magnitude do gradiente, no ajuste da reta.

Um dos problemas neste método é a classificação dos θ e ρ, devido ao erro

na informação direcional. Em imagens muito complexas, com retas muito próximas, podem ocorrer acumulações de retas distintas na mesma classe θ-ρ, gerando erros grosseiros; um passo adicional para a separação destas retas deve ser empregado.

CAPÍTULO III

3. MODELO MATEMÁTICO E PROCESSO RECURSIVO

No Capítulo 2 foram revisados alguns dos métodos existentes de calibração e localização. Em particular, mostrou-se que os métodos usando linhas retas apresentam uma série de vantagens, que os tornam adequados para as aplicações propostas em Visão de Máquina. Além disto, mostrou-se a conveniência da utilização de métodos de Filtragem para a estimação dos parâmetros de localização.

Uma vantagem imediata do uso de técnicas de estimação de parâmetros sobre os métodos analíticos é que os erros de medida e os erros no estado sofrem um tratamento adequado, possibilitando, ainda, uma avaliação da qualidade dos valores estimados.

Outra vantagem do uso de métodos de Filtragem, particularmente os métodos recursivos, é que a cada medida efetuada é possível obter uma estimativa do vetor estado. Com isto, é possível realimentar o processo de extração de feições, através da definição de um polígono de busca na imagem, que contenha a reta a ser detectada. Com isto, cada reta medida na imagem melhora a qualidade da estimativa filtrada e contribui para a redução do espaço de busca da próxima feição. Como já foi discutido, a etapa de extração de feições é que demanda maior tempo de processamento, em comparação com as demais fases. Embora a utilização de técnicas de filtragem implique em aumentar o tempo de processamento na etapa de cálculo dos parâmetros de localização, é esperada uma diminuição mais acentuada do tempo de processamento da etapa de extração de feições.

Entretanto, os modelos matemáticos existentes para o tratamento de linhas

retas são do tipo implícito, o que exige aproximações para a aplicação da Filtragem Kalman, que podem levar à divergência do filtro. Outro problema com alguns dos métodos existentes, notadamente o de Liu e Huang é o efeito de propagação de erros do primeiro para o segundo estágio. Faz-se necessário, portanto, o desenvolvimento de um modelo matemático do tipo explícito e simultâneo usando linhas retas, permitindo o emprego da Filtragem Kalman.

Este capítulo divide-se em três seções:

- na **seção 3.1** apresenta-se o desenvolvimento do **modelo matemático simultâneo** para o tratamento de linhas retas; este modelo é do tipo explícito, adequado para o tratamento por filtragem;
- na **seção 3.2** desenvolve-se a aplicação da **Filtragem Kalman Estendida** no modelo simultâneo;
- na **seção 3.3** é descrita a proposta do **processo recursivo** entre a etapa de estimação de parâmetros e a de extração de feições;

3.1 MODELO MATEMÁTICO

3.1.1 RELAÇÕES ENTRE AS REPRESENTAÇÕES NORMAL E PARAMÉTRICA DA RETA NO PLANO.

As retas a serem utilizadas no modelo matemático simultâneo a ser desenvolvido, serão extraídas da imagem e representadas por suas equações no plano imagem.

A equação de uma reta, representada por dois de seus pontos, p_1 e p_2, é dada por:

$$(y_2 - y_1).x + (x_1 - x_2).y + (x_2.y_1 - x_1.y_2) = 0 \qquad (3.1.1)$$

Seja d a distância entre os pontos p_1 e p_2; dividindo a equação (3.1.1) por d, obtém-se:

$$\frac{(y_2 - y_1)}{d}.x + \frac{(x_1 - x_2)}{d}.y + \frac{(x_2.y_1 - x_1.y_2)}{d} = 0 \qquad (3.1.2)$$

Considerando-se que (figura 3.1.1):

$$\cos\theta = \frac{(y_2 - y_1)}{d} \qquad e \qquad \text{sen}\theta = \frac{(x_1 - x_2)}{d} \tag{3.1.3}$$

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{3.1.4}$$

$$\rho = -(x_2 . y_1 - x_1 . y_2)/d \tag{3.1.5}$$

obtém-se a equação da reta em sua forma normal:

$$\cos\theta . x + \text{sen}\theta . y - \rho = 0 \tag{3.1.6}$$

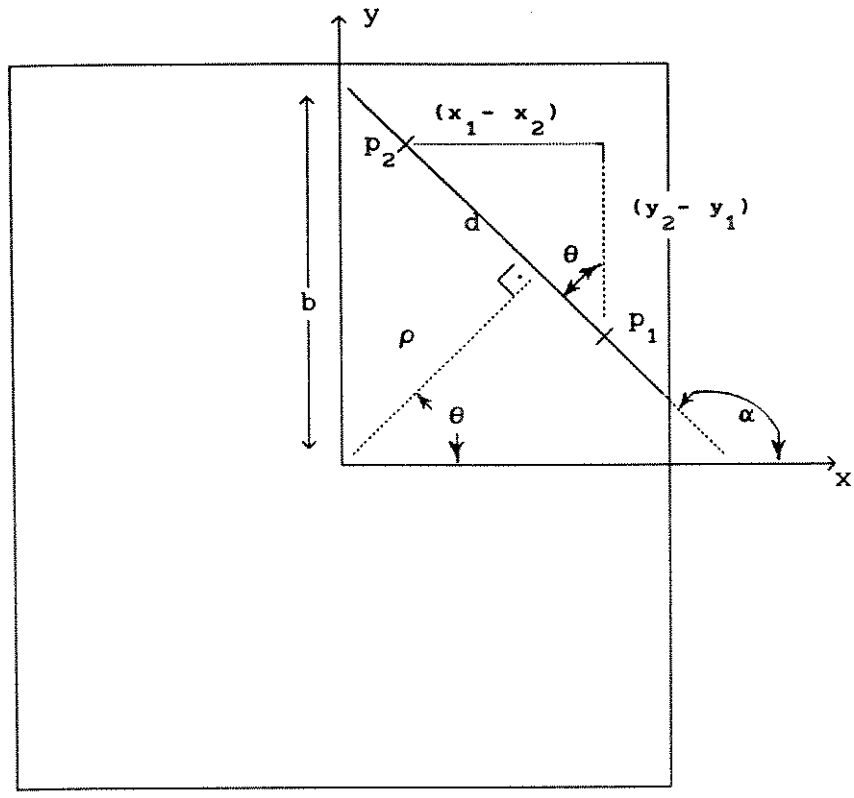


Figura 3.1.1 Elementos das representações normal e paramétrica da reta.

A equação paramétrica da reta no plano é dada por:

$$y = a . x + b \tag{3.1.7}$$

onde:

$a = \text{tg}\ \alpha$ (figura 3.1.1)

b é a ordenada do ponto de intersecção entre a reta e o eixo y (figura 3.1.1);

É possível, portanto, estabelecer relações entre a equação paramétrica e a equação normal da reta, uma vez que $\alpha = \theta + 90^{\circ}$ (figura 3.1.1):

$$tg\ \theta = -cotg\ \alpha \qquad e \qquad tg\ \alpha = -cotg\ \theta \tag{3.1.8}$$

A equação paramétrica pode ser escrita, então, como:

$$y = -cotg\theta . x + \rho/sen\theta \tag{3.1.9}$$

portanto:

$$a = -cotg\theta = tg\alpha \tag{3.1.10}$$

$$b = \rho/sen\theta = -\rho/cos\alpha \tag{3.1.11}$$

Esta representação paramétrica não é definida quando $\theta \cong 0$, ou seja, para retas aproximadamente verticais. Para estes casos é necessário estabelecer uma nova parametrização para a reta na imagem. Seja a equação normal da reta:

$$cos\theta . x + sen\theta . y - \rho = 0$$

Dividindo esta equação por $cos\theta$, obtém-se:

$$x = -tg\theta . y + \rho/cos\theta \tag{3.1.12}$$

Com esta expressão é possível descrever retas aproximadamente verticais, o que não seria possível com a representação anterior. Neste caso:

$$x = a^{*} . y + b^{*} \tag{3.1.13}$$

com:

$$a^{*} = -tg\theta = cotg\alpha \tag{3.1.14}$$

$$b^{*} = \rho/cos\theta = -\rho/sen\alpha \tag{3.1.15}$$

3.1.2 Plano de interpretação.

O plano de interpretação é definido como sendo o plano que contem a reta no espaço objeto, a reta projetada no espaço imagem e o centro perspectivo da

câmara.

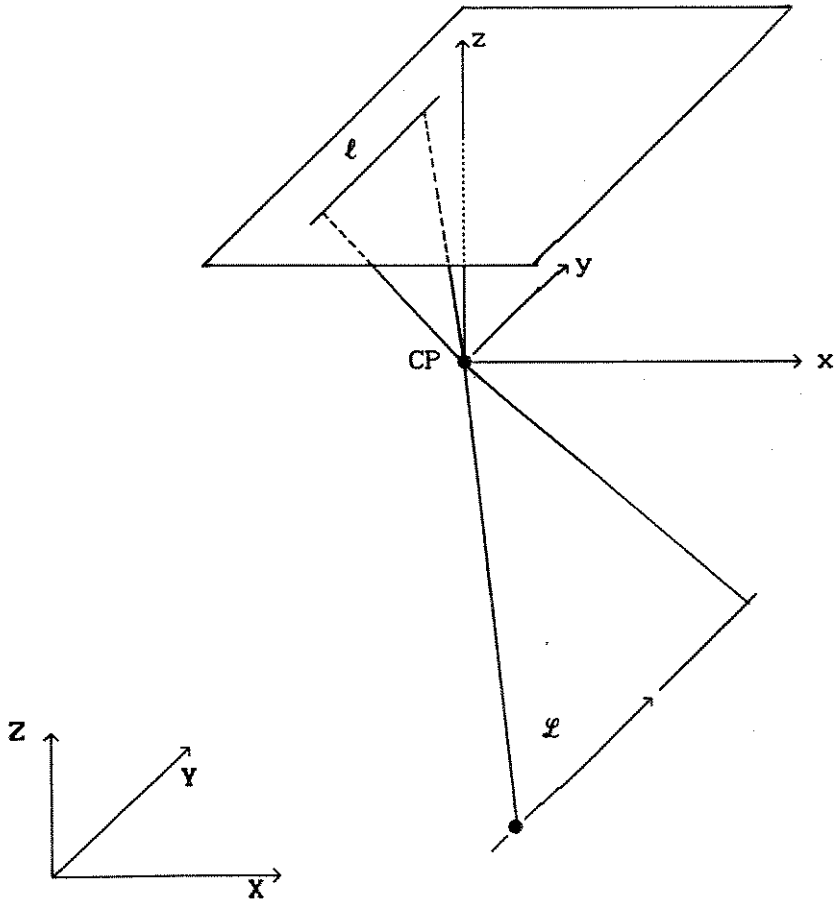


Figura 3.1.2 Plano de interpretação.

3.1.2.1 PLANO DE INTERPRETAÇÃO DESCRITO NO ESPAÇO IMAGEM.

O plano de interpretação é definido, no espaço imagem, pela linha reta no plano imagem e pelo Centro Perspectivo, que é uma abstração pontual do sistema de lentes. O mesmo plano pode ser definido no espaço objeto, por uma reta 3D e pelo Centro Perspectivo, ambos no referencial do espaço objeto (figura 3.1.2). Esta abstração geométrica, entretanto, só é válida se a imagem estiver isenta de erros sistemáticos como, por exemplo, a distorção ótica e os erros de deslocamento do centro da imagem.

Segundo Barnard (1983) este *plano de interpretação* pode ser representado a partir de dois pontos quaisquer pertencentes à reta no plano imagem.

Considerando $\vec{p}_1 = (x_1, y_1, f)$ e $\vec{p}_2 = (x_2, y_2, f)$, onde x_1, y_1 são as coordenadas de dois pontos na imagem e f é a distância focal da câmara, o vetor normal unitário ao plano de interpretação será dado por:

$$\Phi = \frac{\vec{p}_1 \times \vec{p}_2}{|\vec{p}_1||\vec{p}_2|} = (\phi_x, \phi_y, \phi_z)^T \qquad (3.1.16)$$

O vetor normal Φ (3.1.16) define uma família de planos. O plano de interpretação, em particular, é o plano que passa pela origem do referencial imagem e, portanto, possui o termo independente nulo.

A forma apresentada em (3.1.16), contudo, utiliza as coordenadas de dois pontos. Em Visão, é mais comum a representação da reta no espaço imagem através de sua equação normal, tendo como parâmetros θ e ρ, ou da equação paramétrica, com os parâmetros a e b ($y = a.x + b$). Estes parâmetros (θ-ρ ou a-b) são obtidos por algum processo de vetorização.

As componentes do vetor normal não normalizado serão dadas por:

$$\vec{N} = \vec{p}_1 \times \vec{p}_2 = \begin{bmatrix} -f.(y_2 - y_1) \\ -f.(x_1 - x_2) \\ -x_2.y_1 + x_1.y_2 \end{bmatrix} \qquad (3.1.17)$$

A equação do plano de interpretação será, então:

$$-f.(y_2 - y_1).x - f.(x_1 - x_2).y - (x_2.y_1 - x_1.y_2).z = 0 \qquad (3.1.18)$$

É possível encontrar a equação do plano de interpretação em função dos parâmetros da reta, manipulando a equação (3.1.18) e considerando as relações definidas na seção 3.1.1.

Dividindo a (3.1.18) pela distância d entre dois pontos na imagem (equação 3.1.4):

$$-\frac{f.(y_2 - y_1)}{d}.x - \frac{f.(x_1 - x_2)}{d}.y - \frac{(x_2.y_1 - x_1.y_2)}{d}.z = 0 \qquad (3.1.19)$$

Considerando as relações definidas nas equações (3.1.3), verifica-se que a equação do plano de interpretação pode ser escrita em função de θ e ρ como:

$$-f.\cos\theta.x - f.\mathrm{sen}\theta.y + \rho.z = 0 \qquad (3.1.20)$$

Esta é a equação do plano de interpretação em função da equação normal da reta no plano imagem (z = f).

O vetor normal a este plano, não unitário, é dado por:

$$\vec{N} = -\begin{bmatrix} f.\cos\theta \\ f.\mathrm{sen}\theta \\ -\rho \end{bmatrix} \qquad (3.1.21)$$

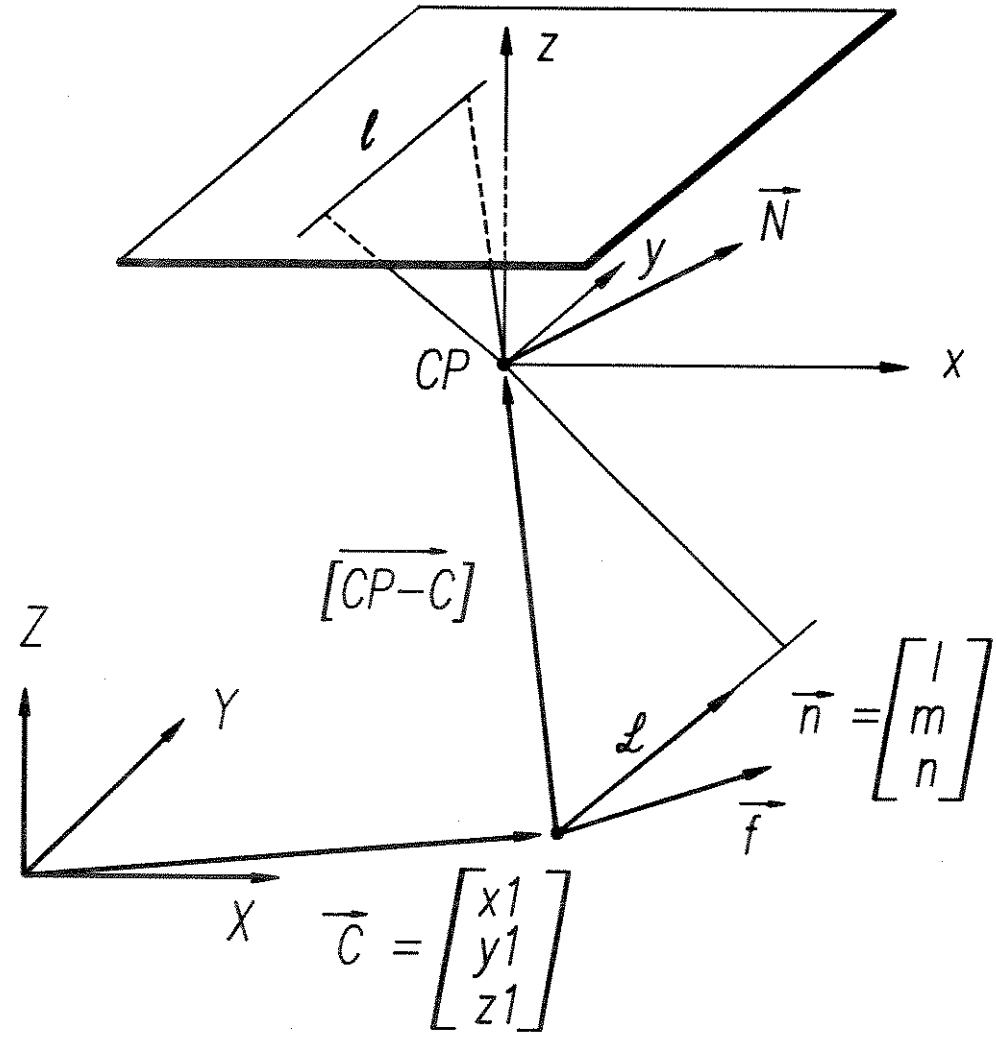


Figura 3.1.3 Os vetores normais aos planos de interpretação no espaço imagem e espaço objeto

3.1.2.2 PLANO DE INTERPRETAÇÃO DESCRITO NO ESPAÇO OBJETO.

Seja uma reta $\mathcal{L}$ no espaço objeto, cuja representação paramétrica é dada por:

$$\mathcal{L}: \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} + t \begin{bmatrix} l \\ m \\ n \end{bmatrix} \qquad (3.1.22)$$

onde X_1, Y_1, Z_1 é um ponto qualquer na reta, l, m, n são os cossenos diretores desta reta e t é um parâmetro livre.

Seja o vetor $\vec{f}$, normal ao plano de interpretação no espaço objeto, definido pelo produto vetorial entre o vetor $\vec{n}$, dos cossenos diretores da reta $\mathcal{L}$, e o vetor diferença $(\overrightarrow{CP - C})$, como pode ser visto na figura 3.1.3.

O vetor $\vec{f}$ é definido através do produto vetorial:

$$\vec{f} = \vec{n} \times [\overrightarrow{CP - C}] = \begin{bmatrix} l \\ m \\ n \end{bmatrix} \times \begin{bmatrix} X_c - X_1 \\ Y_c - Y_1 \\ Z_c - Z_1 \end{bmatrix} \qquad (3.1.23)$$

ou:

$$\begin{bmatrix} f_x \\ f_y \\ f_z \end{bmatrix} = \begin{bmatrix} -n.(Y_c - Y_1) + m.(Z_c - Z_1) \\ n.(X_c - X_1) - l.(Z_c - Z_1) \\ -m.(X_c - X_1) + l.(Y_c - Y_1) \end{bmatrix} \qquad (3.1.24)$$

3.1.3 MODELO FUNCIONAL

Usando o conceito de plano de interpretação será desenvolvido um modelo matemático do tipo explícito e que permita a determinação simultânea dos 6 parâmetros extrínsecos.

Rotacionando o referencial objeto, são compensadas as diferenças angulares entre os referenciais imagem e objeto e, portanto, o vetor normal $\vec{N}$ ao plano de interpretação no espaço imagem e o vetor normal $\vec{f}$ ao plano de interpretação no espaço objeto tornam-se equivalentes. Isto significa dizer que estes vetores possuem mesma direção, porém, normas diferentes. Isto ocorre porquê $\vec{N}$ e $\vec{f}$ não foram normalizados previamente.

Pré-multiplicando o vetor $\vec{f}$ pela matriz de rotação R, com o objetivo de

compensar as diferenças angulares entre o referencial objeto e o referencial imagem, observa-se que seus componentes serão múltiplos escalares do vetor $\vec{N}$. Definindo um fator de escala λ, entre estes vetores, pode-se descrever a relação acima, analiticamente, na forma:

$$\vec{N} = \lambda \, R \, \vec{f} \tag{3.1.25}$$

onde **R** é a matriz de rotação supondo-se a sequência de rotações $Mz(\kappa)\ My(\phi)\ Mx(\omega)$ (Apêndice 1).

Desenvolvendo os termos, e multiplicando-os por (-1), obtem-se:

$$\begin{bmatrix} f.\cos\theta \\ f.\mathrm{sen}\theta \\ -\rho \end{bmatrix} = -\lambda \begin{bmatrix} fx' \\ fy' \\ fz' \end{bmatrix} \tag{3.1.26}$$

onde fx', fy' e fz' são as componentes do vetor normal ao plano de interpretação no espaço objeto, rotacionado. Os elementos deste vetor são:

$$\begin{bmatrix} fx' \\ fy' \\ fz' \end{bmatrix} = \begin{bmatrix} r11.fx + r12.fy + r13.fz \\ r21.fx + r22.fy + r23.fz \\ r31.fx + r32.fy + r33.fz \end{bmatrix} \tag{3.1.27}$$

onde os r_{ij} são os elementos da matriz de rotação.

Para obter os parâmetros extrínsecos, a partir das equações (3.1.26), deve-se considerar a introdução da incógnita λ para cada linha reta medida. O cálculo de λ como incógnita é incoveniente, pois introduz uma variável desnecessária, aumentando a ordem do sistema a ser resolvido.

Existem duas maneiras de contornar este problema: a **primeira** é encontrar uma expressão analítica para λ e utilizá-la nas equações (3.1.26). Para isto deve-se considerar a normalização dos vetores normais tanto no espaço imagem quanto no espaço objeto. Se os vetores normais fossem unitários, não haveria a necessidade do fator de escala λ. Portanto, normalizando $\vec{N}$ e $\vec{f}$, o modelo poderia ser reescrito como:

$$\frac{1}{\|N\|} \vec{N} = \frac{1}{\|f\|} R \, \vec{f} \tag{3.1.28}$$

Contudo, uma das equações, das três possíveis de obter a partir da expressão (3.1.28), deve ser eliminada, pois a normalização cria uma dependência linear entre estas equações. O termo λ pode ser dado, portanto, por:

$$\lambda = \frac{\|N\|}{\|f\|} \qquad (3.1.29)$$

onde: $\|\vec{N}\| = (f^2 + \rho^2)^{1/2}$ e $\|\vec{f}\| = (f_x^2 + f_y^2 + f_z^2)^{1/2}$ (3.1.30)

Esta alternativa, além da questão de dependência linear, implica em tornar o modelo mais complexo, o que dificultaria sua linearização

A **segunda** alternativa, mais simples, que elimina a necessidade de normalizações, baseia-se na manipulação algébrica das equações (3.1.26). Para eliminar λ, basta dividir as equações entre si. Dividindo a 1ª e a 3ª pela 2ª, obtém-se:

$$\begin{aligned} \cotg\theta &= \frac{r11.fx + r12.fy + r13.fz}{r21.fx + r22.fy + r23.fz} = \frac{fx'}{fy'} \\ \frac{-\rho}{f.\mathrm{sen}\theta} &= \frac{r31.fx + r32.fy + r33.fz}{r21.fx + r22.fy + r23.fz} = \frac{fz'}{fy'} \end{aligned} \qquad (3.1.31)$$

Contudo, das expressões (3.1.10) e (3.1.11), sabe-se que:

$$\cotg\theta = -a \qquad \text{e} \qquad -\rho/\mathrm{sen}\theta = -b \qquad (3.1.32)$$

As equações (3.1.31) podem, então, ser reescritas como:

$$\begin{aligned} a &= -\frac{r11.fx + r12.fy + r13.fz}{r21.fx + r22.fy + r23.fz} = -\frac{fx'}{fy'} \\ b &= -f.\frac{r31.fx + r32.fy + r33.fz}{r21.fx + r22.fy + r23.fz} = -f.\frac{fz'}{fy'} \end{aligned} \qquad (3.1.33)$$

Existe um problema básico ao utilizar esta formulação. Quando $\theta \cong 0^\circ$ (ou $\alpha \cong 90^\circ$) a e b se tornam muito grandes (ou $\to \infty$) inviabilizando a formulação. É estabelecido a priori que as equações (3.1.31) ou (3.1.33) serão empregadas quando $45^\circ < \theta < 135^\circ$ ou $225^\circ < \theta < 315^\circ$. Para contornar este problema basta desenvolver um novo conjunto de equações.

Para obter um outro conjunto de equações, válidos para $315^\circ < \theta < 45^\circ$ ou

$135^{\circ} < \theta < 225^{\circ}$, manipula-se as equações (3.1.26) de maneira diferente. Dividindo a 2ª e a 3ª equação pela 1ª, obtém-se:

$$\begin{aligned} tg\theta &= \frac{r21.fx + r22.fy + r23.fz}{r11.fx + r12.fy + r13.fz} = \frac{fy'}{fx'} \\ \frac{-\rho}{f.\cos\theta} &= \frac{r31.fx + r32.fy + r33.fz}{r11.fx + r12.fy + r13.fz} = \frac{fz'}{fx'} \end{aligned} \qquad (3.1.34)$$

Adotando a parametrização a^*-b^* e considerando as relações (3.1.14) e (3.1.15), obtém-se:

$$\begin{aligned} a^* &= -\frac{r21.fx + r22.fy + r23.fz}{r11.fx + r12.fy + r13.fz} = -\frac{fy'}{fx'} \\ b^* &= -f.\frac{r31.fx + r32.fy + r33.fz}{r11.fx + r12.fy + r13.fz} = -f.\frac{fz'}{fx'} \end{aligned} \qquad (3.1.35)$$

As equações (3.1.33) e (3.1.35) são equações do tipo projetiva, originadas pela eliminação do fator de escala. Ao aplicar estas equações é preservada a coplanaridade entre a reta no espaço objeto, a reta no espaço imagem e o centro perspectivo.

O modelo representado pelas equações (3.1.33) e (3.1.35) é bastante compacto e merece algumas considerações adicionais. A primeira observação é a capacidade de relacionar funcionalmente, de forma explícita, os parâmetros extrínsecos com os parâmetros a e b (ou a^* e b^*), da reta na imagem. Estes valores são medidos na imagem através de algum processo de vetorização. Esta representação explícita permite a aplicação da Filtragem Kalman, tornando o modelo adequado a aplicações dinâmicas. Não foi encontrado na literatura modelo semelhante. A segunda observação é que os elementos a e b são obtidos como uma razão entre os elementos do vetor $\vec{f}$, normal ao plano de interpretação no espaço objeto. A terceira observação é que a primeira equação não é função da focal, o que indica que o ângulo θ da reta não depende deste parâmetro, como pode ser notado intuitivamente.

Neste modelo simultâneo os parâmetros extrínsecos serão determinados em um único estágio, a partir de um dos modelos apresentados e de um conjunto de retas, cujas equações paramétricas no espaço sejam conhecidas e suas correspondentes na imagem tenham sido medidas.

3.2 FILTRAGEM KALMAN APLICADA AO MODELO SIMULTÂNEO

O modelo simultâneo, desenvolvido na seção 3.1, é do tipo explícito e não linear. A aplicação da Filtragem Kalman, neste caso, requer uma linearização usando uma aproximação linear por série de Taylor. Torna-se necessário o desenvolvimento de derivadas parciais de primeira ordem, das equações do modelo em relação aos elementos do vetor estado.

Para tornar este desenvolvimento mais conciso, as seguintes notações serão consideradas. A matriz de rotação (Apêndice 1) será definida através de três matrizes linha:

$$\begin{aligned} \mathbf{R}_1 &= [\ r_{11} \quad r_{12} \quad r_{13}\ ] \\ \mathbf{R}_2 &= [\ r_{21} \quad r_{22} \quad r_{23}\ ] \\ \mathbf{R}_3 &= [\ r_{31} \quad r_{32} \quad r_{33}\ ] \end{aligned} \tag{3.2.1}$$

Com esta notação pode-se reescrever as equações (3.1.33) e (3.1.35), como:

$$\begin{aligned} a &= -\frac{r_{11}.f_x + r_{12}.f_y + r_{13}.f_z}{r_{21}.f_x + r_{22}.f_y + r_{23}.f_z} = -\frac{\mathbf{R}_1.\mathbf{f}}{\mathbf{R}_2.\mathbf{f}} \\ b &= -f.\frac{r_{31}.f_x + r_{32}.f_y + r_{33}.f_z}{r_{21}.f_x + r_{22}.f_y + r_{23}.f_z} = -f.\frac{\mathbf{R}_3.\mathbf{f}}{\mathbf{R}_2.\mathbf{f}} \end{aligned} \tag{3.2.2}$$

$$\begin{aligned} a^* &= -\frac{r_{21}.f_x + r_{22}.f_y + r_{23}.f_z}{r_{11}.f_x + r_{12}.f_y + r_{13}.f_z} = -\frac{\mathbf{R}_2.\mathbf{f}}{\mathbf{R}_1.\mathbf{f}} \\ b^* &= -f.\frac{r_{31}.f_x + r_{32}.f_y + r_{33}.f_z}{r_{11}.f_x + r_{12}.f_y + r_{13}.f_z} = -f.\frac{\mathbf{R}_3.\mathbf{f}}{\mathbf{R}_1.\mathbf{f}} \end{aligned} \tag{3.2.3}$$

onde: f é a distância focal da câmara e **f** é o vetor normal ao plano de interpretação no espaço objeto.

Adotando uma nomenclatura mais concisa, e introduzindo os elementos u, v e w, pode-se, ainda, reescrever as equações anteriores:

$$\begin{aligned} a &= G_1(\mathbf{x}_k) = -\frac{\mathbf{R}_1.\mathbf{f}}{\mathbf{R}_2.\mathbf{f}} = -\frac{v}{u} \\ b &= G_2(\mathbf{x}_k) = -f.\frac{\mathbf{R}_3.\mathbf{f}}{\mathbf{R}_2.\mathbf{f}} = -f.\frac{w}{u} \end{aligned} \tag{3.2.4}$$

$$a^* = F_1(x_k) = -\frac{R_2.f}{R_1.f} = -\frac{u}{v}$$
$$b^* = F_2(x_k) \quad -f.\frac{R_3.f}{R_1.f} = -f.\frac{w}{v} \qquad (3.2.5)$$

onde:

a e b (a^* e b^*) são as observações (medidas);

x_k é o vetor estado (os parâmetros extrínsecos);

F_1 e F_2 (G_1 e G_2) são as funções do modelo matemático.

O vetor estado no instante t_k, será constituido por:

$$x_k = [\ \kappa,\ \phi,\ \omega,\ Xc,\ Yc,\ Zc\ ]^T$$

A matriz **M**, será a matriz das derivadas parciais. Para as equações (3.2.5) (ou (3.1.35)) esta matriz terá os seguintes elementos:

$$M_1 = \begin{bmatrix} \frac{\partial F_1}{\partial \kappa} & \frac{\partial F_1}{\partial \phi} & \frac{\partial F_1}{\partial \omega} & \frac{\partial F_1}{\partial Xc} & \frac{\partial F_1}{\partial Yc} & \frac{\partial F_1}{\partial Zc} \\ \frac{\partial F_2}{\partial \kappa} & \frac{\partial F_2}{\partial \phi} & \frac{\partial F_2}{\partial \omega} & \frac{\partial F_2}{\partial Xc} & \frac{\partial F_2}{\partial Yc} & \frac{\partial F_2}{\partial Zc} \end{bmatrix}$$

onde:

$$\frac{\partial F_1}{\partial \kappa} = -\frac{u'.v - u.v'}{v^2} = -\frac{R_2'f.R_1.f - R_2.f.R_1'f}{(R_1.f)^2}$$

Analogamente para F_2:

$$\frac{\partial F_2}{\partial \kappa} = -f.\frac{w'.v - w.v'}{v^2} = -f.\frac{R_3'f.R_1.f - R_3.f.R_1'f}{(R_1.f)^2}$$

Adotando as abreviações "c" para cosseno e "s" para seno, as derivadas das linhas da matriz de rotação são dadas por:

$$\frac{\partial R_1}{\partial \kappa} = [\ -c\phi.s\kappa \quad c\omega.c\kappa - s\omega.s\phi.s\kappa \quad s\omega.c\kappa + c\omega.s\phi.s\kappa\ ] = R_2$$
$$\frac{\partial R_2}{\partial \kappa} = [\ -c\phi.c\kappa \quad -c\omega.s\kappa - s\omega.s\phi.c\kappa \quad -s\omega.s\kappa + c\omega.s\phi.c\kappa\ ] = -R_1$$
$$\frac{\partial R_3}{\partial \kappa} = [\ 0 \quad 0 \quad 0\ ] = 0$$

$$\frac{\partial R_1}{\partial \phi} = [\ -s\phi.c\kappa \quad s\omega.c\phi.c\kappa \quad -c\omega.c\phi.c\kappa\ ] = -R_3.c\kappa$$
$$\frac{\partial R_2}{\partial \phi} = [\ s\phi.s\kappa \quad -s\omega.c\phi.s\kappa \quad c\omega.c\phi.s\kappa\ ] = R_3.s\kappa$$

$$\frac{\partial R3}{\partial \phi} = [\ \ c\phi \qquad\qquad s\omega.s\phi \qquad\qquad -c\omega.s\phi\ \ ]$$

$$\frac{\partial R1}{\partial \omega} = [\ 0 \quad -s\omega.s\kappa + c\omega.s\phi.c\kappa \quad c\omega.s\kappa + s\omega.s\phi.c\kappa\ ] = [\ 0 \quad -r13 \quad r12\ ]$$
$$\frac{\partial R2}{\partial \omega} = [\ 0 \quad -s\omega.c\kappa - c\omega.s\phi.s\kappa \quad c\omega.c\kappa - s\omega.s\phi.s\kappa\ ] = [\ 0 \quad -r23 \quad r22\ ]$$
$$\frac{\partial R3}{\partial \omega} = [\ 0 \quad -c\omega.c\phi \quad -s\omega.c\phi\ ] = [\ 0 \quad -r33 \quad r32\ ]$$

As derivadas em relação ao elementos de rotação:

$$\frac{\partial F1}{\partial \kappa} = -\frac{-(R1.f).R1.f - R2.f.(R2.f)}{v^2} = -\frac{-u^2 - v^2}{v^2} = 1 + \frac{u^2}{v^2}$$

$$\frac{\partial F2}{\partial \kappa} = -f.\frac{-R3.f.R2.f}{v^2} = f.\frac{w.u}{v^2}$$

$$\frac{\partial F1}{\partial \phi} = -\frac{s\kappa.(R3.f).(R1.f) - (R2.f).c\kappa.(-R3.f)}{v^2} = -\frac{w.(s\kappa.R1 + c\kappa.R2).f}{v^2} =$$

$$= \frac{-(s\kappa.w.v + c\kappa.w.u)}{v^2}$$

ou ainda, desenvolvendo $(s\kappa.R1 + c\kappa.R2).f$ obtem-se:

$$(s\kappa.R1 + c\kappa.R2).f = c\omega.fx + s\omega.fy$$

Substituindo:

$$\frac{\partial F1}{\partial \phi} = \frac{-w.(c\omega.fy + s\omega.fz)}{v^2}$$

$$\frac{\partial F2}{\partial \phi} = -f.\frac{(c\phi.fx + s\omega.s\phi.fy - c\omega.s\phi.fz).v + c\kappa^2.w}{v^2}$$

$$\frac{\partial F1}{\partial \omega} = -\frac{(-r23.fy + r22.fz).v - u.(-r13.fy + r12.fz)}{v^2}$$

$$\frac{\partial F2}{\partial \omega} = -f.\frac{(-r33.fy + r32.fz).v - w.(-r13.fy + r12.fz)}{v^2}$$

Para desenvolver as derivadas em relação às translações, faz-se inicialmente as derivadas em relação ao vetor **f**:

$$\frac{\partial f}{\partial Xc} = [\ 0 \quad -n \quad m\ ]^T$$
$$\frac{\partial f}{\partial Yc} = [\ n \quad 0 \quad -l\ ]^T$$
$$\frac{\partial f}{\partial Zc} = [-m \quad l \quad 0\ ]^T$$

As derivadas em relação a u, v e w serão então:

$$\frac{\partial u}{\partial Xc} = (-r22.n + r23.m)$$
$$\frac{\partial v}{\partial Xc} = (-r12.n + r13.m)$$
$$\frac{\partial w}{\partial Xc} = (-r32.n + r33.m)$$

$$\frac{\partial u}{\partial Yc} = (r21.n - r23.l)$$
$$\frac{\partial v}{\partial Yc} = (r11.n - r13.l)$$
$$\frac{\partial w}{\partial Yc} = (r31.n - r33.l)$$

$$\frac{\partial u}{\partial Zc} = (-r21.m + r22.l)$$
$$\frac{\partial v}{\partial Zc} = (-r11.m + r12.l)$$
$$\frac{\partial w}{\partial Zc} = (-r31.m + r32.l)$$

Finalmente, as derivadas em relação a Xc, Yc e Zc serão:

$$\frac{\partial F_1}{\partial Xc} = -\frac{\frac{\partial u}{\partial Xc} v - u \frac{\partial v}{\partial Xc}}{v^2} \qquad \frac{\partial F2}{\partial Xc} = -f.\frac{\frac{\partial w}{\partial Xc} v - w \frac{\partial v}{\partial Xc}}{v^2}$$

$$\frac{\partial F_1}{\partial Yc} = -\frac{\frac{\partial u}{\partial Yc} v - u \frac{\partial v}{\partial Yc}}{v^2} \qquad \frac{\partial F2}{\partial Yc} = -f.\frac{\frac{\partial w}{\partial Yc} v - w \frac{\partial v}{\partial Yc}}{v^2}$$

$$\frac{\partial F_1}{\partial Zc} = -\frac{\frac{\partial u}{\partial Zc} v - u \frac{\partial v}{\partial Zc}}{v^2} \qquad \frac{\partial F2}{\partial Zc} = -f.\frac{\frac{\partial w}{\partial Zc} v - w \frac{\partial v}{\partial Zc}}{v^2}$$

Para as equações (3.2.2) (ou (3.1.33)) as derivadas parciais podem ser obtidas de forma análoga ao grupo anterior. Como as derivadas em relação aos

R_1 e a $\mathbf{f}$ são as mesmas, basta agrupá-las novamente para obter as derivadas das funções G_1 e G_2 em relação aos elementos do vetor estado.

$$\frac{\partial G_1}{\partial \kappa} = -\frac{v'.u - v.u'}{u^2} = -\frac{R_1'\mathbf{f}.R_2.\mathbf{f} - R_1.\mathbf{f}.R_2'\mathbf{f}}{(R_2.\mathbf{f})^2}$$

$$\frac{\partial G_2}{\partial \kappa} = -f.\frac{w'.u - w.u'}{u^2} = -f.\frac{R_3'\mathbf{f}.R_2.\mathbf{f} - R_3.\mathbf{f}.R_2'\mathbf{f}}{(R_2.\mathbf{f})^2}$$

$$\frac{\partial G_1}{\partial \kappa} = -\frac{(R_2.\mathbf{f}).u - v.(-R_1.\mathbf{f})}{u^2} = -\frac{u^2 + v^2}{u^2} = -1 - \frac{v^2}{u^2}$$

$$\frac{\partial G_2}{\partial \kappa} = -f.\frac{-w.(-R_1.\mathbf{f})}{u^2} = -f.\frac{w.v}{u^2}$$

$$\frac{\partial G_1}{\partial \phi} = -\frac{-c\kappa.(R_3.\mathbf{f}).u - v.s\kappa.(R_3.\mathbf{f})}{u^2} = -\frac{-w.u.c\kappa - w.v.s\kappa}{u^2}$$

$$\frac{\partial G_2}{\partial \phi} = -f.\frac{(c\phi.fx + s\omega.s\phi.fy - c\omega.s\phi.fz).u - s\kappa.w^2}{u^2}$$

$$\frac{\partial G_1}{\partial \omega} = -\frac{(-r13.fy + r12.fz).u - v.(-r23.fy + r22.fz)}{u^2}$$

$$\frac{\partial G_2}{\partial \omega} = -f.\frac{(-r33.fy + r32.fz).u - w.(-r23.fy + r22.fz)}{u^2}$$

As derivadas das funções G_1 em relação às translações serão:

$$\frac{\partial G_1}{\partial Xc} = -\frac{\frac{\partial v}{\partial Xc}u - v\frac{\partial u}{\partial Xc}}{u^2} \qquad \frac{\partial G_2}{\partial Xc} = -f.\frac{\frac{\partial w}{\partial Xc}u - w\frac{\partial u}{\partial Xc}}{u^2}$$

$$\frac{\partial G_1}{\partial Yc} = -\frac{\frac{\partial v}{\partial Yc}u - v\frac{\partial u}{\partial Yc}}{u^2} \qquad \frac{\partial G_2}{\partial Yc} = -f.\frac{\frac{\partial w}{\partial Yc}u - w\frac{\partial u}{\partial Yc}}{u^2}$$

$$\frac{\partial G_1}{\partial Zc} = -\frac{\frac{\partial v}{\partial Zc}u - v\frac{\partial u}{\partial Zc}}{u^2} \qquad \frac{\partial G_2}{\partial Zc} = -f.\frac{\frac{\partial w}{\partial Zc}u - w\frac{\partial u}{\partial Zc}}{u^2}$$

3.2.1 SEQUÊNCIA DE PROCEDIMENTOS PARA O CÁLCULO DO IEKF.

Nesta seção será descrita a sequência de procedimentos para a estimação do estado usando a Filtragem Kalman. Esta sequência será apresentada supondo-se que a câmara está acoplada à garra de um manipulador, e que o objetivo do procedimento é **localizar** a câmara continuamente para verificar a posição e orientação da garra.

Como o modelo cinemático do manipulador é conhecido, é possível estimar a posição da câmara em um dado instante t_k a partir dos ângulos de junta do manipulador. Esta informação é usada, em conjunto com as demais transformações entre referenciais (seção 2.4) para estimar um estado predito inicial para o processo de filtragem.

No caso do problema de localização de objetos, uma estimativa predita inicial deve ser obtida através do processo de reconhecimento.

A sequência de procedimentos para calcular o IEKF é descrita a seguir

1) Num instante inicial o vetor estado $\hat{\mathbf{x}}_{1;0}$ é conhecido, com uma matriz covariância $\mathbf{P}_{1;0}$;

2) No instante t_k é obtida a primeira observação, usando a estimativa predita $\hat{\mathbf{x}}_{k;k-1}$ para reduzir o espaço de busca da reta na imagem. Esta primeira reta fornecerá a e b, bem como suas variâncias. Estas duas observações serão usadas para obter uma estimativa de estado. Como o filtro é recursivo poder-se-ia processar em primeiro lugar a e depois b, sem necessidade de inversão de matrizes. Como a estimativa intermediária, baseada somente em "a" não teria utilidade prática, é mais conveniente processar simultaneamente a e b. Neste caso a ordem das matrizes é:

$$ {}_6\hat{\mathbf{x}}_1 \quad {}_6\eta_1 \quad {}_2\mathbf{M}_6 \quad {}_2\mathbf{R}_2 \quad {}_6\mathbf{P}_6 \quad {}_6\mathbf{K}_2 \quad {}_2\mathbf{z}_1 $$

Pode-se admitir **R** diagonal, isto é, que a e b sejam estatísticamente independentes (não correlacionadas). Ainda nesta etapa, mesmo antes de calcular a estimativa filtrada, compara-se o resíduo predito em relação ao seu desvio padrão. Se o resíduo padronizado for maior que 3, supõe-se que há divergência do filtro.

3) Uma vez obtida a estimativa filtrada para o estado, usando a j-ésima observação, calcula-se sua matriz covariância, com base na equação (2.3.12);

4) Para o processamento da observação (j+1) considera-se que não houve alteração no estado do sistema, não havendo necessidade, portanto, de fazer a transição de estado. A estimativa filtrada obtida com a j-ésima observação (estado e matriz covariância) é usada diretamente como predição para a observação (j+1). A estimativa do estado na observação (j) é usada para reduzir o espaço de busca da observação (j+1) na imagem. Uma vez obtida a observação (j+1) pode-se obter uma estimativa de estado usando esta observação, bem como sua matriz covariância. Repete-se o procedimento para todas as retas disponíveis na imagem no instante t_k;

5) Faz-se, agora, uma predição do estado para o instante t_{k+1}, com base na estimativa de estado atual. Esta predição é feita usando o modelo do sistema (transição do estado). Obtém-se, portanto:

$\hat{\mathbf{x}}_{k+1|k}$ e $\mathbf{P}_{k+1|k}$

5) Faz-se k = k + 1 e volta-se ao passo 2;

3.3 PROCESSO RECURSIVO ESTIMAÇÃO-EXTRAÇÃO DE FEIÇÕES

3.3.1 CONCEITO E ESTRUTURA DO PROCESSO RECURSIVO.

O modelo matemático desenvolvido na seção 3.1 permite a implementação de um processo recursivo entre a etapa de estimação de parâmetros e a etapa de extração de feições, com base nas estimativas geradas pelo processamento sequencial de cada reta extraída na imagem.

A utilização de estimativas a priori para reduzir o espaço de busca das feições na imagem recebe alguma menção na literatura (Feddema e Lee, 1990). Fica claro, entretanto, que nas abordagens existentes, uma mesma estimativa é usada para extrair todas as feições na imagem. A idéia de utilizar os resultados parciais (estimativas filtradas) do processo de filtragem como "feedback" para a etapa de extração de feições é nova, não sendo encontrada

referência na literatura consultada.

Nas abordagens existentes, que utilizam o conhecimento a priori para reduzir o espaço de busca, obtem-se uma estimativa inicial para os parâmetros da câmara (κ, ϕ, ω, Xc, Yc, Zc) e suas respectivas variâncias e, a partir destes elementos, são definidos polígonos de busca (janelas) na imagem. É realizado, então, o processo de classificação, isto é, a determinação das correspondências e, finalmente, é calculada uma nova estimativa para os parâmetros da câmara, usando todas as retas (ou pontos) disponíveis.

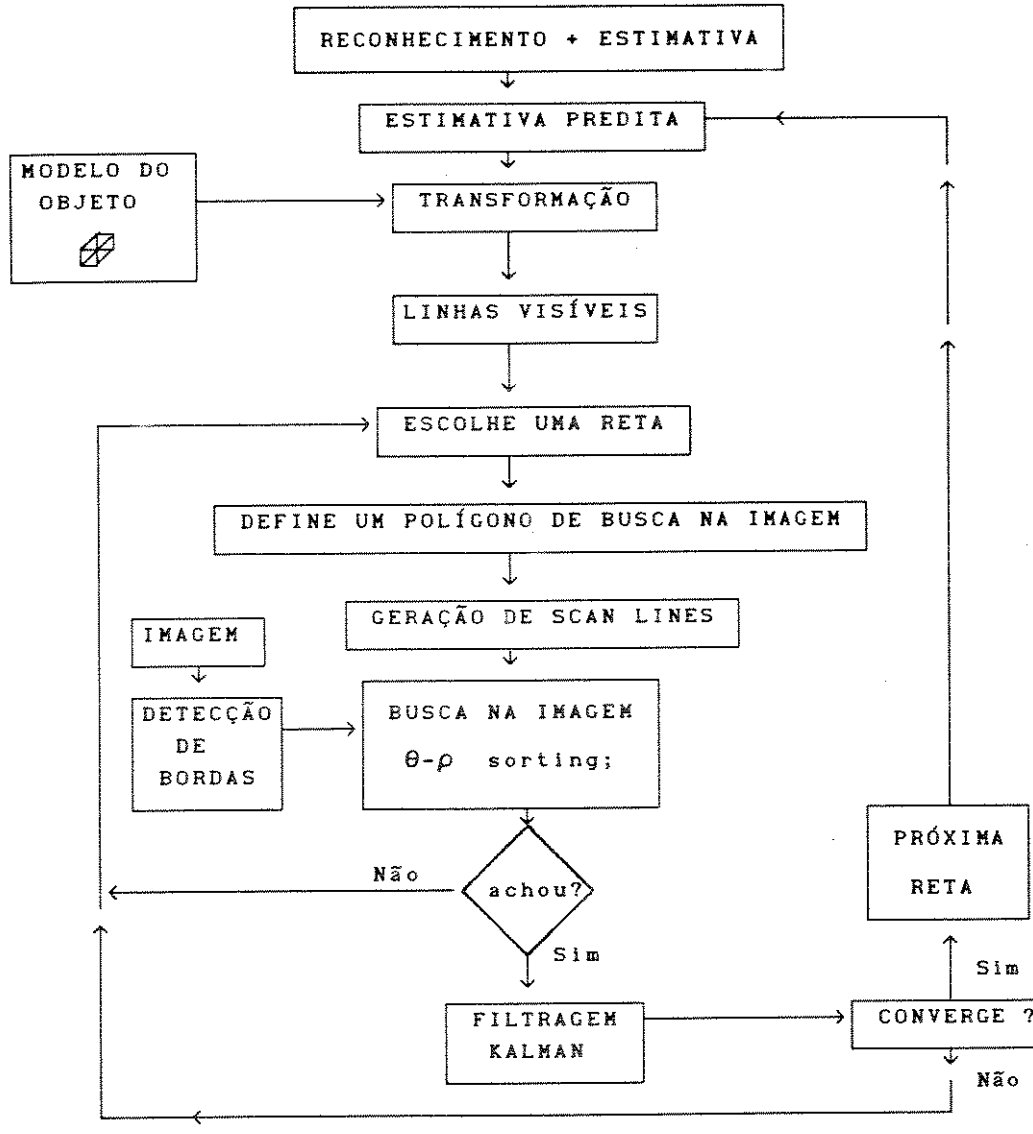


Figura 3.3.1 Fluxo do procedimento recursivo entre filtragem e extração de feições.

No processo ora proposto, estabelece-se um processo iterativo, onde estimativas parciais dos parâmetros são obtidas e reutilizadas para reduzir o espaço de busca da próxima feição. Na figura 3.3.1 estão representadas as etapas necessárias para este procedimento recursivo (iterativo).

Nas seções seguintes serão descritas cada uma das etapas presentes no fluxo apresentado na figura 3.3.1.

3.3.2 Reconhecimento e Estimativa Predita.

Num instante inicial, quando o sistema de Visão captura a primeira imagem de um objeto, é necessário reconhecê-lo. Este reconhecimento pode ser feito de várias maneiras e usando diferentes critérios. Alguns aspectos do reconhecimento foram discutidos no Capítulo 2. A definição de uma estimativa predita (a priori) para a localização do objeto pode ser feita em conjunto com o reconhecimento ou separadamente. Caso o objeto tenha sido reconhecido previamente, devem ser estabelecidas algumas correspondências mínimas entre feições no objeto e na imagem para ser possível calcular uma localização inicial. Estas correspondências podem ser obtidas usando primitivas de indexação, da mesma maneira que no reconhecimento. A rigor, pode-se afirmar que a técnica é a mesma que do reconhecimento, embora o espaço de busca seja significativamente menor, pois o objeto já é conhecido.

Como, nesta etapa, deseja-se apenas uma estimativa a priori, pode-se trabalhar com a imagem em nível de resolução mais grosseiro e usando métodos de localização mais rápidos, como os apresentados no Capítulo 2, com o objetivo de acelerar o processamento.

Esta etapa está além do escopo deste trabalho. Supõe-se que o objeto tenha sido reconhecido e que uma estimativa de localização tenha sido obtida. Os requisitos de precisão desta estimativa a priori serão estudados experimentalmente.

No problema de localização da câmara, ao contrário da localização de objetos, esta etapa inicial é extremamente simples, pois o modelo cinemático do robô fornece uma estimativa a priori para os parâmetros da câmara com razoável precisão.

3.3.3 MODELO DO OBJETO E TRANSFORMAÇÃO.

O modelo do objeto é dado por uma representação geométrica, descrevendo as primitivas básicas do objeto e seus relacionamentos. O modelo matemático desenvolvido neste trabalho baseia-se em linhas retas para a localização de objetos ou da câmara. Supõe-se, então, que o objeto possua algumas linhas retas, embora, possa não ser, necessariamente, um poliedro.

Para fins de experimentação, e com a finalidade de adotar técnicas de modelamento de sólidos mais simples, os objetos serão caracterizados pelas coordenadas de seus vértices, pelas arestas deste objeto e pelas suas faces, constituídas, por sua vez, de uma lista de arestas modelo B-rep).

Nesta etapa a transformação projetiva é aplicada às coordenadas dos vértices do objeto, transformando-as para o referencial imagem. Neste caso, a coordenada z (imagem) também deve ser calculada (Apêndice 1) para permitir a determinação das faces escondidas. Esta transformação é aplicada usando os parâmetros dados pela estimativa predita. As coordenadas imagem dos vértices do objeto também serão utilizadas posteriormente para definir o polígono de busca da reta na imagem.

3.3.4 DETERMINAÇÃO DE LINHAS VISÍVEIS E ESCOLHA DE UMA RETA.

Uma vez que todas as linhas presentes no modelo do objeto tenham sido projetadas para o referencial imagem, usando a estimativa predita, é necessário verificar quais destas linhas estarão escondidas pelas faces do objeto.

O algoritmo de Roberts, descrito em Rogers 1985), embora seja um dos mais antigos, é adequado a esta abordagem, pois trata o problema usando descrições de objetos poliédricos convexos. No caso de objetos côncavos, pode-se subdividi-los recursivamente até formar um conjunto de objetos convexos. Neste método, calculam-se as equações dos planos que definem as faces do objeto (volume) e, em seguida verifica-se para um determinado ponto de vista, quais são as faces invisíveis. Isto é feito mediante um simples produto escalar do vetor posição do ponto de vista com os vetores que compõem a matriz volume. Este procedimento permite calcular as faces escondidas. Como, neste trabalho, supõe-se a presença de objetos isolados, não foi implementada a parte do algoritmo que trata da interação com outros objetos.

Uma vez definidas as faces visíveis (e, consequentemente as linhas visíveis) deve-se escolher uma das retas visíveis para buscar sua homóloga na

imagem de bordas. Esta escolha baseia-se em uma consulta à uma lista de retas visíveis que ainda não tenham sido usadas no processo.

3.3.5 DEFINIÇÃO DO POLÍGONO DE BUSCA.

Uma das etapas iniciais do processo de busca na imagem é a determinação da posição aproximada da reta na imagem com base na estimativa existente dos parâmetros da câmara. Uma hipótese inicial é utilizar diretamente as equações (3.1.33) e (3.1.35) e obter a e b (ou a^* e b^*). Entretanto, não estariam determinados o início e o fim do segmento, o que tornaria necessária a busca na direção da reta até os limites da imagem. A segunda possibilidade, mais eficiente, é calcular as coordenadas imagem correspondentes aos extremos do segmento de reta (vértices no modelo do objeto) usando a equação de colinearidade (isto já foi feito na etapa de transformação).

Considerando a segunda alternativa, isto é, determinar as coordenadas imagem dos vértices da reta, usando a equação de colinearidade, torna-se necessário estabelecer uma estimativa para o erro esperado na imagem, considerando o erro existente na estimativa dos parâmetros da câmara. Isto pode ser obtido através da propagação de covariâncias.

Seja $\hat{\mathbf{x}} = [\ \kappa\ \ \phi\ \ \omega\ \ Xc\ \ Yc\ \ Zc\ ]$ o vetor estimado dos parâmetros da câmara e $\sum\mathbf{x}$ sua matriz covariância. Admitindo, por simplicidade, que os parâmetros são estatisticamente independentes, as covariâncias tornam-se nulas e $\sum\mathbf{x}$ pode ser escrita como:

$$\sum\mathbf{x} = \begin{bmatrix} \sigma^2_{\kappa} & & & & & 0 \\ & \sigma^2_{\phi} & & & & \\ & & \sigma^2_{\omega} & & & \\ & & & \sigma^2_{Xc} & & \\ & 0 & & & \sigma^2_{Yc} & \\ & & & & & \sigma^2_{Zc} \end{bmatrix}$$

Lembrando que as equações de colinearidade (equações projetivas) são dadas por duas funções (Apêndice 1):

$$x_i = F_x(\mathbf{x}) \quad \text{e} \quad y_i = F_y(\mathbf{x}) \tag{3.3.1}$$

O objetivo do procedimento de propagação de covariâncias é determinar as

variâncias de x_i e y_i a partir de $\sum \mathbf{x}$. O procedimento de propagação das covariâncias pode ser resumido através da equação:

$$\sum \mathbf{z} = D \cdot \sum \mathbf{x} \cdot D^T \qquad (3.3.2)$$

onde: $\sum \mathbf{z}$ é a matriz covariância das coordenadas imagem x_i e y_i;

D é a matriz das derivadas parciais em relação aos parâmetros extrínsecos (Capítulo 2, seção 2.1.1);

Considerando as correlações nulas na matriz $\sum \mathbf{x}$ as variâncias de x_i e y_i podem ser dadas por:

$$\begin{aligned} \sigma_x^2 &= D_{11}^2 \cdot \sigma_\kappa^2 + D_{12}^2 \cdot \sigma_\phi^2 + \ldots + D_{16}^2 \cdot \sigma_{Xc}^2 \\ \sigma_y^2 &= D_{21}^2 \cdot \sigma_\kappa^2 + D_{22}^2 \cdot \sigma_\phi^2 + \ldots + D_{26}^2 \cdot \sigma_{Xc}^2 \end{aligned} \qquad (3.3.3)$$

Ou, em forma mais concisa:

$$\begin{aligned} \sigma_x^2 &= \sum_{j=1}^{6} D_{1j}^2 \cdot \hat{\sigma}_{\mathbf{x}_j}^2 \\ \sigma_y^2 &= \sum_{j=1}^{6} D_{2j}^2 \cdot \hat{\sigma}_{\mathbf{x}_j}^2 \end{aligned} \qquad (3.3.4)$$

Com estes elementos é possível estabelecer as variâncias dos pontos extremos de um segmento de reta na imagem, a partir das coordenadas destes extremos no espaço objeto (vértices no modelo do objeto) e dos parâmetros da câmara e suas respectivas variâncias (estimativa predita).

Pode-se, então, delimitar um polígono contornando o segmento provável (calculado). Para simplificar a definição deste polígono é estabelecido um valor médio para o erro em x e y. Este valor médio pode ser entendido como a metade da largura da janela ao redor do ponto. Para aumentar a probabilidade de que o segmento de reta real esteja dentro do polígono, pode-se aumentar a largura da janela em 2 ou 3 vezes ($2x\sigma$ ou $3x\sigma$). Uma vez estabelecida a largura da janela é possível delimitar o polígono de busca contornando o segmento de reta.

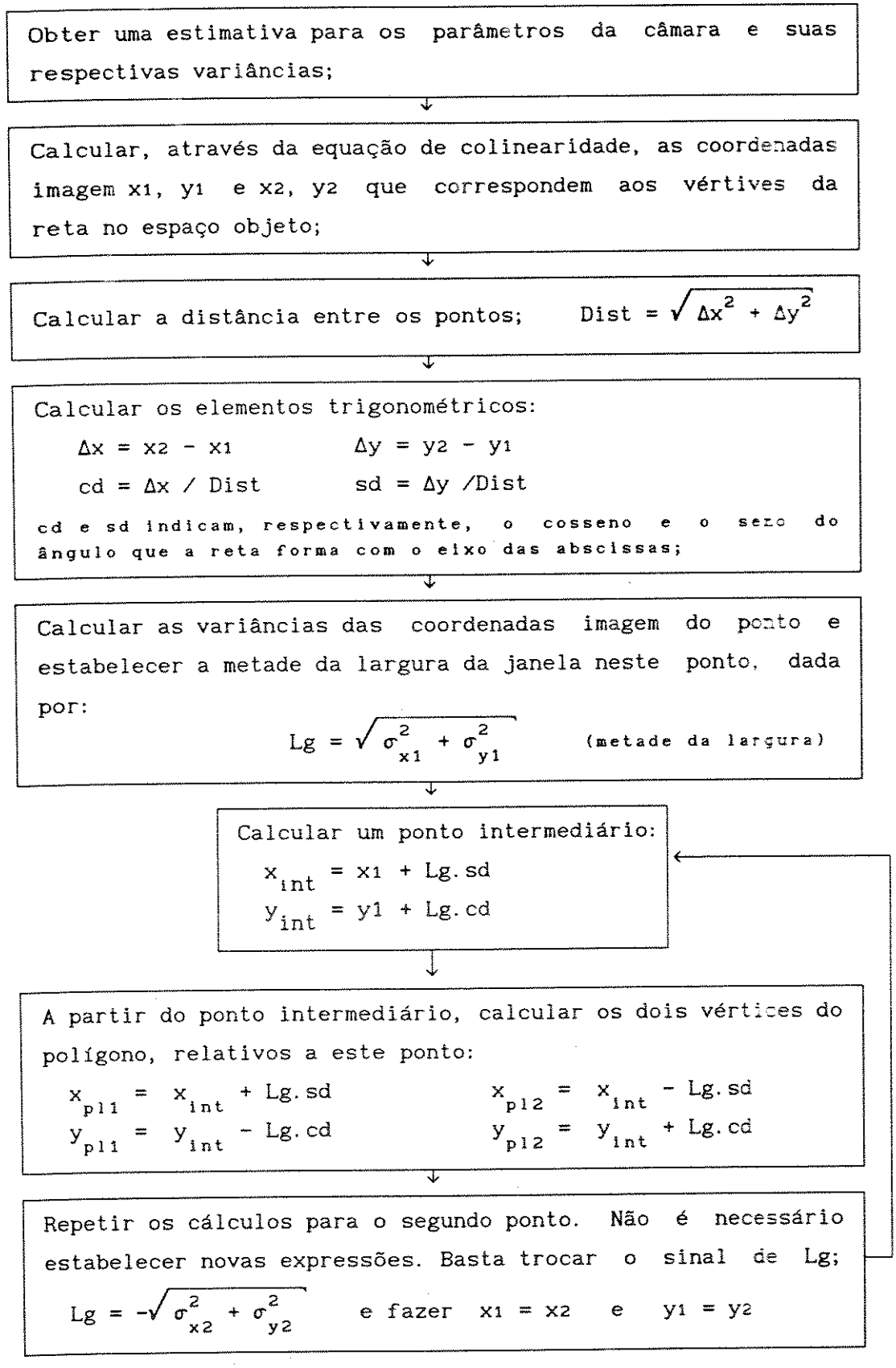


Figura 3.3.2 Sequência de cálculos dos vértices do polígono de busca.

A sequência de cálculos necessários para definir os vértices do polígono de busca é apresentada na figura 3.3.2. Na figura 3.3.3 são mostrados os elementos geométricos utilizados para o cálculo do polígono.

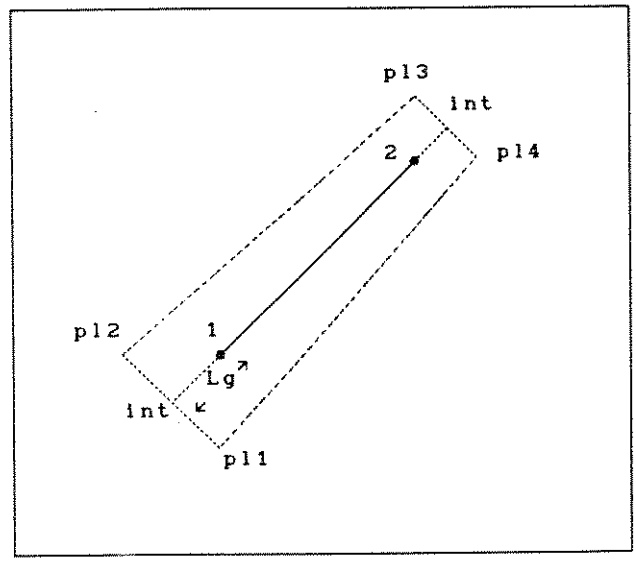


Figura 3.3.3 Elementos do polígono de busca.

3.3.6 GERAÇÃO DAS "SCAN LINES"

Uma vez definidos os vértices do polígono de busca é necessário verificar se estão dentro dos limites da imagem e, então, definir, para cada linha do "frame buffer", o começo e o fim do espaço de busca. Isto corresponde à geração de "scan-lines".

Caso o polígono de busca ultrapasse os limites da imagem é necessário realizar um recorte ("clipping") deste em relação aos limites da imagem. Para evitar esta operação o algorítmo de geração de "scan-lines" foi modificado para realizar simultâneamente este "clipping".

Inicialmente, transformam-se as coordenadas que definem o polígono para o referencial de máquina, isto é, transforma-se coordenadas imagem para valores que representem linhas e colunas no "frame-buffer".

Em seguida ordenam-se as arestas que definem o polígono, montando uma lista de arestas ativas (Rogers, 1985). Para o algoritmo proposto esta lista é mais simples, pois somente duas arestas estarão ativas de cada vez.

Verifica-se se o polígono está totalmente fora dos limites da imagem, simplesmente comparando-se as coordenadas dos vértices do polígono com as dos

limites da imagem.

Verifica-se se as arestas que definem o polígono de busca são paralelas aos limites da imagem. Neste caso, a geração de scan lines é simples, bastando adotar um valor único para todas as linhas. Caso contrário, definir a primeira "scan-line", verificando se o primeiro vértice está dentro ou fora da imagem. Se estiver dentro então a primeira "scan-line" começa neste vértice; se estiver fora, então a "scan-line" começa no limite superior da imagem.

A partir de então são calculadas as intersecções entre a linha horizontal que define a "scan-line" e as duas arestas ativas. Verifica-se se estas intersecções estão dentro da imagem e, para cada caso, adota-se o valor adequado. Esta verificação elimina a necessidade de um procedimento de "clipping" após a definição do polígono de busca.

O resultado final deste processo é uma lista com as "scan lines" que definem o polígono de busca, com as respectivas colunas no "frame-buffer" indicando o início e fim do polígono de busca. A figura 3.3.4 ilustra este conceito, enquanto que na figura 3.3.5 estão representadas as etapas mais importantes deste processo.

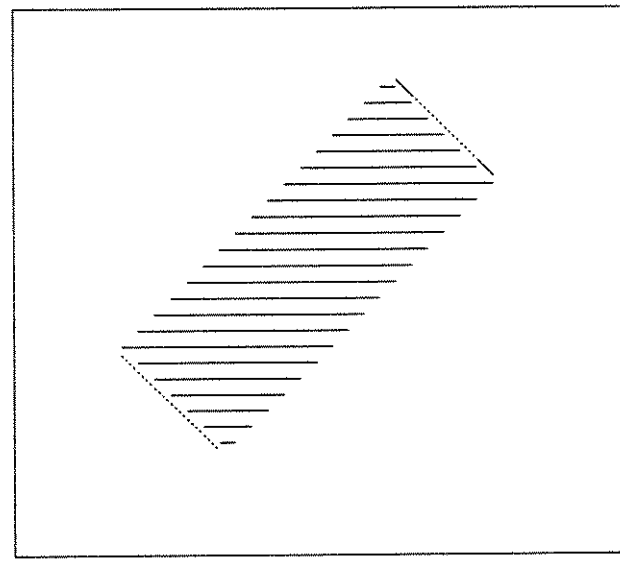


Figura 3.3.4 "Scan lines" para o polígono de busca.

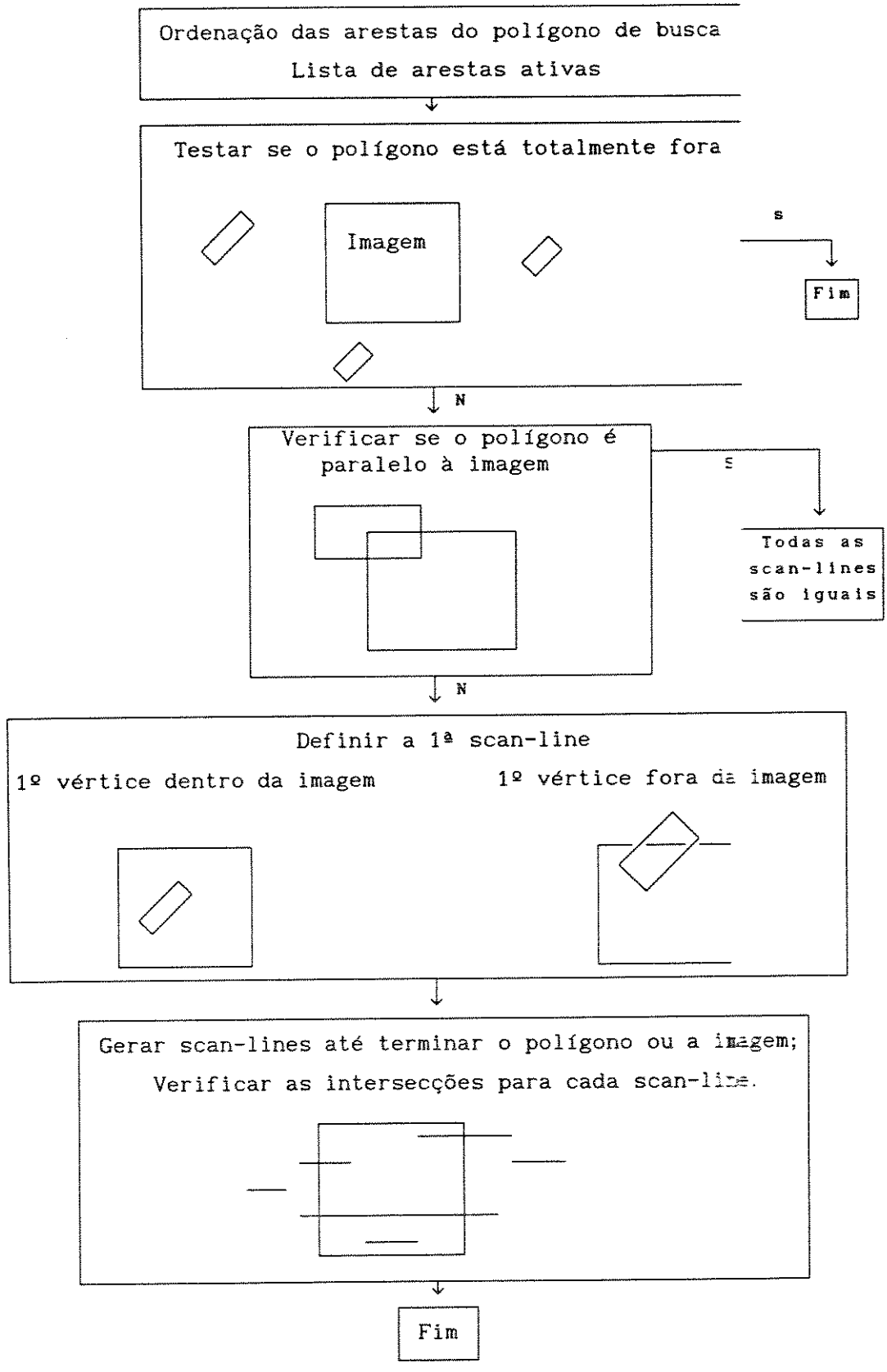


Figura 3.3.5 Sequência para a geração de "scan-lines"

3.3.7 BUSCA NA IMAGEM - VETORIZAÇÃO

Nesta fase será efetuada uma busca na imagem real coletada, tendo como domínio o polígono de busca definido através das "scan lines". Pode-se atuar sobre uma imagem de bordas, previamente processada ou, então, aplicar o procedimento de detecção somente sobre a área definida pelo polígono de busca.

Para esta janela devem ser extraidas as retas existentes, usando o agrupamento θ-ρ. O resultado deste procedimento é um conjunto de pixels com θ e ρ semelhantes. A este conjunto ajusta-se uma reta, eliminando-se sucessivamente os pontos cujos resíduos ultrapassarem um determinado critério.

Para o processo recursivo proposto neste trabalho podem ser estabelecidas adaptações neste método de agrupamento θ-ρ, que aumentem sua eficiciência. O objetivo do processo de vetorização, neste trabalho, é extrair **uma única** reta de cada vez. Assim, para uma determinada janela na imagem, calcula-se o histograma em θ e seleciona-se apenas o agrupamento θ que mais se aproximar do θ predito. Para este agrupamento θ calcula-se a distribuição ρ e seleciona-se somente o ρ mais próximo do ρ predito. Somente depois de definidos os agrupamentos θ e ρ faz-se a acumulação nesta estrutura. Este agrupamento consistirá de todos os pixels que possuam θ e ρ próximos aos valores selecionados.

Uma vez agrupados os pixels em uma classe θ-ρ pode-se passar ao ajuste da reta pelo MMQ. Antes de ajustar a reta é conveniente eliminar o efeito dos erros sistemáticos nestes pontos armazenados e reduzi-los ao referencial imagem.

3.3.7.1 REFINAMENTO DAS COORDENADAS DE MÁQUINA.

O refinamento das coordenadas de máquina consiste na redução ao referencial imagem (centro da imagem) e na eliminação dos erros sistemáticos.

A imagem armazenada no "frame buffer" está referenciada ao canto inferior esquerdo da imagem. Aplica-se, inicialmente uma translação ao centro aproximado da imagem (256, 200) e multiplica-se por um fator de escala, que corresponde ao tamanho do pixel (0.010), obtendo-se as coordenadas em milímetros. Aplicam-se, então, reflexões nos eixos, para tornar o referencial imagem compatível com a posição da câmara no espaço (caso contrário seria obtido o complemento de κ).

Uma vez calculadas estas reduções aproximadas pode-se eliminar o efeito

dos erros sistemáticos, através das seguintes expressões, apresentadas no Capítulo 2, seção 2.1.1:

$$x_i = x_f - c_x + (x_f - c_x)(k_1 r^2) + (x_f - c_x)ds_x$$

$$y_i = y_f - c_y + (y_f - c_y)(k_1 r^2)$$

onde: x_i e y_i são as coordenadas imagens de um ponto, isentas de erros sistemáticos e reduzidas ao centro da imagem;

c_x, c_y, k_1 e ds_x são parâmetros intrínsecos, obtidos no processo de calibração da câmara;

3.3.7.2 AJUSTE DA RETA PELO M.M.Q.

Uma das modificações introduzidas no método proposto originalmente por Dudani e Luk (1978) diz respeito ao ajuste da reta pelo MMQ. Ao invés de ajustar uma reta usando a equação normal (parâmetros θ e ρ) será ajustada uma reta usando a forma paramétrica (parâmetros a e b, ou a^* e b^*, dependendo do valor do ângulo θ).

Para ajustar uma reta do tipo $y = a.x + b$ pelo método dos mínimos quadrados considera-se y como observação e x isento de erro. As expressões que permitem calcular a e b, minimizando o somatório do quadrado dos resíduos em y, são dadas em (3.3.5).

$$a = \frac{1}{det} . (N.\sum x_i . y_i - \sum x_i . \sum y_i) \qquad (3.3.5)$$

$$b = \frac{1}{det} . (-\sum x_i . \sum x_i . y_i + \sum x_i^2 . \sum y_i)$$

onde:

N é o número de pixels;

$det = (\sum x_i^2 . N - \sum x_i . \sum x_i)$

As variâncias de a e b são dadas por:

$$\sigma_a^2 = \frac{N}{det} \qquad e \qquad \sigma_b^2 = \frac{\sum x^2}{det} \qquad (3.3.6)$$

De modo análogo, quando o θ aproximado indicar que a reta a ser ajustada

é o tipo $x = a^*.y + b^*$ (Capítulo 3), usam-se as equações (3.3.7).

$$a^* = \frac{1}{det^*}.(N.\sum y_i.x_i - \sum y_i.\sum x_i)$$
$$b^* = \frac{1}{det^*}.(-\sum y_i.\sum x_i.y_i + \sum y_i^2.\sum x_i) \qquad (3.3.7)$$

onde:

N é o número de pixels;

$det^* = (\sum y_i^2.N - \sum y_i.\sum y_i)$

As variâncias de a^* e b^* são dadas por:

$$\sigma^2_{a^*} = \frac{N}{det} \qquad e \qquad \sigma^2_{b^*} = \frac{\sum y^2}{det} \qquad (3.3.8)$$

Para evitar a presença de pontos espúrios pode-se fazer um teste de resíduos ponto a ponto, elimando-se aqueles pontos que apresentarem um erro superior a um critério pré-estabelecido. Elimina-se o ponto que apresentar o maior resíduo, refazendo-se o ajuste para o novo conjunto de pontos.

As variâncias de a e b (ou a^* e b^*) indicam a precisão do ajuste e são estatísticas relacionadas com a dispersão dos pontos em relação ao valor ajustado.

Outra estatística importante, que permite avaliar a qualidade das observações e, de certa maneira, a dispersão em torno do valor ajustado, é a variância a posteriori ($\hat{\sigma}^2_0$), também conhecida como variância da observação de peso unitário. Como os pesos adotados são unitários, esta estatística reflete a variância das observações, no caso, as coordenadas y (ou x no caso a^*-b^*) dos pontos pertencentes à reta.

O $\hat{\sigma}^2_0$ é calculado através da expressão (3.3.9):

$$\hat{\sigma}^2_0 = \frac{V^T V}{n-u} \qquad (3.3.9)$$

onde: V é o vetor dos resíduos, n é o número de observações e u é o número de incógnitas;

Para o caso específico da reta a-b é possível desenvolver uma expressão específica. Os resíduos são dados por:

$$v_i = a.x_i + b_i - y_i \qquad (3.3.10)$$

O somatório do quadrado dos resíduos é, consequentemente:

$$V^TV = \sum (a.x_i + b_i - y_i)\ (a.x_i + b_i - y_i) \qquad (3.3.11)$$

Finalmente:

$$V^TV = a^2\sum x_i^2 + 2.a.b\ \sum x_i - 2.a\ \sum x_i.y_i - 2.b\ \sum y_i + N.b^2 + \sum y^2 \qquad (3.3.12)$$

É importante observar que θ e ρ obtidos no agrupamento inicial são definidos no referencial de máquina, enquanto que os parâmetros a e b obtidos no ajuste são definidos no referencial imagem.

Outro aspecto importante a observar é que as variâncias obtidas através das expressões (3.3.6) e (3.3.8) representam apenas a precisão dos parâmetros. Para obter a exatidão destes parâmetros, é preciso levar em conta os erros nos parâmetros intrínsecos, através de propagação de covariâncias. Esta avaliação é apresentada a seguir.

3.3.7.3 Efeito dos erros nos parâmetros intrínsecos na avaliação da precisão da reta na imagem.

A estimativa de precisão obtida pela matriz covariância, no ajuste da reta imagem pelo MMQ é apenas um indicativo de dispersão. É necessário considerar os erros devidos à incerteza nos parâmetros intrínsecos, para evitar uma super-estimação da precisão das retas na imagem. Esta super-estimação da qualidade das observações pode resultar em divergência do filtro.

Para levar em conta os erros nos parâmetros intrínsecos basta realizar uma propagação de covariâncias.

Inicialmente, é necessário estabelecer as expressões para os parâmetros a e b (a^* e b^*) em função das coordenadas de máquina e dos parâmetros intrínsecos.

Sejam a' e b' os parâmetros da reta no referencial imagem, isto é, corrigidas dos erros sistemáticos; a e b são os parâmetros da reta ainda com estes erros. Para efeito de simplificação não será considerada a distorção

ótica, devido à pequena magnitude de seu erro. Além disto, a distorção ótica, por ser radial e simétrica, não teria um valor único para toda a reta. Com isto, o eventual erro no parâmetro k_1 se refletiria na dispersão da reta.

A equação da reta no referencial de máquina é dada por:

$$(y_2 - y_1).x - (x_2 - x_1).y - (x_1.y_2 - y_1.x_2) = 0 \qquad (3.3.13)$$

onde x_i, y_i são pontos pertencentes à reta;

Na forma paramétrica esta reta é dada por:

$$y = a.x + b \qquad (3.3.14)$$

Os parâmetros a e b podem ser escritos em função das coordenadas como:

$$a = \frac{(y_2 - y_1)}{(x_2 - x_1)} \qquad b = \frac{(y_1.x_2 - x_1.y_2)}{(x_2 - x_1)} \qquad (3.3.15)$$

Se as coordenadas imagem estiverem no referencial de máquina, então devem ser introduzidas as correções aos erros sistemáticos. Com estas correções serão obtidas as coordenadas imagem. Os parâmetros a' e b' são os correspondentes a a e b, mas no referencial imagem:

$$a' = \frac{(y_2 - c_y - (y_1 - c_y))}{(x_2-c_x+(x_2-c_x)d_s - x_1+c_x-(x_1-c_x)d_s)} = \frac{(y_2 - y_1)}{(x_2 - x_1).(1 + d_s)}$$

finalmente:

$$a' = a.\frac{1}{(1+d_s)} \qquad (3.3.16)$$

Verifica-se, portanto, que o parâmetro a' não depende das coordenadas do centro da imagem. Para o parâmetro b:

$$b' = \frac{(y_1-c_y)(x_2-c_x+(x_2-c_x)d_s) - (x_1-c_x+(x_1-c_x)d_s)(y_2-c_y)}{(x_2 - x_1).(1 + d_s)} =$$

$$= \frac{b}{(1+d_s)} + \frac{a.c_x}{(1+d_s)} + \frac{b.d_s}{(1+d_s)} - \frac{c_y}{(1+d_s)} + \frac{a.c_x.d_s}{(1+d_s)} - \frac{c_y.d_s}{(1+d_s)}$$

Reagrupando os termos, chega-se, finalmente a:

$$b' = b + a.c_x - c_y \qquad (3.3.17)$$

Desenvolvendo um raciocínio análogo para a^* e b^* obtém-se $a^{*'}$ e $b^{*'}$, parâmetros corrigidos dos erros sistemáticos:

$$a^{*'} = a^*.(1 + ds) \tag{3.3.18}$$

$$b^{*'} = (b^* - cx + a^*.cy).(1 + ds) \tag{3.3.19}$$

onde:

$$a^* = \frac{(x2 - x1)}{(y2 - y1)} \qquad b^* = \frac{(x1.y2 - y1.x2)}{(y2 - y1)} \tag{3.3.20}$$

Considerando-se, agora, que foram obtidos os parâmetros a e b (ou a^* e b^*) e suas respectivas variâncias, através do ajuste de retas pelo MMQ , deseja-se calcular a' e b' e suas respetivas variâncias. Os parâmetros a' e b' (ou $a^{*'}$ e $b^{*'}$) são calculados através das expressões (3.3.16 , (3.3.17), (3.3.18) e (3.3.19) e as suas variâncias são obtidas por meio de propagação a partir das variâncias de a e b e das variâncias dos parâmetros intrínsecos.

$$\sigma^2_{a'} = \frac{1}{(1+ds)^2}\sigma^2_a + \frac{a^2}{(1+ds)^4}\sigma^2_{ds} \tag{3.3.21}$$

$$\sigma^2_{b'} = cx^2.\sigma^2_a + \sigma^2_b + a^2.\sigma^2_{cx} + \sigma^2_{cy} \tag{3.3.22}$$

e, para $a^{*'}$ e $b^{*'}$:

$$\sigma^2_{a^{*'}} = (1+ds)^2\,\sigma^2_{a^*} + a^{*2}\,\sigma^2_{ds} \tag{3.3.23}$$

$$\sigma^2_{b^{*'}} = (1+ds)^2(cy^2\sigma^2_{a^*} + \sigma^2_{b^*} + \sigma^2_{cx} + a^{*2}\sigma^2_{cy}) + (b^* - cx + a^*cy)\sigma^2_{ds} \tag{3.3.24}$$

Na prática, o que se faz é refinar as coordenadas dos pontos que constituem as bordas, isto é, corrigir os erros sistemáticos, ajustar a reta e calcular as variâncias dos parâmetros ajustados. Usam-se, então, as expressões (3.3.21), (3.3.22), (3.3.23) e (3.3.24) para calcular as variâncias reais dos parâmetros da reta, levando-se em conta o efeito dos erros nos parâmetros

intrínsecos. O único erro que não foi considerado é o erro em k_1, porque este erro tem pequena magnitude e também, devido à sua natureza radial e simétrica, já se refletir na dispersão da reta.

Uma vez definidos os erros nas medidas, encerra-se o processo de extração de feições. Pode-se, então, passar ao procedimento de filtragem, que procurará obter uma estimativa filtrada para o estado, com base nas medidas extraídas. É importante ressaltar que no caso do agrupamento θ-ρ serão calculados os parâmetros a e b (ou a^* e b^*) e suas variâncias como resultado do ajuste da reta. Os valores θ e ρ são usados apenas como critério para agrupamento de pixels e escolha da reta mais provável.

3.3.8 FILTRAGEM

As medidas extraídas no processo de vetorização são usadas para filtrar o vetor estado. Uma análise de convergência permite verificar se as observações foram obtidas de modo adequado, se não houve erro grosseiro na seleção da reta, ou se a precisão das observações foi estimada corretamente.

O resultado do processo de filtragem pode, então, ser usado para realimentar novamente a etapa de extração de feições. Se houver divergência do filtro, procura-se extrair nova reta e repetir o processo.

O processo recursivo definido anteriormente é repetido até que todas as retas disponíveis no modelo do objeto tenham sido utilizadas.

3.4 CONCLUSÃO

Este capítulo procurou agrupar as contribuições propostas neste trabalho. Um modelo matemático do tipo explícito e que permita o tratamento simultâneo dos parâmetros de localização da câmara, foi desenvolvido.

A Filtragem Kalman Estendida (IEKF), que é um método adequado a aplicações dinâmicas, foi utilizada para tratar os erros observacionais.

A natureza recursiva deste filtro levou à definição de um processo iterativo em que as estimativas parciais do vetor estado, obtidas com o processamento de cada observação, são usadas para realimentar a etapa de extração de feições. Este processo recursivo permite uma redução do custo computacional da etapa de extração de feições.

CAPÍTULO IV

4. ANÁLISE EXPERIMENTAL DO MODELO

4.1 INTRODUÇÃO

Neste capítulo será analisado o modelo simultâneo, desenvolvido anteriormente. Inicialmente, será feito um estudo considerando várias situações, usando dados simulados. Posteriormente, uma análise usando dados reais será apresentada, abordando um problema protótipo, em ambiente parcialmente controlado.

4.2 ANÁLISE COM DADOS SIMULADOS.

4.2.1 METODOLOGIA DA SIMULAÇÃO.

O processo de gerar dados simulados, neste caso, consiste em estabelecer um conjunto de retas no espaço objeto, descritas por dois pontos cada uma, arbitrar uma posição para a câmara e calcular as coordenadas imagem dos pontos usando as equações projetivas (Apêndice 1).

A estas coordenadas imagem geradas, que representam os vértices de uma reta são, então, introduzidos erros randômicos, com a finalidade de simular o erro de medida. A partir das coordenadas imagem dos vértices calculam-se as equações das retas na imagem (representação a-b, ou a^*-b^*).

As expressões que permitem calcular a e b (ou a^* e b^*) em função das coordenadas imagem dos vértices são:

$$a = \frac{(y_2 - y_1)}{(x_2 - x_1)} \qquad b = \frac{(y_1.x_2 - x_1.y_2)}{(x_2 - x_1)} \tag{3.3.15}$$

$$a^* = \frac{(x_2 - x_1)}{(y_2 - y_1)} \qquad b^* = \frac{(x_1.y_2 - y_1.x_2)}{(y_2 - y_1)} \tag{3.3.20}$$

onde os elementos x_i e y_i são as coordenadas imagem dos vértices da reta;

Aplicando o procedimento de propagação de covariâncias, e supondo-se que $\sigma_x^2 = \sigma_y^2$, pode-se obter as variâncias dos parâmetros da reta imagem. Desprezando as correlações entre as coordenadas dos vértices pode-se expressar:

$$\sigma_a^2 = \left(\left(\frac{\partial a}{\partial x_1}\right)^2 + \left(\frac{\partial a}{\partial x_2}\right)^2 + \left(\frac{\partial a}{\partial y_1}\right)^2 + \left(\frac{\partial a}{\partial y_2}\right)^2 \right) \sigma_x^2 \tag{4.2.1}$$

$$\frac{\partial a}{\partial x_1} = \frac{(y_2 - y_1)}{(x_2 - x_1)^2} \qquad \frac{\partial a}{\partial x_2} = -\frac{(y_2 - y_1)}{(x_2 - x_1)^2}$$

$$\frac{\partial a}{\partial y_1} = -\frac{1}{(x_2 - x_1)} \qquad \frac{\partial a}{\partial y_2} = \frac{1}{(x_2 - x_1)}$$

Substituindo e agrupando, obtém-se:

$$\sigma_a^2 = \frac{2.(a^2 + 1)}{(x_2 - x_1)^2} \sigma_x^2 \tag{4.2.2}$$

Para o elemento b:

$$\sigma_b^2 = \left(\left(\frac{\partial b}{\partial x_1}\right)^2 + \left(\frac{\partial b}{\partial x_2}\right)^2 + \left(\frac{\partial b}{\partial y_1}\right)^2 + \left(\frac{\partial b}{\partial y_2}\right)^2 \right) \sigma_x^2 \tag{4.2.3}$$

$$\frac{\partial b}{\partial x_1} = \frac{-x_2.a}{(x_2 - x_1)} \qquad \frac{\partial b}{\partial x_2} = \frac{x_1.a}{(x_2 - x_1)}$$

$$\frac{\partial b}{\partial y_1} = \frac{x_2}{(x_2 - x_1)} \qquad \frac{\partial b}{\partial y_2} = \frac{-x_1}{(x_2 - x_1)}$$

Agrupando os termos, chega-se a:

$$\sigma_b^2 = \frac{(x_2^2 - x_1^2)}{(x_2 - x_1)^2} (a^2 + 1)\, \sigma_x^2 \qquad (4.2.4)$$

Para a representação a^*-b^* :

$$\sigma_{a^*}^2 = \left(\left(\frac{\partial a^*}{\partial x_1}\right)^2 + \left(\frac{\partial a^*}{\partial x_2}\right)^2 + \left(\frac{\partial a^*}{\partial y_1}\right)^2 + \left(\frac{\partial a^*}{\partial y_2}\right)^2 \right) \sigma_x^2 \qquad (4.2.5)$$

$$\frac{\partial a^*}{\partial x_1} = \frac{-1}{(y_2 - y_1)} \qquad \frac{\partial a^*}{\partial x_2} = \frac{1}{(y_2 - y_1)}$$

$$\frac{\partial a^*}{\partial y_1} = -\frac{(x_2 - x_1)}{(y_2 - y_1)^2} \qquad \frac{\partial a^*}{\partial y_2} = \frac{(x_2 - x_1)}{(y_2 - y_1)^2}$$

Substituindo e agrupando, obtém-se:

$$\sigma_{a^*}^2 = \frac{2.(a^{*2}+1)}{(y_2 - y_1)^2}\, \sigma_x^2 \qquad (4.2.6)$$

Para o elemento b^*:

$$\sigma_{b^*}^2 = \left(\left(\frac{\partial b^*}{\partial x_1}\right)^2 + \left(\frac{\partial b^*}{\partial x_2}\right)^2 + \left(\frac{\partial b^*}{\partial y_1}\right)^2 + \left(\frac{\partial b^*}{\partial y_2}\right)^2 \right) \sigma_x^2 \qquad (4.2.7)$$

$$\frac{\partial b^*}{\partial x_1} = \frac{y_2}{(y_2 - y_1)} \qquad \frac{\partial b^*}{\partial x_2} = \frac{-y_1}{(y_2 - y_1)}$$

$$\frac{\partial b^*}{\partial y_1} = \frac{-y_2.a^*}{(y_2 - y_1)} \qquad \frac{\partial b^*}{\partial y_2} = \frac{y_1.a^*}{(y_2 - y_1)}$$

Agrupando os termos, chega-se a:

$$\sigma_{b^*}^2 = \frac{(y_2^2 - y_1^2)}{(y_2 - y_1)^2} (a^{*2} + 1)\, \sigma_x^2 \qquad (4.2.8)$$

Com as expressões desenvolvidas acima, é possível estimar os elementos a e b (ou a^* e b^*) suas respectivas variâncias σ_a^2 e σ_b^2, a partir das coordenadas imagem dos vértices da reta, e do desvio padrão dos erros randômicos introduzidos nestas coordenadas (σ_x).

4.2.2. CARACTERÍSTICAS DOS EXPERIMENTOS.

Para analisar o efeito da variação das características da câmara, como distância focal, tamanho do pixel, erros sistemáticos, etc.., serão conduzidos vários experimentos. Os valores típicos para a maioria dos experimentos são:

- a distância focal da câmara é de 20mm;
- o tamanho do pixel é de 10x10 μm^2, e a área de imageamento é de 5x4 mm^2;
- o desvio padrão dos erros a serem introduzidos nas coordenadas imagem é de 3 μm. Este erro de 3 μm corresponde a 1/3 do tamanho do pixel. Assume-se esta precisão a partir da constatação de que é possível extrair retas da imagem com precisão sub-pixel (Liu e Huang, 1991);
- adota-se como objeto um cubo com 70mm de lado;
- a distância entre a câmara e o objeto varia de 0.6 m a 1.8 m;
- os critérios de convergência das iterações no IEKF são: 1.0e-6 rad para as rotações e 0.001 mm para as translações;

Cada experimento é repetido 50 vezes; são gerados os dados simulados e a filtragem é executada, sendo obtido o valor filtrado final. O erro verdadeiro no estado é, então, calculado como a diferença entre o valor filtrado e o valor verdadeiro, conhecido da simulação. São obtidos, portanto, os erros verdadeiros para o mesmo experimento repetido 50 vezes, com erros de medida diferentes em cada repetição.

Finalmente, é calculado o erro médio quadrático dos erros verdadeiros nos 50 experimentos, através da expressão (4.2.9):

$$m = \sqrt{\frac{\sum \varepsilon_i^2}{n}} \qquad (4.2.9)$$

onde: ε_i é o erro verdadeiro no i-ésimo experimento;
n é o número de experimentos;
m é o erro médio quadrático;

Serão realizados vários experimentos, visando verificar o comportamento

do modelo matemático diante de várias situações. A tabela 4.2.1 resume as características dos experimentos com dados simulados.

Tabela 4.2.1 Características dos experimentos com dados simulados.

Seção	Características
4.2.3	Caso típico - Analise de convergência
4.2.4	Comparação dos resultados do IEKF com o Filtro em Batch;
4.2.5	Efeito da variação da distância câmara-objeto;
4.2.6	Efeito da distância focal da câmara;
4.2.7	Estudo de configurações deficientes;
4.2.8	Efeito da variação da posição do objeto na imagem;
4.2.9	Efeito do erro observacional;
4.2.10	Efeito da avaliação incorreta do erro observacional
4.2.11	Efeito da variação na estimativa predita inicial;
4.2.12	Efeito de erros grosseiros nas observações;
4.2.13	Efeito de erros sistemáticos não eliminados;
4.2.14	Sequência de imagens (caso dinâmico);

4.2.3 Análise de um caso típico.

Neste experimento procura-se mostrar os valores numéricos utilizados no experimento; apresenta-se ainda, uma análise do comportamento do IEKF, com relação à sua convergência, à medida em que as observações são processadas.

O objeto a ser utilizado como exemplo, e cujas arestas no espaço objeto são conhecidas, é um cubo representado por um modelo "wireframe" (não é realizada a eliminação das linhas escondidas, nos experimentos com dados simulados), de 70 mm de lado, posicionado a, aproximadamente, 86 cm de uma câmara com distância focal 20 mm. A figura 4.2.1 (a) mostra a imagem simulada, composta de 12 retas, do cubo "wireframe".

Na tabela 4.2.2 são mostrados os valores de a e b e suas respectivas variâncias, obtidas por propagação, correspondentes ao cubo mostrado na figura 4.2.1.(a). Nesta tabela, o código **g**, indicado ao lado dos vértices que compõem a reta, indica que os parâmetros são **a** e **b**, enquanto que o código **f** indica que os parâmetros são $\mathbf{a}^*$ e $\mathbf{b}^*$.

Tabela 4.2.2 Parâmetros da reta e respectivas variâncias.
g → a e b f → a* e b*

RETA	a (a*)	b (b*)mm	σ_a^2	σ_b^2
32 g	0.389645	1.1342	0.0000097	0.0000067
21 g	-0.559306	1.3702	0.0000311	0.0000168
10 g	0.362515	0.3385	0.0000103	0.0000080
03 g	-0.461155	0.1101	0.0000256	0.0000215
36 f	-0.121935	-1.1256	0.0000068	0.0000046
65 g	0.334066	-0.6220	0.0000086	0.0000069
52 f	-0.160486	0.4460	0.0000065	0.0000053
67 g	-0.353579	-1.3123	0.0000234	0.0000093
74 g	0.320839	-1.2748	0.0000095	0.0000070
45 g	-0.437648	-0.2236	0.0000278	0.0000350
41 f	-0.170815	1.2430	0.0000075	0.0000054
07 f	-0.139685	-0.2463	0.0000078	0.0000064

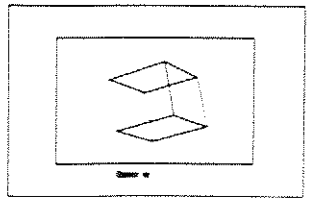
Tabela 4.2.3 Resultados obtidos com o IEKF.

	VALORES PREDITOS			VALORES FILTRADOS		
	Estado Predito	Desvio Padrão	Erro Verdad.	Estado Filtrado	Desvio Padrão	Erro m. quad.
κ	2.720	0.08660	-0.080	2.80068	0.000963	0.00112
ϕ rad	0.560	0.08660	0.090	0.47022	0.000981	0.00149
ω	-1.100	0.08660	0.060	-1.16081	0.001032	0.00127
X_c	392.0	10.0	-8.000	400.331	0.739	1.147
Y_c mm	710.0	10.0	10.000	700.718	0.789	0.969
Z_c	312.0	10.0	-8.000	319.645	0.739	0.950

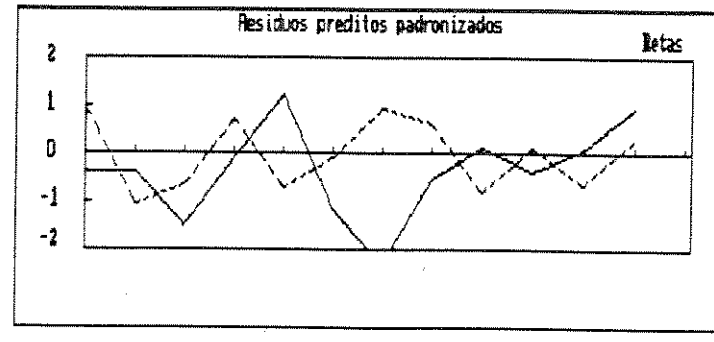
No IEKF as observações são processadas sequencialmente, a partir de uma estimativa predita. Com isto, é possível verificar, a cada nova reta introduzida, a convergência da solução. A Tabela 4.2.3 indica o valor verdadeiro para o vetor estado (parâmetros da câmara), a estimativa predita inicial, a matriz covariância predita e os valores filtrados finais do vetor estado e da matriz covariância, obtidos após o processamento de 12 retas.

Uma verificação mais imediata do efeito de convergência do IEKF, poderá ser obtida analisando os gráficos condensados na Figura 4.2.1. Na Figura 4.2.1 (b) são mostrados os resíduos preditos padronizados; este elemento é essencial para verificar se há divergência do filtro. Na figura 4.2.1.(b) verifica-se que apenas em uma reta o resíduo predito padronizado foi maior que

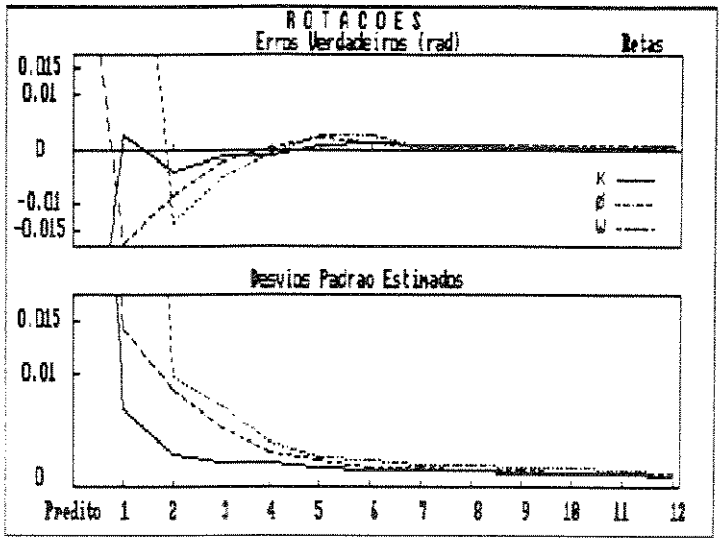
1, o que é aceitável, pois isto indica que os resíduos possuem as propriedades estatísticas esperadas.



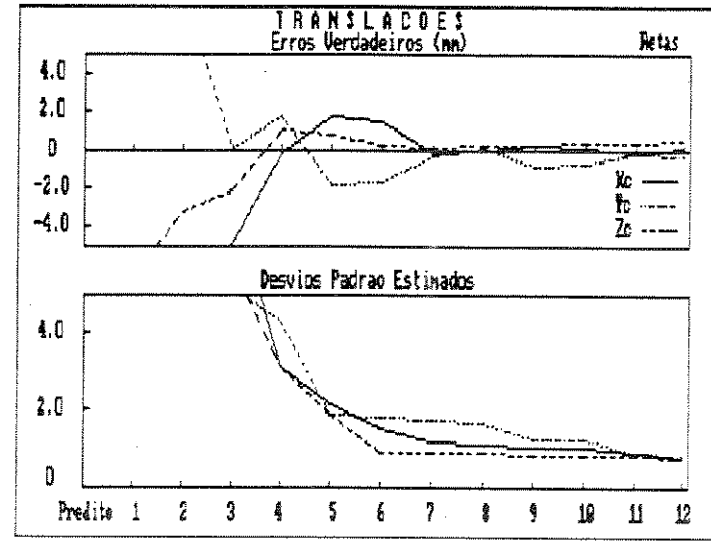
(a)



(b)



(c)



(d)

Figura 4.2.1 Gráficos correspondentes à convergência do IEKF.

Os gráficos 4.2.1.(c) e 4.2.1.(d) mostram os erros verdadeiros e os desvios padrão estimados, começando com o valor predito, e passando por cada uma das retas utilizadas, tanto para as rotações (c) quanto para translações (d).

A análise destes gráficos e da tabela 4.2.3 permite concluir:

. o filtro convergiu, como pode-se inferir através da análise dos erros verdadeiros e dos resíduos padronizados;

. o erro final no estado é da ordem de 3' para as rotações e de 0.9 mm para

as translações;

- o valor do desvio padrão estimado é coerente com o erro verdadeiro obtido;
- com relação às rotações, verifica-se que o parâmetro κ (rotação em z) converge rapidamente, logo com o processamento da 1ª e 2ª observações. Com o processamento de 4 retas as rotações atingem um valor satisfatório e com 7 retas alcançam um valor equivalente ao valor final;
- em relação às translações, observa-se um comportamento semelhante dos três elementos. Embora a matriz covariância indique uma convergência na 6ª reta, observa-se que os erros verdadeiros apresentam valores aceitáveis já após o processamento da 4ª reta, como nas rotações.

4.2.4 Comparação dos resultados obtidos com o IEKF com os resultados do Filtro em Batch

Nesta seção serão comparados os resultados obtidos com o Filtro Kalman Estendido Iterativo (IEKF) e com o Filtro em Batch (equivalente ao MMQ), para verificar as diferenças entre o método de estimação sequencial e o simultâneo.

Devido ao modelo não linear, sabe-se que o IEKF pode produzir resultados espúrios, especialmente quando a estimativa predita é muito grosseira ou quando existe algum erro grosseiro nas observações. A tabela 4.2.4 apresenta os resultados obtidos com o Filtro em Batch (MMQ), onde todas as observações foram processadas simultâneamente; os resultados do IEKF são reproduzidos da tabela 4.2.3, permitindo uma comparação entre os métodos.

Tabela 4.2.4 Comparação entre o Filtro em Batch e o IEKF

	FILTRO EM BATCH			I E K F		
	Estado Filtrado	Desvio Padrão	Erro m. quad.	Estado Filtrado	Desvio Padrão	Erro m. quad.
κ	2.80035	0.00134	0.00118	2.80068	0.000963	0.00112
ϕ rad	0.47011	0.00136	0.00150	0.47022	0.000981	0.00149
ω	-1.16055	0.00145	0.00123	-1.16081	0.001032	0.00127
X_c	400.201	1.044	1.120	400.331	0.739	1.147
Y_c mm	700.579	1.051	1.009	700.718	0.789	0.969
Z_c	319.579	1.046	0.821	319.645	0.739	0.950
Tempo de process.	1.04 s			2.30 s		

Analisando a tabela 4.2.4 pode-se observar que:

- os desvios padrão estimados (matriz covariância) possuem magnitude semelhante nos dois experimentos, embora ligeiramente menores no IEKF;
- os erros médios quadráticos também possuem ordem de grandeza análoga, apesar de pequenas diferenças entre os dois métodos de estimação;
- na última linha da tabela 4.2.4 são mostrados os tempos de processamento do Filtro em Batch e do IEKF. Verifica-se que o tempo de processamento do IEKF é, praticamente, o dobro do Filtro em Batch. O cálculo do IEKF, entretanto, pode ser substancialmente acelerado utilizando formas fatoradas da matriz covariância (Bierman, 1977);

Isto permite concluir que o IEKF e o Filtro em Batch (MMQ) são equivalentes em relação à precisão obtida, para a situação estudada.

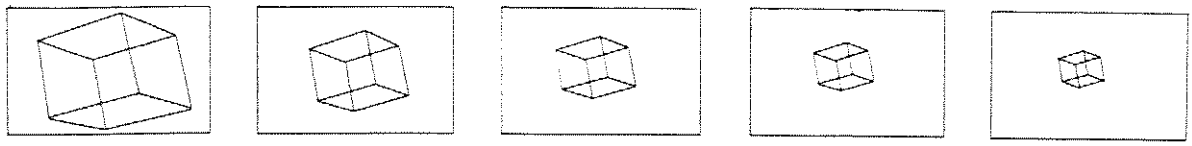
4.2.5 Efeito da variação da distância câmara-objeto.

Neste ítem serão analisados os efeitos causados pela variação dadistância câmara-objeto.

Na figura 4.2.2 são apresentados os gráficos comparando os resultados obtidos deslocando-se a câmara da posição inicial, a 60cm da câmara, até 1.80m. No gráfico 4.2.2.(b) são mostrados os erros médios quadráticos nas estimativas filtradas finais (50 experimentos) para as rotações e seus respectivos desvios padrão estimados, para os vários casos estudados. Analogamente, no gráfico 4.2.2.(c), são apresentadas as mesmas estatísticas para as translações.

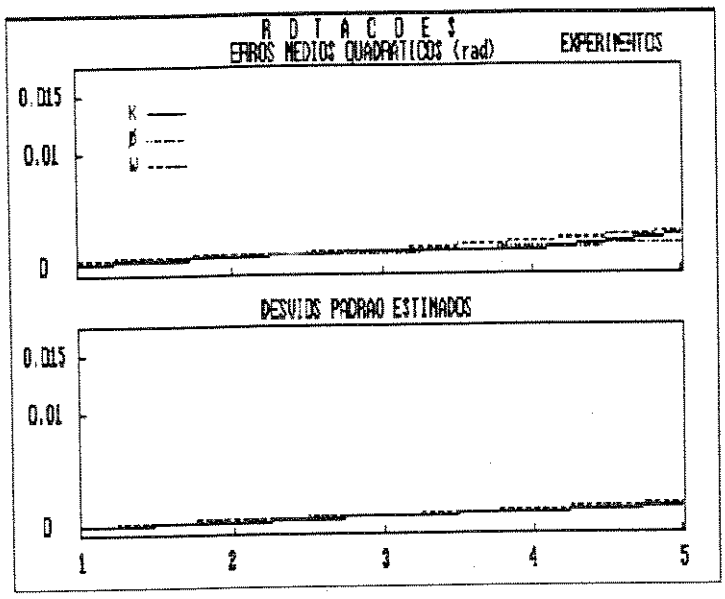
Foram realizados 5 experimentos, considerando uma câmara de focal 20 mm, posicionando-a à distância de 60cm, 88cm, 1.10m, 1.40m e 1.82m, sucessivamente.

Com relação à convergência do filtro, observou-se que, nos experimentos em que a câmara estava mais próxima do objeto (60cm e 80cm), os elementos κ e Zc convergem antes dos demais. À medida que a câmara se afasta (o cubo fica menor na imagem) esta característica fica menos definida.

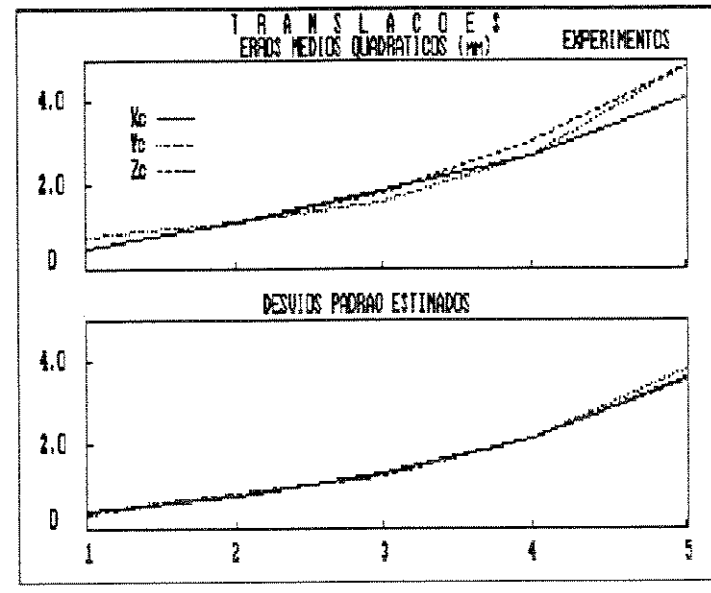


1 2 3 4 5

(a)



(b)



(c)

Figura 4.2.2 Efeito da variaçâ: da distância nos experimentos com o IEKF.

Observando os gráficos apresentados na Figura 4.2.3, pode-se concluir que:

. os elementos de rotação possuem praticamente a mesma precisão, e esta precisão diminui com o aumento da distância, de maneira quase que linear;

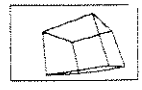
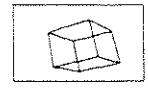
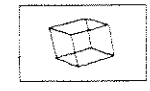
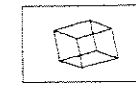
. esta mesma característica observa-se nas translações, embora a perda de precisão, neste caso, seja mais acentuada com o aumento da distância;

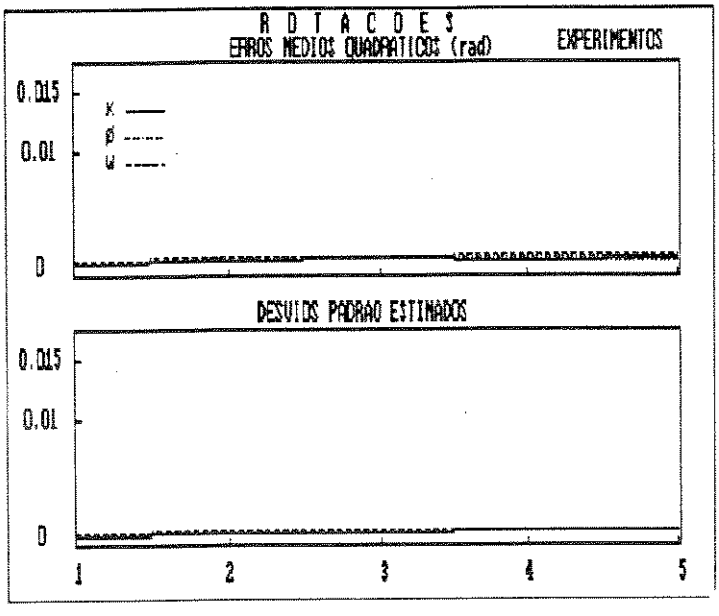
. uma diminuição na escala da imagem, para o mesmo objeto e mesma distância focal, implica em diminuição da precisão. Diminuir a escala equivale a aumentar a distância câmara-objeto. Uma diminuição nas dimensões do objeto causaria idêntico efeito;

4.2.6 EFEITO DA VARIAÇÃO DA DISTÂNCIA FOCAL DA CÂMARA.

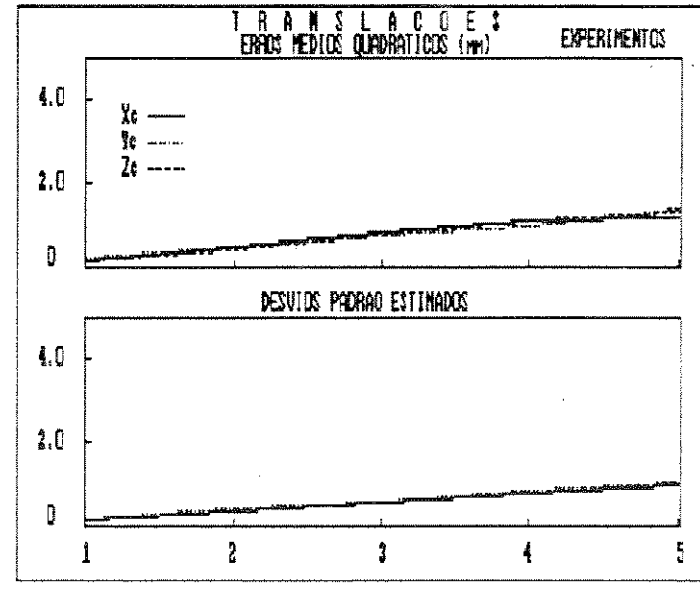
Numa segunda série de experimentos procurou-se variar a distância focal da câmara em conjunto com a variação da distância, de tal maneira que a escala da imagem se mantivesse fixa. Os resultados destes experimentos são mostrados na Figura 4.2.3, e a tabela 4.2.5 resume suas características.

Tabela 4.2.5 Características dos experimentos com a distância focal.

Nº	1	2	3	4	5
focal (mm)	5.	10.	15.	20.	25.
distância (m)	0.21	0.43	0.65	0.87	1.08
Escala	1:43	1:43	1:43	1:43	1:43
Imagens					



(a)



(b)

Figura 4.2.3 Efeito da variação da focal nos experimentos com o IEKF.

Analisando os gráficos na figura 4.2.3 pode-se observar que:

- imagens com a mesma escala, mas com distâncias focais diferentes, conduzem a resultados diferentes. Quanto menor a distância focal, maior é a precisão;
- os erros nas translações aumentam mais sensivelmente com o aumento da distância focal do que os das rotações;

. este aumento dos erros nas translações pode ser explicado pela piora na qualidade geométrica das intersecções que definem o centro perspectivo; quanto maior a distância focal menor é o ângulo de intersecção;

4.2.7 CONFIGURAÇÕES DEFICIENTES.

Neste ítem serão apresentados 5 experimentos visando investigar o efeito causado pela configuração das retas. O primeiro experimento é estabelecido como um padrão, e trata-se de um cubo "wireframe" (12 retas) a 88 cm de distância, imageado por uma câmara com distância focal de 20 mm.

O 2º e 3º experimentos são configurações deficientes (ou degeneradas) porque todas as retas são paralelas. Estas retas foram geradas a partir do cubo original, selecionando-se as retas horizontais, para o 2º experimento e as verticais para o 3º.

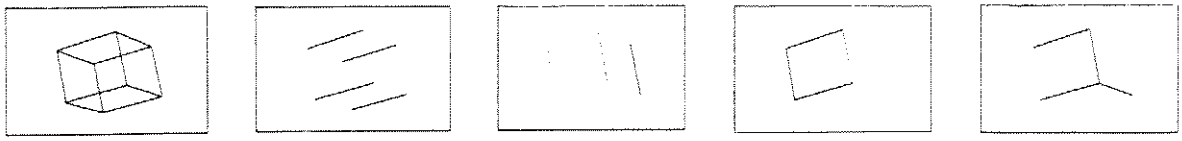
O 4º experimento foi realizado selecionando-se as 4 retas que compõem uma das faces do cubo, enquanto que no 5º experimento foram selecionadas três retas formando um terno ortogonal, mais uma quarta reta para estabelecer um sistema reduntante de equações.

Na figura 4.2.4 (a) são mostradas as imagens simuladas para cada uma das 5 configurações. Na figura 4.2.4.(b) são apresentados os resultados obtidos para as rotações; os erros médios quadráticos de 50 experimentos e o desvios padrão a partir da matriz covariância filtrada. Na figura 4.2.4.(b) são mostrados os resultados para as translações.

Da análise da figura 4.2.4 pode-se concluir que:

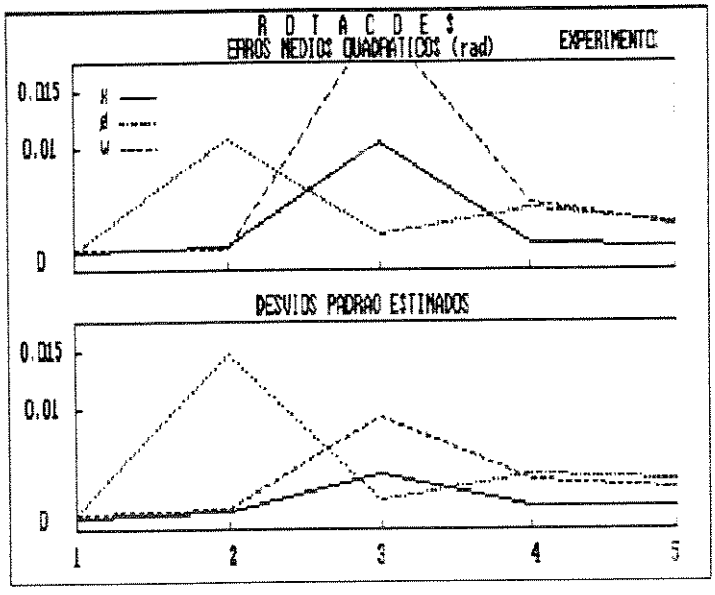
. em nenhum dos experimentos verificou-se divergência do fitro;
. os melhores resultados foram conseguidos com o cubo completo, o que explica-se pela configuração geométrica e pelo número de retas;
. no experimento 2, em que as retas são horizontais e paralelas, verifica-se que os parâmetros de rotação κ e ω possuem o mesmo nível de precisão que no experimento 1, com o cubo; o parâmetro Zc é um pouco menos preciso, enquanto que Xc e Yc apresentam erros de grande magnitude;
. no experimento 3, com retas verticais, a precisão de κ sofre uma pequena deterioração; em relação ao experimento 2 existe uma inversão de situação: o parâmetro ϕ agora é mais preciso enquanto que ω perdeu precisão; contudo, o erro médio quadrático de ω é bastante superior ao seu desvio padrão estimado. Nas translações aumentou significativamente o erro em Zc, mas diminuiram os erros em Xc e Yc;

. foram observados níveis de precisão semelhantes nos experimentos 4 e 5, com 4 retas; estes níveis de precisão são superiores aos alcançados com os configurações deficientes 2 e 3 mas são ainda inferiores aos conseguidos com o cubo, no experimento 1;

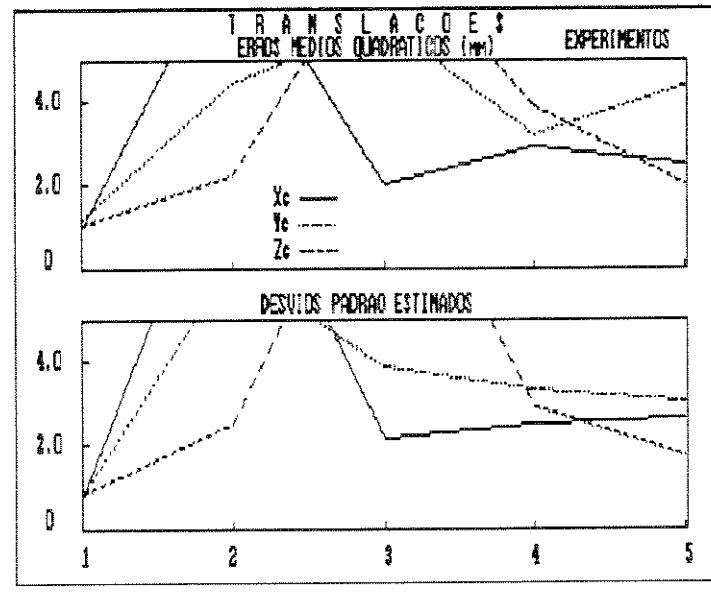


1 2 3 4 5

(a)



(b)



(c)

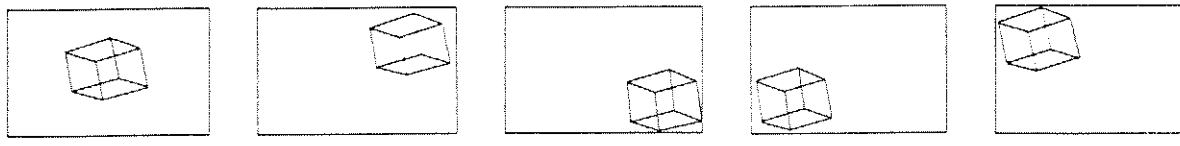
Figura 4.2.4 Resultados obtidos com diferentes configurações.

4.2.8 VARIAÇÃO DA POSIÇÃO DO OBJETO NA IMAGEM.

Neste ítem serão desenvolvidos 5 experimentos, com a finalidade de verificar o efeito causado pela concentração de linhas em uma parte da imagem. Para a realização destes experimentos foram simuladas imagens com um cubo "wireframe" a cerca de 1,10m de uma câmara com distância focal de 20mm.

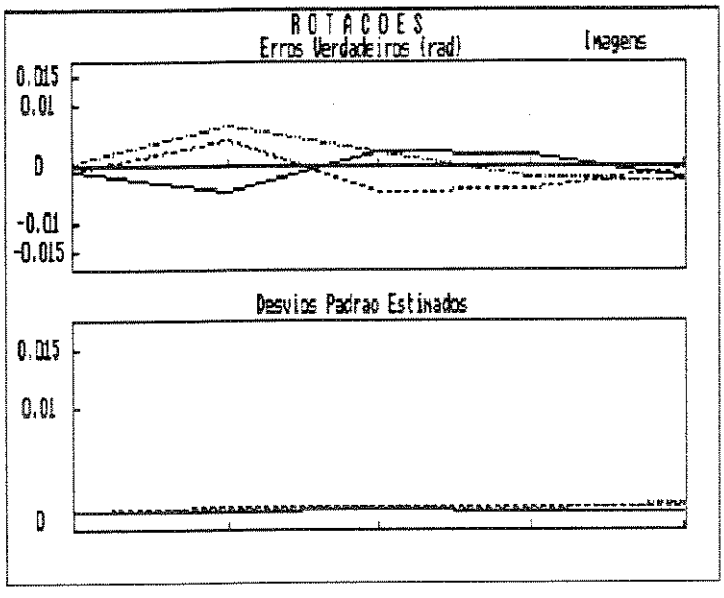
No experimento 1, considerado como uma situação ideal, o cubo está no centro da imagem (figura 4.2.5.(a)), enquanto que nos demais experimentos o cubo foi deslocado para os cantos da imagem, mediante pequenas rotações introduzidas na câmara.

O efeito desta variação na posição relativa das feições em uma porção da imagem pode ser verificado nos gráficos 4.2.5.(b) e 4.2.5.(c).

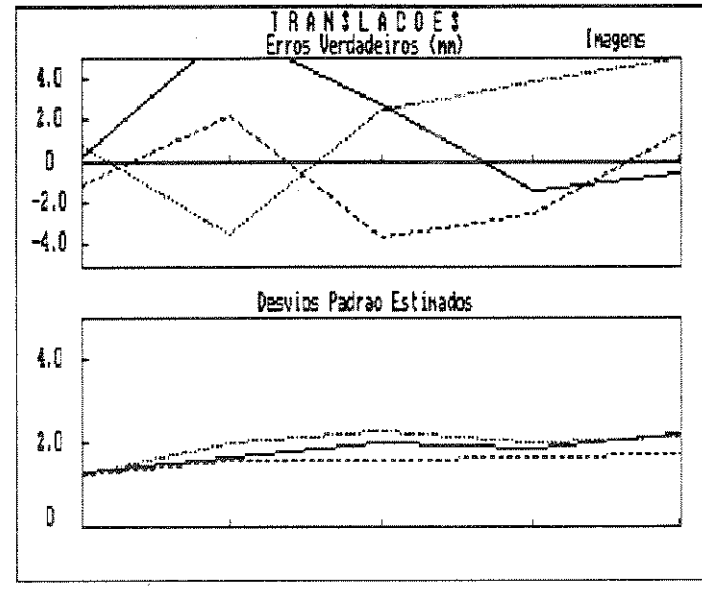


1 2 3 4 5

(a)



(b)



(c)

Figura 4.2.5 Variação da posição do cubo na imagem.

A análise dos experimentos resumidos na figura 4.2.5, permite observar que:

. os melhores resultados foram obtidos com o cubo no centro da imagem;

. as translações são mais afetadas pelas mudanças na posição do cubo; observa-se que, embora os desvios estimados sofram uma variação pequena, os erros médios quadráticos indicam um erro maior do que o estimado; o mesmo ocorre para as rotações, mas com menor efeito;

. o parâmetros Zc é o menos afetado pelas variações na posição do cubo;
. nos experimentos com o cubo nos cantos da imagem observou-se que alguns resíduos padronizados foram superiores a 2, e em alguns casos, superiores a 3, o que indicaria divergência do filtro. Na realidade, as situações em que as feições se agrupam numa pequena parte da imagem, podem ser encaradas também como configurações deficientes, porque os ângulos de intersecção são menores;

4.2.9 EFEITO DA MAGNITUDE DOS ERROS OBSERVACIONAIS.

Um dos dados mais importantes na estimação de parâmetros usando Filtragem Kalman, é a qualidade das observações. A magnitude dos erros observacionais reflete-se diretamente na qualidade do estado filtrado final.

Os experimentos a serem apresentados a seguir procuram analisar o efeito dos erros observacionais variando sua magnitude: na primeira série de experimentos os erros nos eixos x e y são considerados iguais; na segunda série, adota-se uma hipótese mais realista, considerando o erro em x maior que o erro em y.

Na primeira série de experimentos foram geradas 5 imagens com o cubo a 88 cm de uma câmara de distância focal 20mm. A tabela 4.2.6 mostra o valor do desvio padrão do erro introduzido nas observações, para cada experimento.

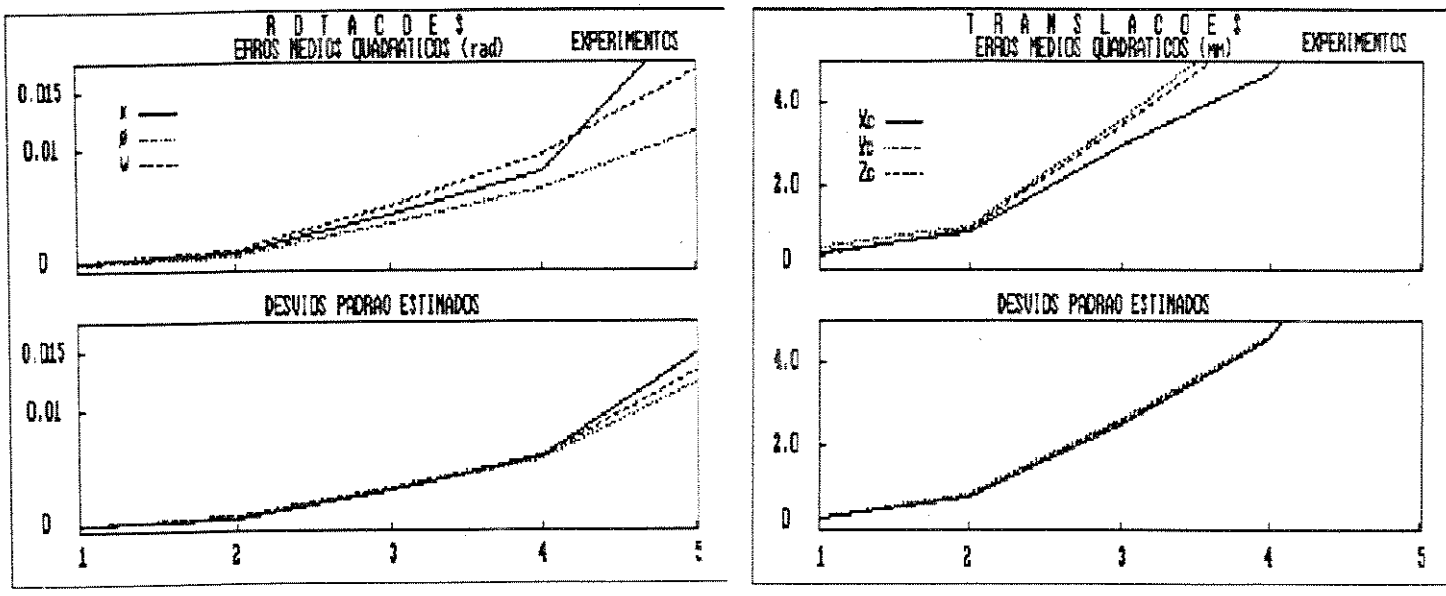
Tabela 4.2.6 Desvios padrão dos erros observacionais nos vértices das retas. ($\sigma x = \sigma y$)

Experim.	1	2	3	4	5
σx (μm)	1	3	10	20	50
σy (μm)	1	3	10	20	50

A figura 4.2.6 mostra o resultado dos experimentos utilizando os dados resumidos na tabela 4.2.6.

Pode-se concluir, com uma análise dos gráficos que:
. como era esperado, a maior precisão foi obtida com o experimento em que foram introduzidos os menores erros observacionais (experimento 1, com 1 μm);
. observa-se o aumento proporcional do erro no estado filtrado à medida que aumenta o erro observacional;

. em nenhum dos casos observou-se divergência do filtro;



(a) (b)

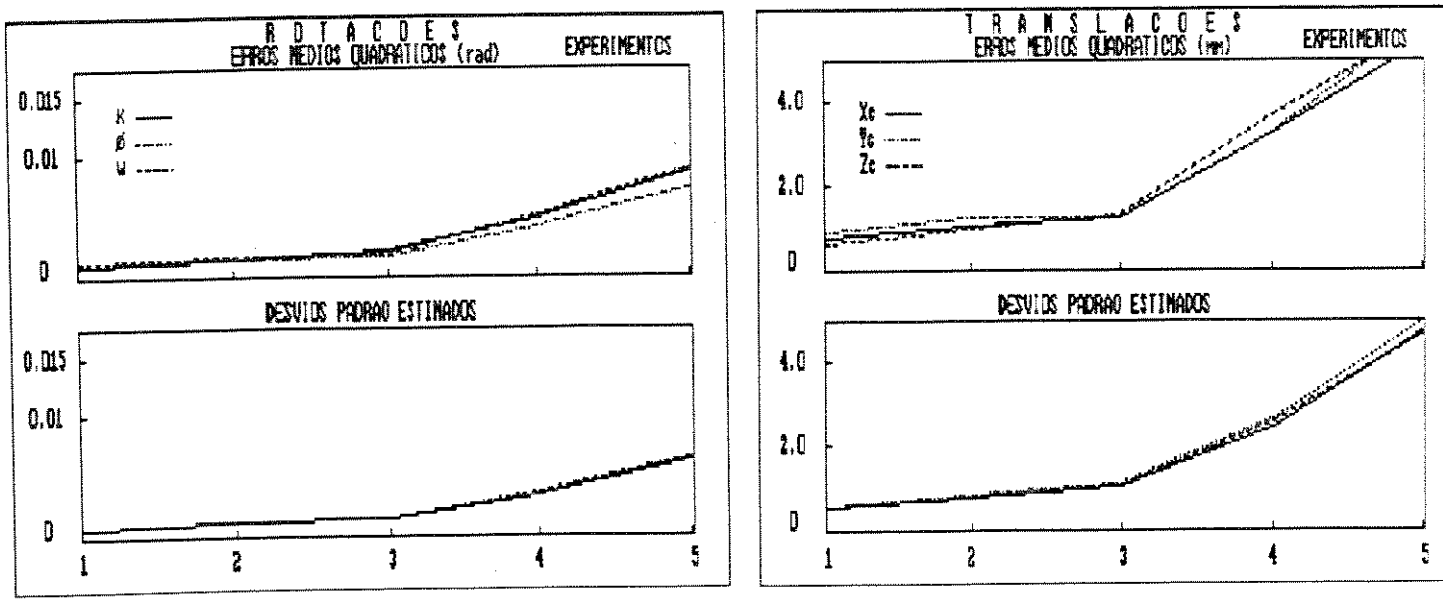
Figura 4.2.6 Efeito da magnitude dos erros observacionais.

Os experimentos anteriores mostraram que o aumento do erro observacional provoca aumento no erro do estado, mas sem provocar divergência do filtro.

Uma análise semelhante pode ser conduzida, agora considerando os erros em x diferentes dos erros em y. A tabela 4.2.7 mostra os desvios padrão dos erros observacionais utilizados para realizar 5 experimentos. Na figura 4.2.7 são apresentados os resultados obtidos com os dados gerados considerando estes erros.

Tabela 4.2.6 Desvios padrão dos erros observacionais nos vértices das retas. ($\sigma x \neq \sigma y$)

Experim.	1	2	3	4	5
σx (μm)	5	7	10	25	50
σy (μm)	2	3	4	10	20



(a) (b)

Figura 4.2.7 Efeito da magnitude dos erros observacionais para $\sigma x \neq \sigma y$.

Analisando a figura 4.2.7 e comparando-a com a figura 4.2.6, verifica que não uma perda sensível de precisão quando os erros observacionais são diferentes em x e y. Não ocorre nenhuma diferença significativa entre os elementos do vetor estado, devido a esta diferença.

4.2.10 Efeito da avaliação incorreta do erro observacional.

Um dos dados mais importantes no processo de estimação usando Filtragem Kalman, é a qualidade das observações. Uma das causas frequentes de divergência do filtro é a super-estimação da precisão das observações.

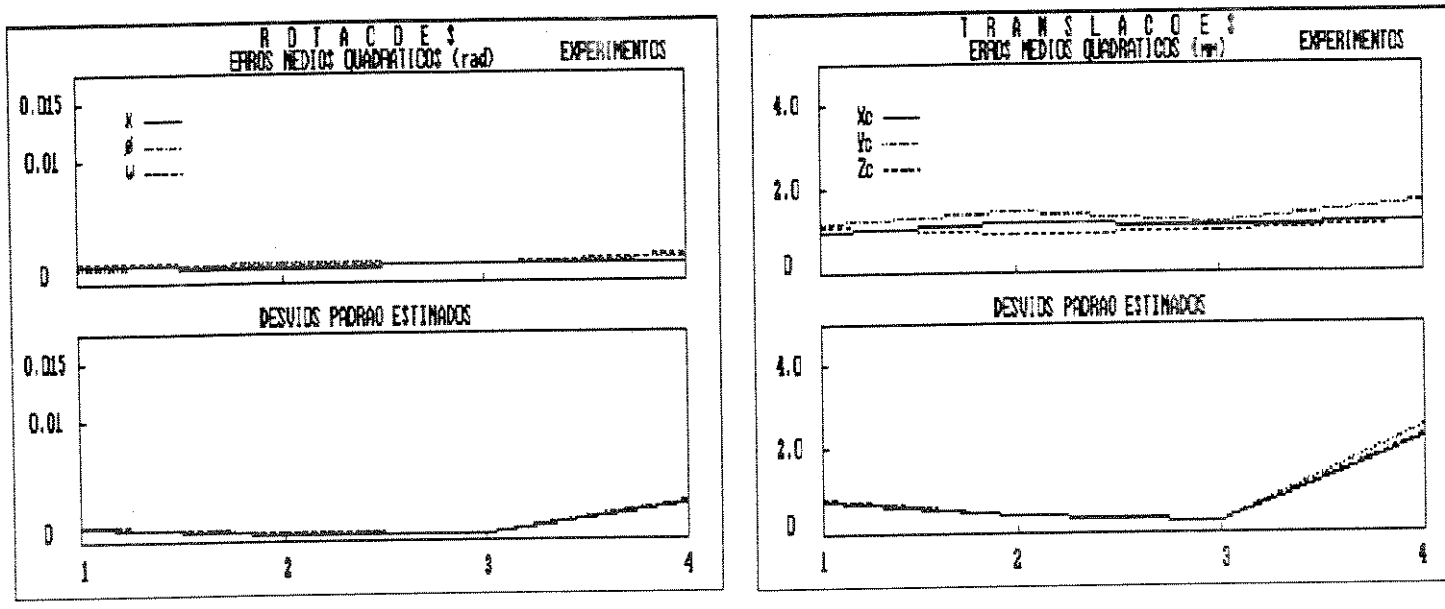
Nos experimentos a seguir procura-se investigar os efeitos de uma avaliação incorreta da variância das observações. Este estudo é importante para prever situações potencialmente problemáticas quando forem processados dados reais.

Foram realizados 4 experimentos, considerando um cubo a 88cm de distância de uma câmara de focal 20mm. Foram introduzidos erros randômicos nos vértices das retas, cujo desvio padrão é de 3μm, tanto em x quanto em y. As mesmas observações foram usadas em todos os experimentos, variando apenas a variância estimada destas observações.

No experimento 1 foram obtidas as variâncias dos parâmetros a e b, da

reta, considerando-se o desvio padrão real de 3μm nas coordenadas imagem dos vértices. No experimento 2 a precisão das observações foi super-estimada, supondo-se que as coordenadas dos vértices das retas na imagem estivessem com desvio padrão de 1.5μm (metade do valor real). No experimento 3 estas variâncias foram ainda menores, com um valor de 1μm (1/3 do valor real). No experimento 4, finalmente, adotou-se uma estratégia inversa, subestimando a precisão das observações, adotando-se um valor de 9μm (3 vezes o valor real).

Os valores finais dos erros médios quadráticos no vetor estado e os desvios padrão estimados para as rotações e translações são apresentados na figura 4.2.8. (a) e (b), respectivamente. Na figura 4.2.9 são mostradas as sequências dos resíduos preditos padronizados, para cada um dos experimentos, com o objetivo de avaliar a possibilidade de detecção de divergência.



(a) (b)

Figura 4.2.8 Efeito da estimação da variância das observações.

Analisando a figura 4.2.8 pode-se verificar que:

- os erros médios quadráticos nas rotações praticamente não se alteram com as mudanças nas variâncias das observações; um pequeno aumento no erro foi verificado no 4º experimento, nas rotações ϕ e ω; os desvios padrão estimados diminuem, especialmente no 3º experimento, e aumentam no 4º experimento, onde a precisão das observações foi sub-estimada;
- os erros médios quadráticos nas translações mantém-se praticamente os mesmos nos 3 primeiros experimentos, aumentando no último; os desvios

padrão estimados diminuem nos experimentos 2 e 3, e aumentam no 4º, de maneira proporcional às variações introduzidas nas variâncias das observações;

- devido à proposital super-estimativa da precisão das observações nos experimentos 2 e 3 verifica-se que, nestes casos, os desvios padrão estimados são menores que os erros médios quadráticos, tanto para as rotações quanto para as translações;

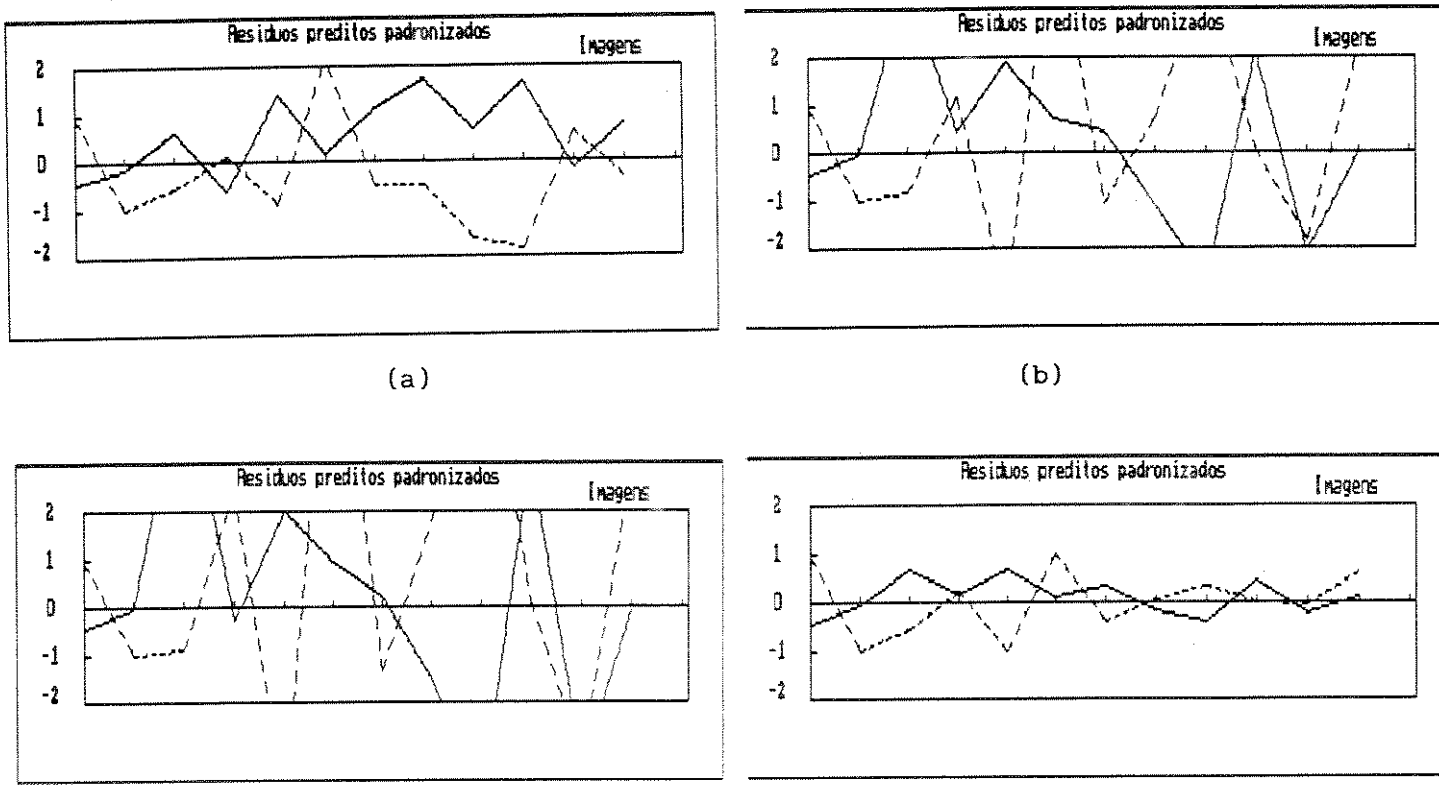


Figura 4.2.9 Resíduos preditos padronizados para os 4 experimentos.

A análise dos gráficos dos resíduos permite verificar que nos experimentos 2 e 3 os resíduos padronizados assumem valores muito acima de 1; isto ocorre porque a matriz covariância dos resíduos torna-se menor com as variâncias super-estimadas. Estes efeitos de aumento dos resíduos padronizados e diminuição da matriz covariância do estado filtrado, verificados na figura 4.2.8, poderiam ser interpretados como divergência do filtro. No experimento 4, verifica-se que os resíduos padronizados são todos menores que 1, embora esta sub-estimação da variância das observações também tenha produzido erros verdadeiros maiores.

4.2.11 EFEITO DA VARIAÇÃO NA ESTIMATIVA PREDITA INICIAL.

Uma das características básicas do IEKF é a sua natureza recursiva, iniciando-se o processo com base em um estado predito inicial e em uma matriz covariância predita. Os experimentos a serem discutidos nesta seção procuram investigar os efeitos causados pelas variações nestas estimativas a priori.

Foram realizados 7 experimentos, considerando-se um cubo a 88 cm de uma câmara de distância focal 20mm. Foram usados os mesmos parâmetros verdadeiros da câmara, para simular os 7 experimentos, variando-se apenas o estado predito e a matriz covariância predita. Cada experimento foi executado 50 vezes, calculando-se, então os erros verdadeiros para cada execução e, finalmente, os erros médios quadráticos dos parâmetros.

Os dados a priori para cada experimento estão resumidos na tabela 4.2.8. Na coluna ε_v estão indicados os erros verdadeiros na estimativa predita, isto é, as diferenças entre o estado predito e o valor verdadeiro do estado, conhecido da simulação. Na coluna σx_o estão indicados os desvios padrão preditos.

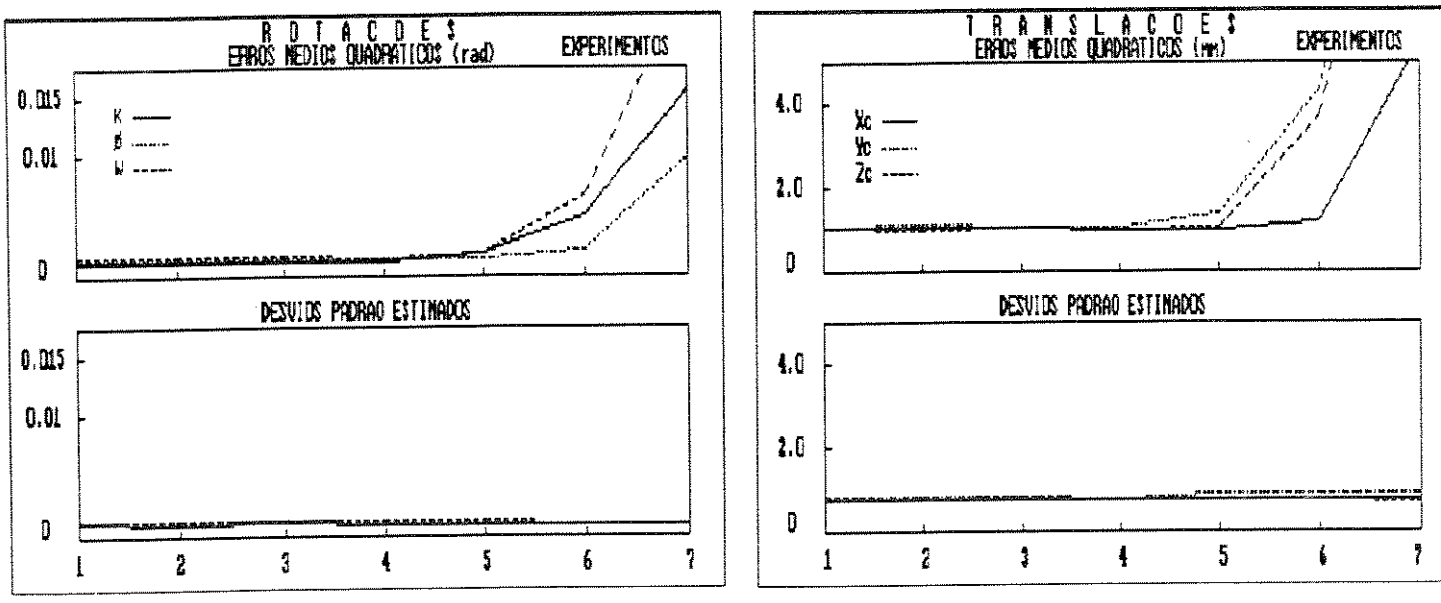
Tabela 4.2.8. Estimativas preditas para os 7 experimentos.

	1		2		3	
	σx_o	ε_v	σx_o	ε_v	σx_o	ε_v
κ	0.08	-0.08	0.17	-0.08	0.17	-0.08
ϕ (rad)	0.08	0.09	0.17	0.09	0.17	0.09
ω	0.08	0.06	0.17	0.06	0.17	0.06
Xc	10.0	-8.00	10.0	-8.00	20.0	-8.00
Yc (mm)	10.0	10.00	10.0	10.00	20.0	10.00
Zc	10.0	-8.00	10.0	-8.00	20.0	-8.00

	4		5		6		7	
	σx_o	ε_v	σx_o	ε_v	σx_o	ε_v	σx_o	ε_v
κ	0.17	-0.08	0.08	-0.17	0.08	-0.26	0.08	-0.43
ϕ (rad)	0.17	0.07	0.08	0.17	0.08	0.26	0.08	0.43
ω	0.17	-0.06	0.08	0.17	0.08	0.26	0.08	0.43
Xc	20.0	8.0	10.0	-20.0	10.0	-30.0	10.0	-50.0
Yc (mm)	20.0	-10.0	10.0	20.0	10.0	30.0	10.0	50.0
Zc	20.0	10.0	10.0	-20.0	10.0	-30.0	10.0	-50.0

Nos experimentos 1, 2 e 3 apenas foram alteradas as variâncias preditas, mantendo-se o estado predito. No 4º experimento alterou-se o estado predito,

mas mantendo o valor dentro do intervalo definido pela variância predita. Nos experimentos 5, 6 e 7, procurou-se simular uma situação onde a estimativa predita estivesse fora do intervalo de $1.\sigma$ (desvio padrão predito); no experimento 5 o erro verdadeiro é de $2.\sigma$, no experimento 6 é de $3.\sigma$ e no 7º é de $5.\sigma$.



(a) (b)

Figura 4.2.10 Efeito das estimativas preditas.

Os resultados obtidos com estas variações nas estimativas preditas estão resumidos nas figuras 4.2.10 e 4.2.11. Na figura 4.2.10 estão representados os erros médios quadráticos do estado filtrado final e desvios padrão estimados, para todos os experimentos.

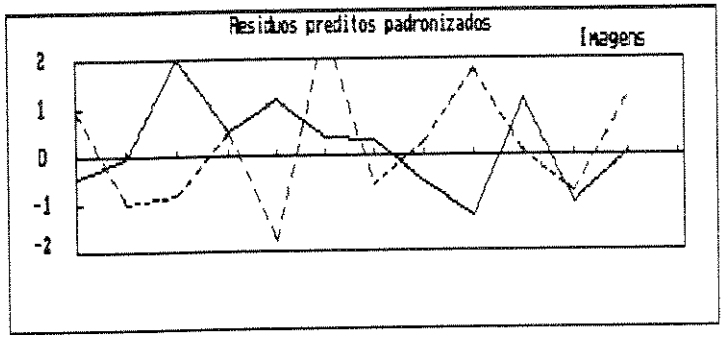
Na figura 4.2.11 são apresentados os gráficos de evolução dos resíduos preditos padronizados, em cada observação, para os experimentos 1 (a), 5 (b), 6 (c) e 7 (d).

A análise dos gráficos 4.2.10 e 4.2.11 em confronto com a tabela 4.2.8 permitem estabelecer as seguintes conclusões:

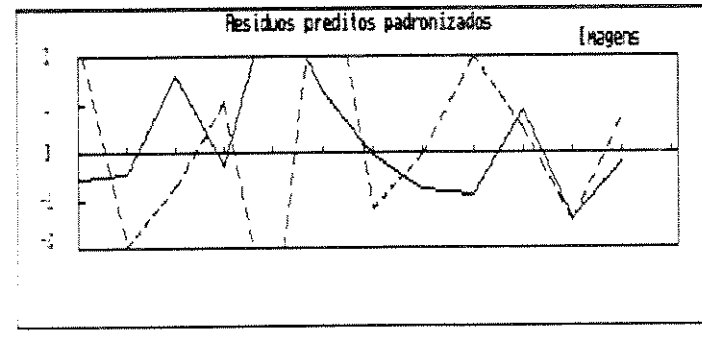
. os quatro primeiros experimentos tem comportamento idêntico, tanto nos valores finais do estado e da matriz covariância filtrada, quanto na evolução dos resíduos; a maioria dos resíduos está dentro de um intervalo de $1.\sigma$; as estimativas de precisão obtidas pela matriz covariância filtrada (desvios padrão estimados) são compatíveis com os erros médios quadráticos

obtidos; não ocorreu divergência;

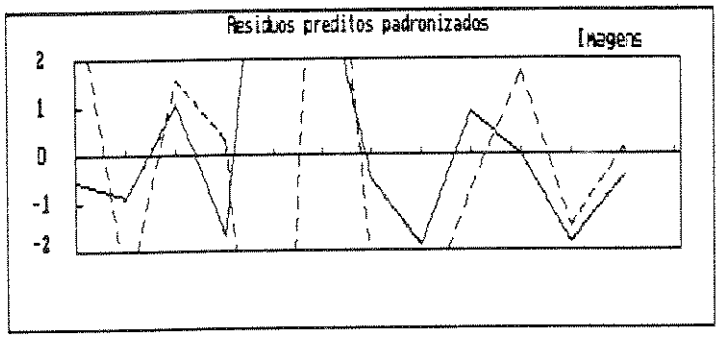
- nos experimentos 5, 6 e 7 propositadamente foram introduzidos erros no estado predito maiores do que o desvio padrão predito; no experimento 5 (erro de $2.\sigma$) verifica-se um aumento do erro médio quadrático do estado filtrado, uma fraca redução na matriz covariância e um aumento nos resíduos preditos padronizados (figura 4.2.11.(b)); nos experimentos 6 e 7 estes efeitos são claramente acentuados; nota-se no experimento 7 o erro médio quadrático aumentando sensívelmente (especialmente o erro nos parâmetros ω, Yc e Zc), a matriz covariância diminuindo e os resíduos padronizados muito maiores do que 1, o que caracteriza a divergência do filtro;
- os experimentos 5, 6 e 7 mostraram que estimativas preditas incorretas podem causar a divergência do filtro; a melhor maneira de detectar esta divergência é a análise dos resíduos preditos padronizados, uma vez que, num sistema real, não é possível calcular os erros verdadeiros;



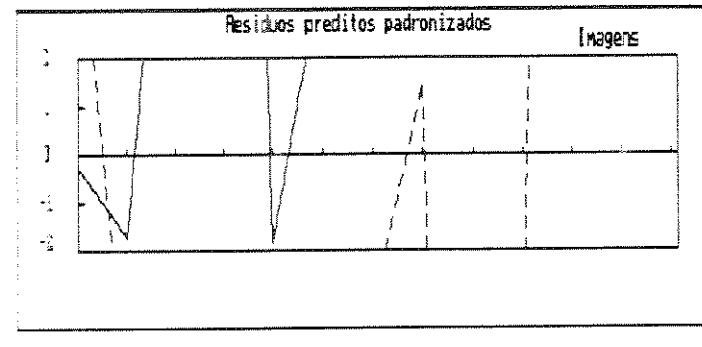
(a) Experimento 1



(b) Experimento 5



(c) Experimento 6



(d) Experimento 7

Figura 4.2.11 Resíduos preditos padronizados.

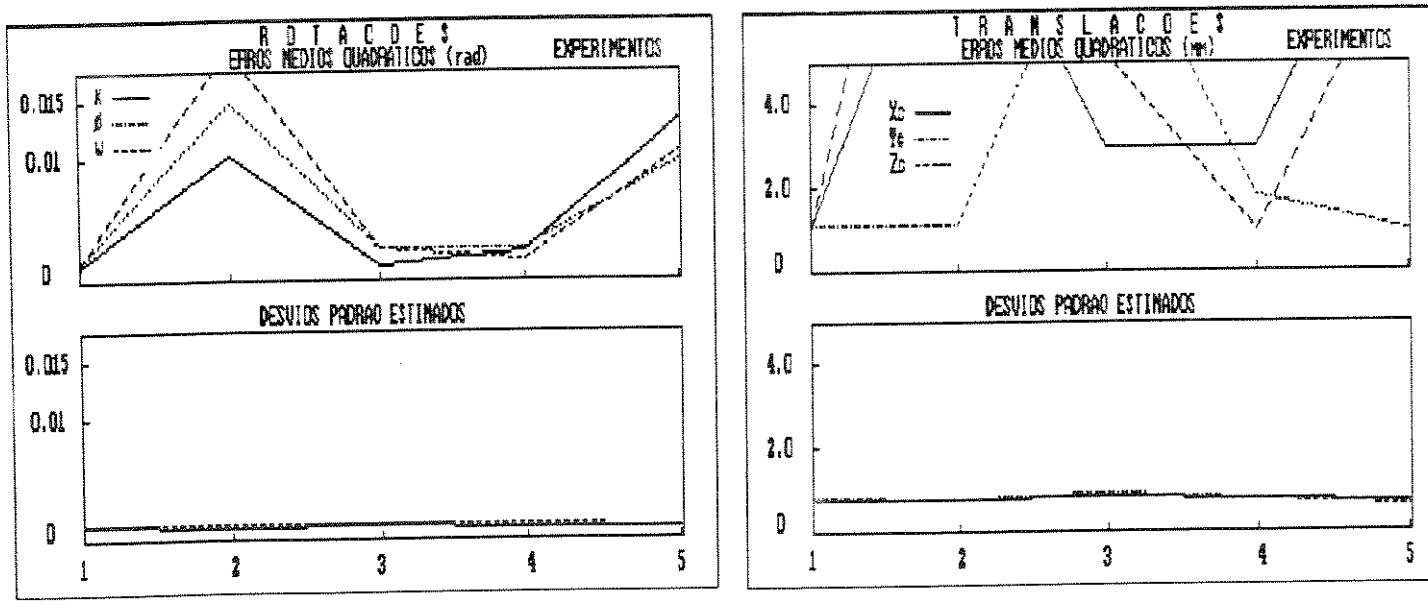
4.2.12 EFEITO DE ERROS GROSSEIROS NAS OBSERVAÇÕES.

Uma das hipóteses adotadas sobre as observações, é a de que estas estão isentas de erros grosseiros ou sistemáticos, que introduziriam uma tendência na distribuição dos erros observacionais. Neste ítem serão conduzidos experimentos com a finalidade de verificar os efeitos causados por erros grosseiros, que não podem ser preditos.

Serão apresentados 5 experimentos, com as seguintes características:

- o 1º experimento é um caso padrão, considerando a câmara (focal de 20mm) a 88cm da câmara, com as observações isentas de erros grosseiros.
- no 2º experimento foi introduzido um erro grosseiro de 500μm (50 pixels) no parâmetro **b** da 10ª reta;
- no 3º experimento foi introduzido um erro grosseiro de 60μm (6 pixels) no parâmetro **b** da 5ª reta;
- no 4º experimento foi introduzido um erro grosseiro de 60μm (6 pixels) no parâmetro **b** da 2ª reta;
- no 5º experimento foi introduzido um erro grosseiro de 10° (0.17 rad) no parâmetro **a** da 5ª reta.

Na figura 4.2.12 são mostrados os resultados para todos os experimentos, tanto para as rotações quanto para as translações.



(a) (b)

Figura 4.2.12 Efeito de erros grosseiros nas observações.

A análise da figura 4.2.12 permite estabelecer as seguintes conclusões:

- a presença de erros grosseiros de grande magnitude nas observações provoca divergência do filtro, como pode-se observar nos experimentos 2 e 5; nos experimentos 3 e 4, em particular, observou-se que o erro introduzido de 60μm no parâmetro b da 5ª e da 2ª reta provocaram um forte aumento nos erros médios quadráticos das translações, mas não alteraram substancialmente as rotações; mesmo assim, alguns resíduos padronizados ultrapassaram o valor 3, o que pode ser interpretado como divergência;
- no 5º experimento, onde é introduzido um erro no parâmetro a da 5ª reta, verifica-se um aumento substancial nos erros médios quadráticos das rotações; as translações são também afetadas, exceto o parâmetro Yc;
- é importante observar que os desvios padrão estimados são praticamente os mesmos em todos os experimentos, enquanto que os erros médios quadráticos são, comparativamente, maiores;

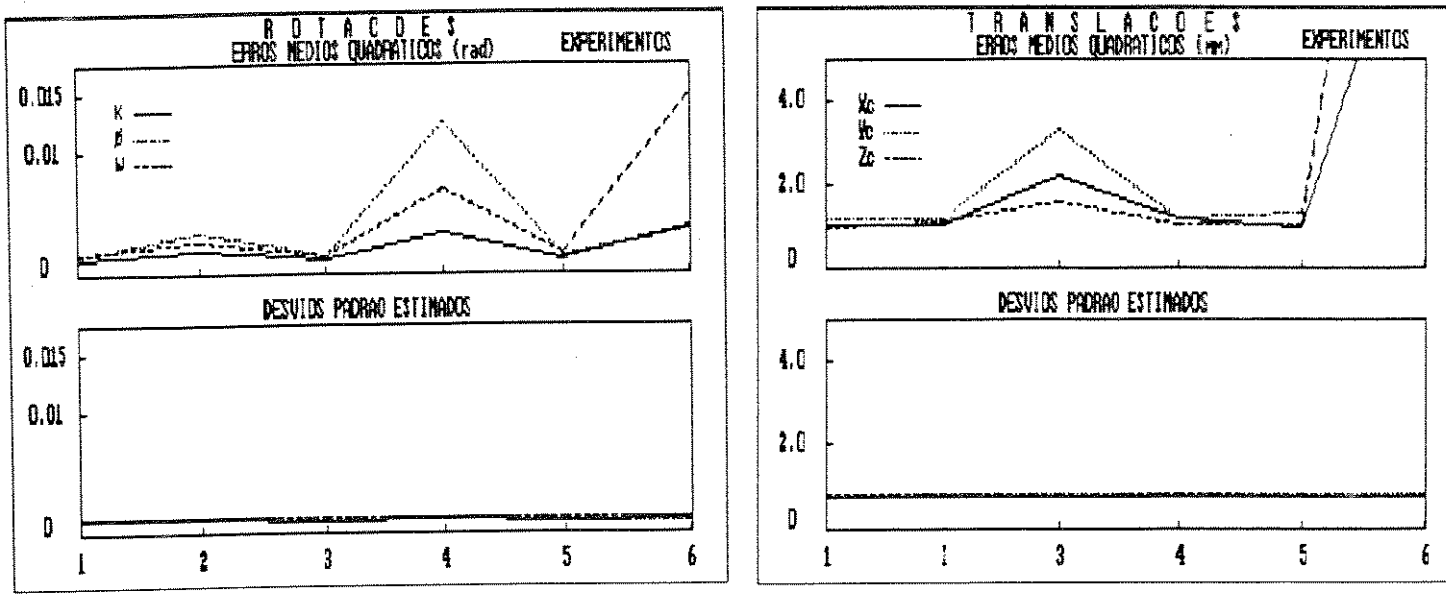
4.2.13 Efeito de erros sistemáticos não eliminados.

Ao contrário dos erros grosseiros ou dos erros randômicos, que não podem ser previstos, os erros sistemáticos tem causas físicas determinadas e podem ser parametrizados através da calibração da câmara. Contudo, a qualidade desta calibração, ou mesmo a sua ausência leva a parâmetros incorretos, o que têm influência no processo de calibração.

Nesta seção, procura-se investigar os efeitos de erros sistemáticos não corrigidos, nos resultados finais do processo de localização. Os 6 experimentos realizados tem as seguintes características:

- o 1º experimento é um caso padrão, considerando a câmara (focal de 20mm) a 88 cm da câmara, com as observações isentas de erros sistemáticos;
- no 2º experimento as imagens foram simuladas com erros no centro da imagem de 50 μm (5 pixels), em x e y;
- no 3º experimento introduziu-se um erro na distância focal de 100μm (10 pixels);
- no 4º experimento introduziram-se erros de 200μm (20 pixels) no centro da imagem, em x e y;
- no 5º experimento introduziu-se um erro de 0.0003 em k_1, o fator de distorção ótica;
- no 6º experimento introduziu-se um erro de 0.0694 no fator de escala em x (ds_x);

Para cada caso os experimentos foram repetidos 50 vezes, sem eliminar os erros sistemáticos. Os resultados destes experimentos são apresentados na figura 4.2.13.



(a) (b)

Figura 4.2.13 Efeito de erros sistemáticos não eliminados.

Analisando a figura 4.2.13 pode concluir que:

- os desvios padrão estimados são semelhantes, para todos os experimentos, enquanto que os erros médios quadráticos sofrem sensíveis variações;
- comparando os experimentos 2 e 4 (erros no centro da imagem) com o experimento 1, verifica-se as translações são menos afetadas que rotações; as translações apresentam praticamente os mesmos erros médios quadráticos em comparação com o experimento padrão (1); já as rotações são mais influenciadas pelos erros no centro da imagem, especialmente no experimento 4, onde este erro tem grande magnitude;
- a matriz covariância filtrada não é compatível com a magnitude dos erros médios quadráticos obtidos, o que poderia causar divergência do filtro no processamento dinâmico;
- comparando o experimento 3 com o experimento 1, verifica-se que o erro na focal causa uma forte tendência sistemática nas translações; as rotações sofrem pouca influência do erro na distância focal, neste caso;
- o experimento 5 mostra que o erro provocado pela não utilização da distorção ótica é muito pequeno, tanto para as rotações quanto para as translações; apenas o erro médio quadrático no parâmetro Yc sofreu um

pequeno aumento;

- o experimento 6 mostra que os efeitos do fator de escala em x são bastante severos; os erros causados pela não consideração deste parâmetro são muito maiores que nos demais experimentos;

4.2.14 RESULTADOS PARA UMA SEQUÊNCIA DE IMAGENS (CASO DINÂMICO).

Um dos pontos fundamentais do processo recursivo proposto neste trabalho, é a possibilidade de tratar o problema de localização dinâmicamente. Com o objetivo de estudar uma situação deste tipo, foram simuladas sequências de imagens, a partir da suposição de que um cubo "wireframe" estivesse se movendo sobre uma esteira, no eixo X do referencial do espaço objeto, com velocidade de 70 mm/s, e mantendo-se a câmara imóvel. Portanto, não haveria mudança nos elementos angulares do vetor estado nem nos elementos Yc e Zc, exceto por perturbações aleatórias no estado. Considerando que o objeto está a 88cm da câmara e adotando um tempo de amostragem de 1s, para cada imagem, foi obtida uma sequência de 4 imagens, mostradas na figura 4.2.14, supondo-se uma câmara com distância focal de 15mm.

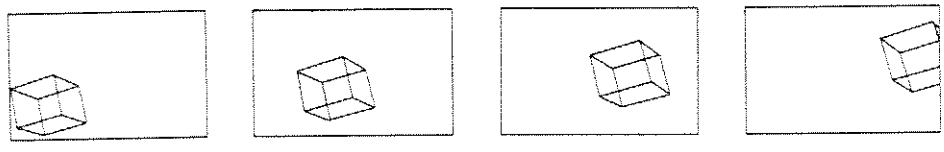
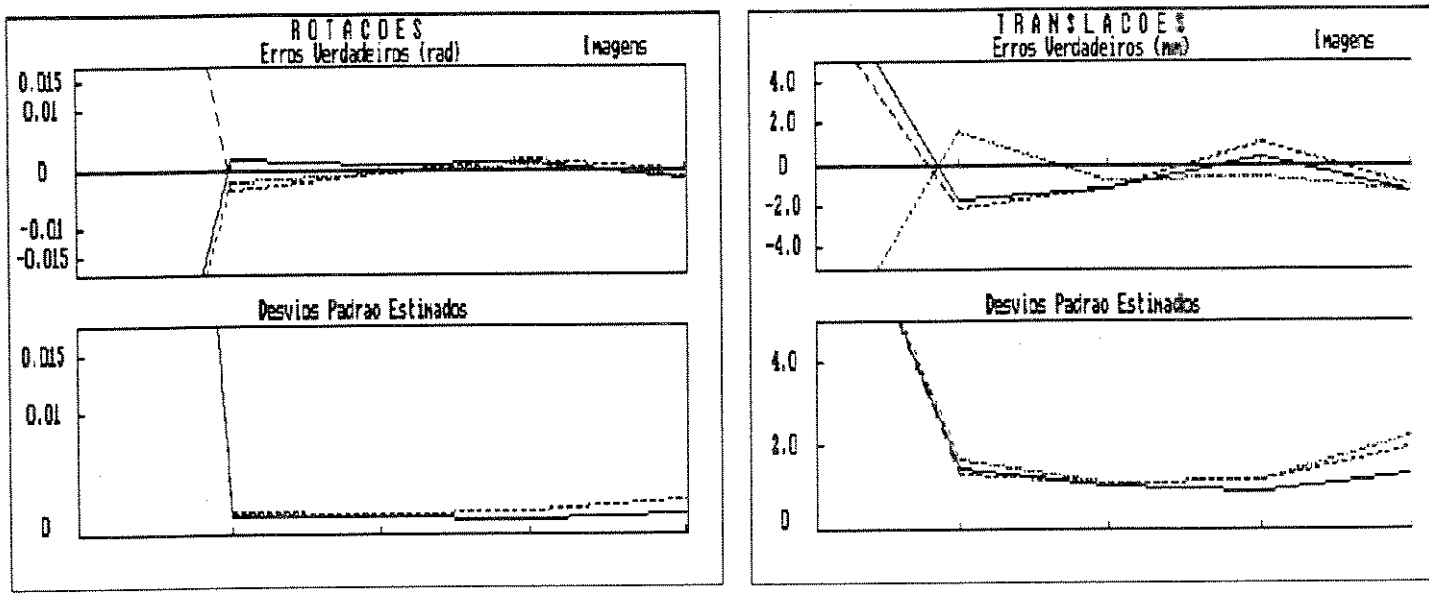


Figura 4.2.14 Sequência de imagens com o cubo a 88 cm da câmara.

Na figura 4.2.15 são mostrados os resultados obtidos pelo processo recursivo, que utiliza o estado filtrado no instante t_{k-1} para compor a estimativa predita do estado no instante t_k, através do modelo do sistema, e da propagação de covariâncias, para calcular a matriz covariância predita no instante t_k. No exemplo proposto, o modelo do sistema constitui-se, simplesmente, de uma translação em X, enquanto que no cálculo da matriz covariância predita para o instante t_k deve-se considerar a perturbação no vetor estado entre os instantes de amostragem. Neste primeiro experimento supõe-se que erro de transição de estado é de 5° (0.087 rad) para as rotações

e de 5mm para as translações.

Na figura 4.2.15 o elemento 0 (zero) indica o erro verdadeiro na estimativa predita para a primeira imagem; em seguida, são apresentados os erros verdadeiros nas estimativas filtradas para cada instante (imagem). Ao contrário dos experimentos anteriores, em que foram apresentados os erros médios quadráticos de 50 repetições, neste caso serão mostrados os erros verdadeiros em uma única sequência, para poder avaliar a possibilidade de tendência nas estimativas após a filtragem de toda a sequência.



(a) (b)

Figura 4.2.15 Resultados para uma sequência de 4 imagens.

Da análise dos gráficos apresentados na figura 4.2.15 é possível inferir que:

- a precisão alcançada com as imagens em que o cubo aparece nos cantos (1 e 4) é ligeiramente inferior às duas demais imagens;
- a sequência convergiu, embora, como já foi ressaltado, com uma ligeira perda de precisão na última amostra;

Uma segunda série de experimentos foi realizada, agora com o cubo a 1.10m, gerando uma sequência de 6 imagens, mostradas na figura 4.2.16. Os resultados obtidos com esta sequência, partindo de condições análogas ao experimento anterior, são mostrados nas figuras 4.2.17 e 4.2.18. Os resultados apresentados na figura 4.2.17 foram obtidos considerando-se os mesmos erros de transição de estado do caso anterior. Para os resultados

mostrados na figura 4.2.18 foram adotadas estimativas preditas um pouco melhores e um erro de transição de estado menor (0.5° (0.0087rad) para as rotações e 1mm para as translações).

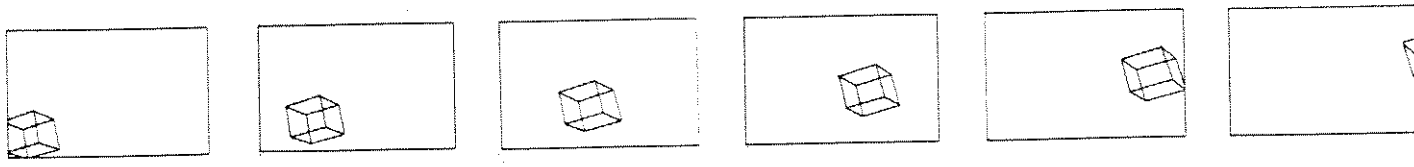
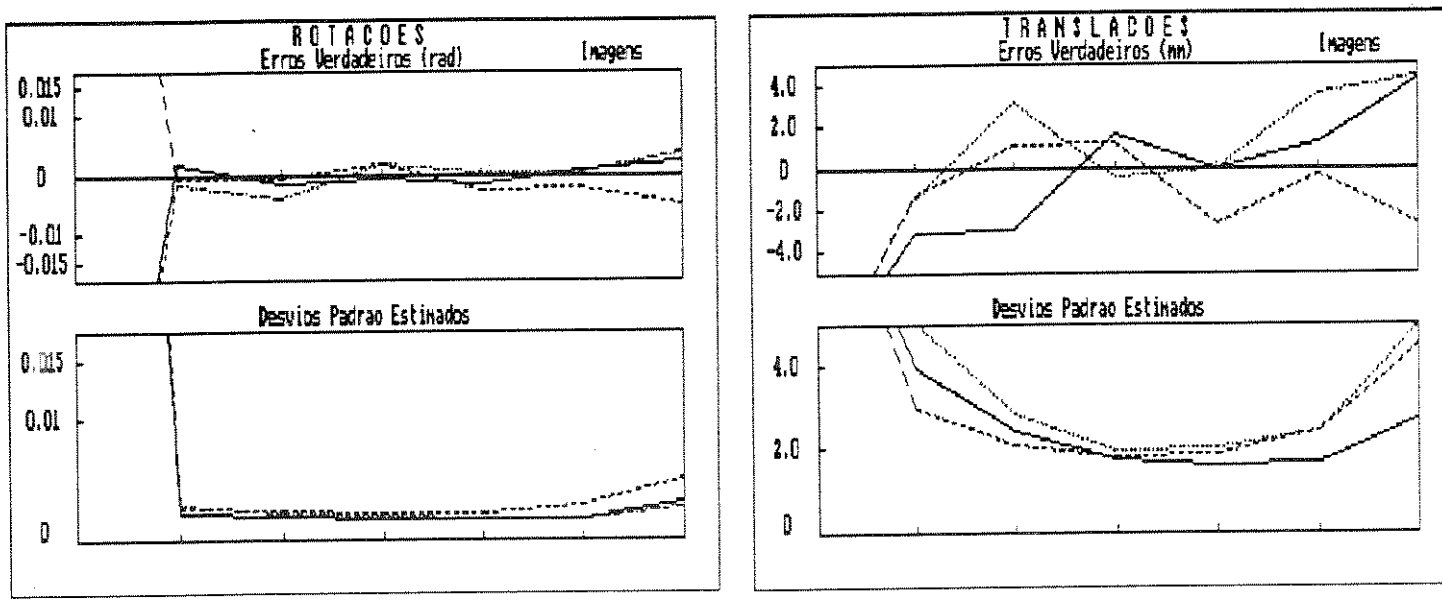


Figura 4.2.16 Sequência de imagens com o cubo a 1.10 m da câmara.

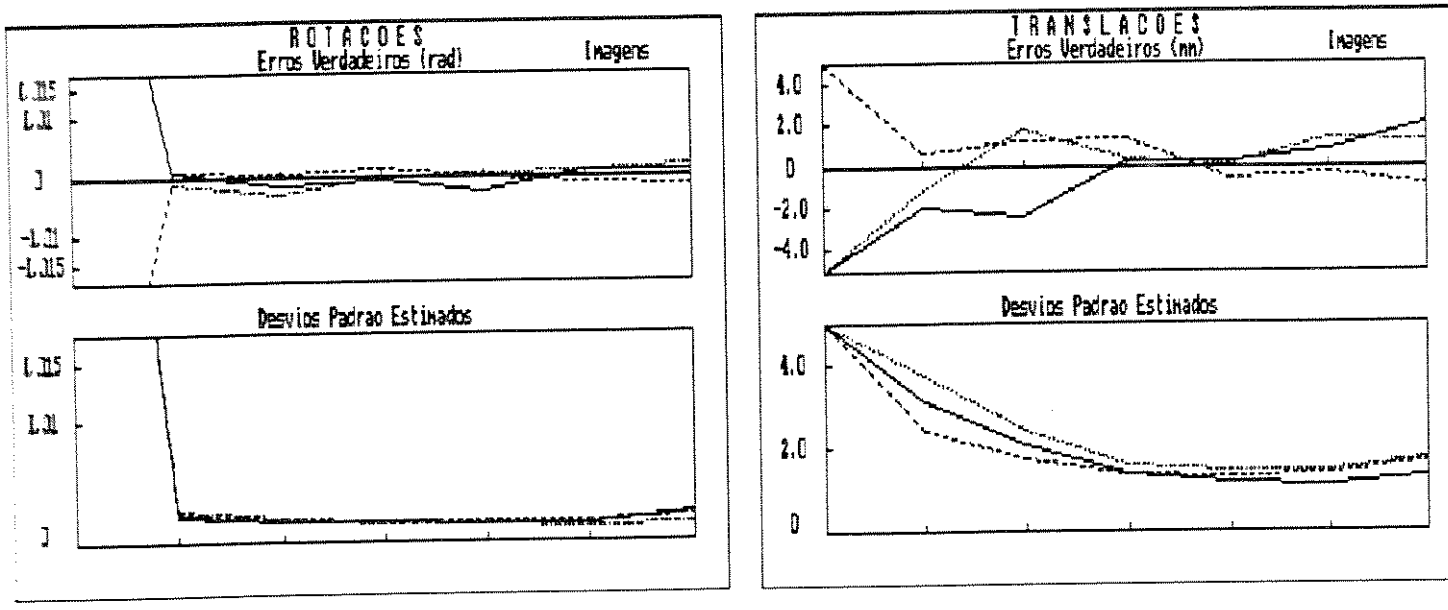


(a) (b)

Figura 4.2.17 Resultados para uma sequência de 6 imagens.

Analisando a figura 4.2.17 e comparando estes resultados com os apresentados na figura 4.2.15 é possível verificar que:

- da mesma maneira que no caso anterior existe, uma perda de precisão nos últimos frames da sequência, tanto nas rotações quanto nas translações;
- com relação ao caso anterior, existe uma perda de precisão ao aumentar a distância câmara-objeto;



(a) (b)

Figura 4.2.18 Resultados para uma sequência de 6 imagens, com erro de transição de estado menor.

A figura 4.2.18 mostra resultados em que o erro de transição de estado foi reduzido, em relação aos dois experimentos anteriores. Pode-se verificar claramente que existe uma tendência de estabilização na estimativa, mesmo nas últimas imagens da sequência. Isto ocorre porque a matriz covariância predita para as imagens posteriores à primeira, é mais precisa, já que o erro de transição de estado é menor;

Contudo, caso a perturbação real no estado fosse maior que erro de transição de estado previsto, poderia haver divergência do filtro. Nas situações dinâmicas este erro de transição de estado é mais uma componente importante do sistema. Se este erro for subestimado perde-se uma informação importante que poderia conferir precisão às imagens finais de uma sequência; contudo, se este erro for super-estimado pode provocar a divergência no filtro.

4.2.15 SUMÁRIO DAS CONCLUSÕES OBTIDAS COM O ESTUDO COM DADOS SIMULADOS.

Do estudo desenvolvido com dados simulados podem ser extraídas algumas conclusões básicas:

- com relação à convergência dos elementos do vetor estado, à medida que as observações são processadas, verifica-se que o parâmetro κ converge mais rapidamente que os demais;
- para duas imagens com mesma escala aquela que for tomada com uma câmara de distância focal menor apresentará, provavelmente, melhores resultados no processo de localização;
- quanto maior a distância câmara-objeto menor a precisão;
- podem existir configurações deficientes mesmo quando o número de retas é maior que 3, como, por exemplo, um conjunto de retas paralelas;
- nas imagens em que as retas observadas se concentram em partes da imagem, notadamente os cantos, existe uma perda de precisão; além disto, a matriz covariância frequentemente subestima os erros no estado;
- quanto maior o erro observacional menor a precisão final da localização;
- a super-estimação da precisão das observações pode causar a divergência do filtro;
- estimativas preditas com erro verdadeiro maior que $3x\sigma$ (σ é o desvio padrão predito) podem provocar divergência do filtro;
- erros grosseiros de grande magnitude nas observações causam divergência do filtro, na maioria dos casos;
- erros sistemáticos não modelados de grande magnitude provocam uma tendência na estimativa do estado, podendo, ainda, causar divergência do filtro, especialmente nos problemas dinâmicos com sequências de imagens;

4.3 RESULTADOS COM DADOS REAIS.

4.3.1 CARACTERÍSTICAS DOS EXPERIMENTOS.

Com o objetivo de testar o procedimento desenvolvido, foram conduzidos alguns experimentos reais em um ambiente parcialmente controlado. Optou-se por desenvolver experimentos independentes do manipulador, colocando a câmara em uma posição arbitrária e efetuando a localização-calibração. Foi utilizada uma câmara Kentek-CCD, com distância focal nominal de 15mm; as imagens foram digitalizadas através de uma placa AT-Vista, com 512 colunas e 400 linhas.

Com o objetivo de obter os parâmetros intrínsecos da câmara, foram tomadas 6 imagens convergentes de um conjunto de 20 alvos circulares. Estes alvos foram colocados sobre um mesmo plano e espaçados de 10cm. As coordenadas de máquina das imagens destes alvos foram extraidas calculando-se o centro de massa dos pixels acima do limiar pré-estabelecido. Aplicou-se, então, o procedimento clássico de calibração, descrito no Capítulo 2, seção 2.1.1, obtendo-se os resultados apresentados na tabela 4.3.1. O centro da imagem refere-se ao centro aproximado da imagem (256, 200).

Tabela 4.3.1 Parâmetros intrínsecos da câmara CCD e seus respectivos desvios padrão estimados, obtidos com o método clássico, utilizando 6 imagens convergentes.

Parâmetro	cx	cy	k1	ds	focal
Valor (mm)	-0.1478	0.0487	0.000323	0.06847	16.3896
σ_x (mm)	0.0165	0.0191	0.000031	0.00021	0.0633

Observa-se, na tabela 4.3.1, o desvio padrão estimado da ordem de 16μm (1,6 pixels) para c_x e de 19μm para c_y, o que é perfeitamente aceitável para a aplicação proposta. Verifica-se, ainda, a alta precisão na determinação do fator de escala em x (ds).

Uma vez calibrados os parâmetros intrínsecos da câmara, o dispositivo de focalização permaneceu imóvel. Caso este dispositivo seja movimentado para focalizar alguma cena, então uma nova calibração deve ser realizada.

Foram adotados dois referenciais para o modelo do sistema: o referencial da câmara e o referencial do objeto. Estabeleceu-se o referencial do objeto coincidente com o referencial global. O objeto de teste, no caso um cubo de

70mm de lado, foi, então, medido em relação ao vértice de origem (figura 4.3.1).

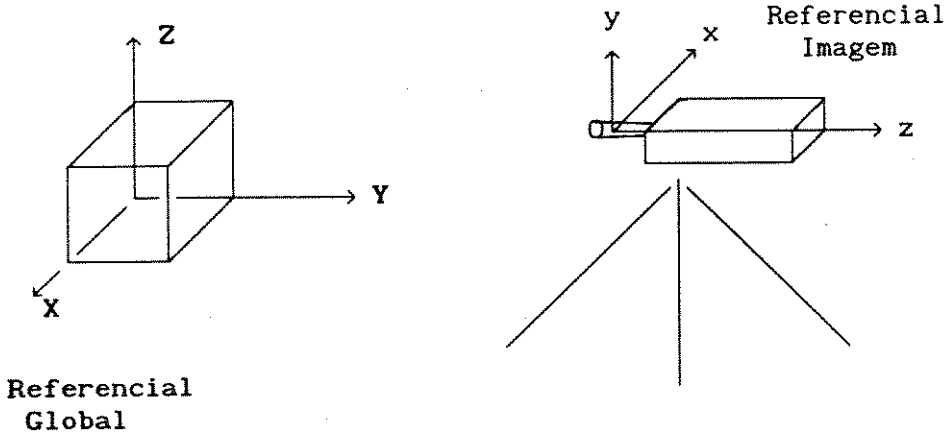


Figura 4.3.1 Referenciais usados nos experimentos reais.

As imagens de bordas do objeto imageado foram geradas em um processo "off-line", calculando-se os gradientes usando os operadores de Sobel; a magnitude e a direção do gradiente foram calculados usando as expressões (2.5.1) e (2.5.3) (Capítulo 2, seção 2.5).

O processo recursivo foi implementado para atuar sobre a imagem de bordas previamente gerada, considerando-se como elemento de borda aquele pixel que possuir a magnitude do gradiente acima do limiar pré-estabelecido. Devido às limitações de memória, e para acelerar o processo, a imagem de bordas é armazenada descartando-se as linhas ímpares, o que resulta em uma imagem 512x200. Esta redução do número de linhas terá reflexos na precisão das observações, uma vez que as retas serão definidas por um número menor de elementos de borda.

As estimativas preditas foram obtidas através de medidas diretas (posição e orientação da câmara em relação ao cubo) e confirmadas através de valores calculados com base em uma cópia impressa da imagem de bordas. Os parâmetros da câmara foram obtidos medindo-se as coordenadas dos vértices do cubo na cópia impressa, com uma escala comum, e realizando-se, então, uma localização (resseção espacial) usando estes pontos.

Devido ao erro de 0.1mm na medida dos vértices na cópia impressa, os valores obtidos para os parâmetros serão, evidentemente, menos precisos do que

aqueles obtidos com a imagem original. Assumindo-se, neste caso, um erro de quantização de 1 pixel, este processo equivaleria à utilização de uma imagem de 50 colunas por 20 linhas. Esta técnica de obtenção de valores preditos é desnecessária num sistema real, uma vez que a posição da câmara na garra do manipulador é conhecida com razoável precisão. No caso da localização de objetos o processo de reconhecimento fornece uma estimativa predita de qualidade compatível com o processo recursivo proposto.

A resolução da imagem é uma das componentes ligada à qualidade das observações, porque o erro de quantização está relacionado com o tamanho do pixel (cerca de 1/4 do pixel). Portanto, para uma mesma área de imageamento, espera-se uma precisão maior das observações para um número maior de amostras (maior resolução). Nos experimentos com dados simulados verificou-se experimentalmente o efeito dos erros observacionais nos valores finais dos parâmetros de localização.

Do ponto de vista da Filtragem Kalman, a estimativa predita não poderia ser muito grosseira. Experimentos com as simulações mostraram que erros da ordem de 5° nas rotações e 50mm nas translações, nas estimativas preditas, permitem uma convergência satisfatória do filtro. Contudo, erros muito grandes na estimativa predita poderiam aumentar tanto a janela de busca inicial que esta abrangeria toda a imagem. Neste caso, a idéia do processo recursivo, em que os resultados da filtragem realimentam a extração de feições, perderia o sentido prático.

Por este motivo, nos experimentos reais procurou-se trabalhar com estimativas preditas com erros angulares da ordem de 3° a 5° nas rotações e 20mm (ou menos) nas translações.

4.3.2. Resultados com o cubo estático.

Neste primeiro experimento procura-se avaliar a precisão obtida em uma configuração típica, coletando-se uma imagem com o cubo imóvel, e à distância aproximada de 1,13 m.

Na figura 4.3.2 mostra-se a imagem de bordas do cubo (pixels com magnitude do gradiente acima do limiar estabelecido) para este experimento. Esta imagem é utilizada no processo recursivo para calcular os parâmetros da câmara.

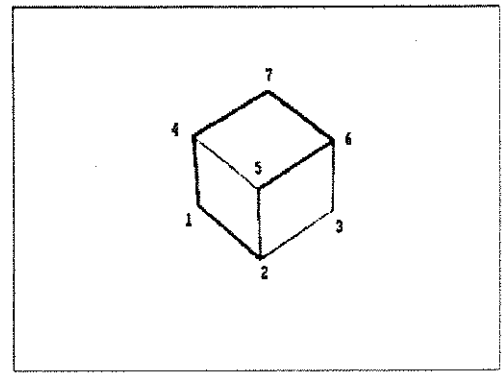


Figura 4.3.2 Imagem de bordas do cubo a 1.13 m da câmara.

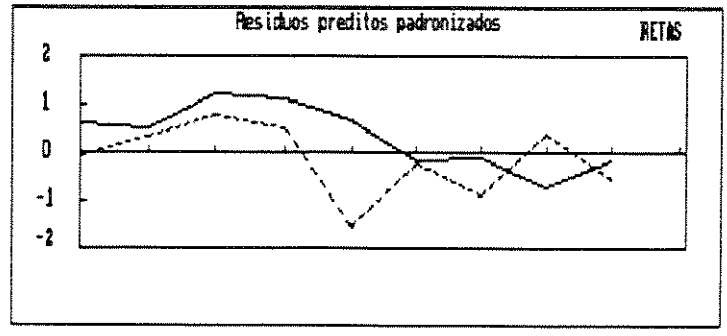
Com a imagem mostrada na figura 4.3.2 foram obtidos excelentes resultados no processo de localização. A tabela 4.3.4 mostra os valores das estimativas predita e filtrada usando o processo recursivo (IEKF), com esta imagem. A matriz covariância do estado filtrado indica erros da ordem de 5' para as rotações e de 1.5 mm para as translações.

Tabela 4.3.2 Resultados com o IEKF, utilizando a imagem 4.3.2; cubo a 1.13m da câmara.

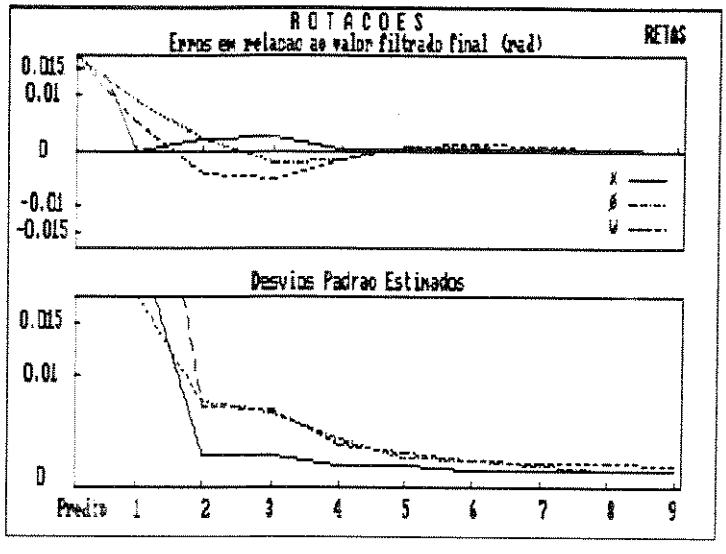
	VALORES PREDITOS		VALORES FILTRADOS	
	Estado Predito	Desvio Padrão	Estado Filtrado	Desvio Padrão
κ	2.6070	0.0523	2.57783	0.00132
ϕ rad	0.4870	0.0523	0.46920	0.00125
ω	-0.6280	0.0523	-0.64528	0.00190
X_c	528.00	10.00	523.37	1.343
Y_c mm	628.00	10.00	621.44	1.561
Z_c	785.00	10.00	789.00	1.623

Na figura 4.3.3 mostra-se a evolução das estimativas com a introdução das retas. Nestes gráficos o erro indicado corresponde à diferença entre o valor filtrado com i-ésima reta e o valor filtrado final. Nota-se, claramente que já na quarta reta as rotações atingem um estado semelhante ao valor final, o que é compatível com as conclusões obtidas no estudo com dados simulados. Os resíduos padronizados, são, em sua maioria, menores que 1, o que indica

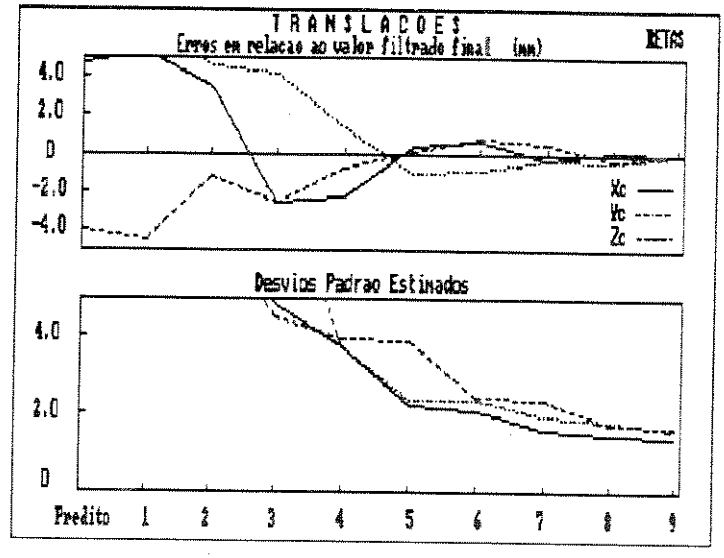
convergência da solução e estimativa adequada das variâncias das observações.



(a)



(b)



(c)

Figura 4.3.3 Resultados para a imagem 4.3.2.

A tabela 4.3.3 mostra os tempos de processamento, obtidos em um AT 386, para as etapas do processo recursivo, além da área da janela de busca. É possível verificar, nesta tabela, a progressiva redução no tempo de processamento, em decorrência da redução da janela de busca. A etapa de determinação das faces visíveis do objeto e escolha de uma reta (hidden lines) tem um tempo de processamento praticamente constante. A etapa de estimação usando Filtragem Kalman tem o tempo variável em função do número de iterações.

É importante ressaltar a grande redução no tempo de processamento na etapa de vetorização, resultado da redução na janela de busca com a

realimentação proveniente do processo de estimação. Esta redução ocorre porque a cada nova reta introduzida a estimativa de localização da câmara é melhorada e, consequentemente, a janela de busca para a próxima feição é também reduzida.

Tabela 4.3.3 Tempos de processamento, em segundos, para cada etapa do processo recursivo e área da janela de busca, em nº de pixels, para cada reta.

Etapa \ Reta	67	74	45	56	52	23	36	41	12	
Hidden Lines	0.05	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.06	
Scan Lines	0.11	0.11	0.06	0.05	0.05	0.05	0.06	0.05	0.05	
Vetorização	6.26	3.46	1.42	1.38	0.88	0.61	0.55	0.77	0.82	
Filtragem	0.22	0.17	0.22	0.17	0.22	0.16	0.22	0.17	0.16	
Tempo Total	6.64	3.80	1.76	1.65	1.20	0.87	0.88	1.04	1.09	18.93s
Área da Janela	28520	18176	3545	2946	1326	1002	713	581	552	

A análise da tabela 4.3.3 permite extrair algumas conclusões:

- a utilização do processo recursivo promove uma substancial redução no tempo de processamento;
- em alguns casos, embora a janela de busca seja menor, o tempo de processamento é maior (retas 23, 26, 41 e 12). Isto ocorre devido à dispersão dos elementos de borda, o que obriga a várias iterações no processo de ajuste da reta e eliminação de pontos espúrios. Outra causa do aumento do tempo de processamento é o número de elementos de bordas presentes dentro de cada janela; embora a área da janela seja maior é possível existir um número menor de elementos de borda;
- a variação nos tempos da filtragem está relacionada ao número de iterações, que depende da convergência do iterador;

Na figura 4.3.4 mostra-se, graficamente, o efeito de redução da janela de busca, para este caso. Nota-se a sensível redução nas dimensões da janela, a partir da 3ª reta. Na última reta, a largura da janela é de apenas alguns pixels.

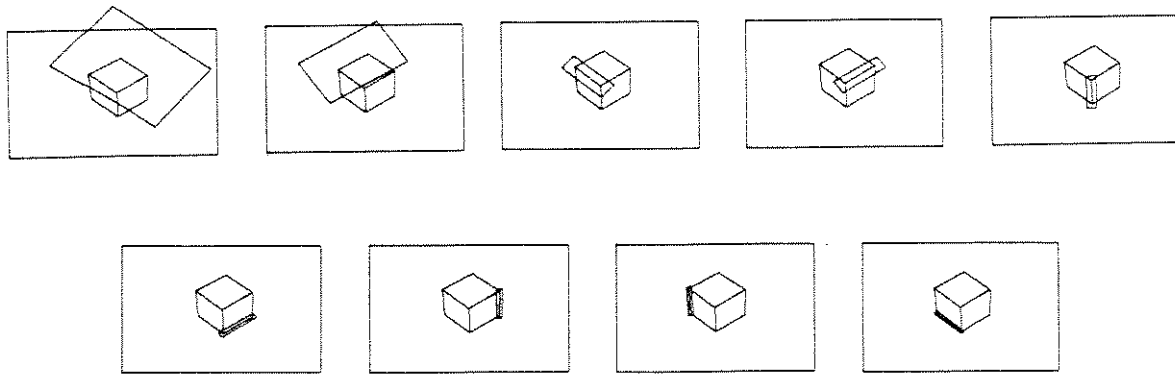


Figura 4.3.4 Redução progressiva da janela de busca.

A adoção do processo recursivo, que permite uma redução gradativa do espaço de busca na etapa de extração de feições, possui duas vantagens básicas, quando comparado aos métodos de extração de feições por toda a imagem:

. o tempo de processamento é menor;
. a possibilidade de seleção de retas espúrias é diminuida, pois a redução do espaço de busca reduz o número retas semelhantes;

Em relação ao tempo de processamento, uma análise simples pode ser feita utilizando os dados da tabela 4.3.3.. O número total de pixels analisados é de 57.361, enquanto que o total de pixels na imagem é de 102.400. Isto significa uma redução de 44% no numero de pixels a serem analisados, embora isto não implique na mesma redução do tempo de processamento, pois nem todos os pixels pertencem a uma borda.

As relações acima podem ser ainda melhoradas se forem consideradas as interseções entre as janelas. O processo proposto neste trabalho pode ser otimizado para processar somente as partes da janela de busca ainda não analisadas. Com relação à Filtragem Kalman, podem ser usados métodos de fatoração da matriz covariância, que aceleram o processamento.

Outra comparação pode ser feita com relação aos resultados obtidos por Lee et al (1990), embora estes autores tenham empregado uma workstation SUN com uma unidade de aceleração aritmética. O algoritmo utilizado pelos autores baseia-se na análise de toda a imagem para a detecção de bordas e extração de formas retangulares. Os tempos de processamento para a extração de feições retangulares são de 4, 3, 8, 14 e 3 segundos para a detecção de bordas, suavização ("smoothing"), afinamento, transformada de Hough e cálculo dos

vértices dos retângulos, respectivamente. O cálculo dos parâmetros de localização leva apenas 1.07ms. Cerca de 32 segundos são necessários para processar a imagem (640x400).

Embora Lee et al tenham utilizado alguns passos adicionais, em relação ao processo recursivo proposto neste trabalho, pode-se verificar que, apenas a transformada de Hough consumiu 14 segundos de processamento. Deve-se considerar, nesta comparação, a eficiência da máquina empregada por Lee et al em comparação com um AT 386. Uma estação SUN é cerca de 16 vezes mais rápida que um AT 386 com coprocessador 80287, utilizado nestes experimentos.

4.3.3 Resultados com o cubo em movimento.

O experimento anterior foi desenvolvido com o cubo estático. Em aplicações industriais, contudo, o objeto pode estar em movimento, durante a aquisição da imagem, o que causa um efeito de arrastamento na imagem, dependendo da velocidade do objeto. O objetivo do experimento a ser apresentado nesta seção é investigar os efeitos causados por este arrastamento, nos parâmetros da câmara.

Para produzir este efeito de arrastamento introduziu-se um movimento linear uniforme de cerca de 1 m/s, movendo-se o cubo na direção X e coletando-se a imagem durante o movimento. O resultado deste movimento é um certo borramento na imagem com a perda de nitidez das bordas. É importante ressaltar que esta velocidade introduzida é superior às condições típicas de um ambiente industrial.

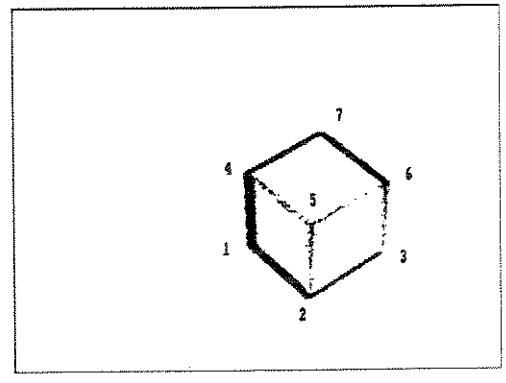


Figura 4.3.5 Imagem de bordas com o efeito de arrastamento causado pelo movimento do objeto.

Na figura 4.3.5 é mostrada a imagem de bordas obtida do cubo em movimento. Pode-se notar a espessura das bordas no contorno do objeto, e verificar o efeito de borramento, e compará-las com as obtidas com o objeto estático (figura 4.3.2).

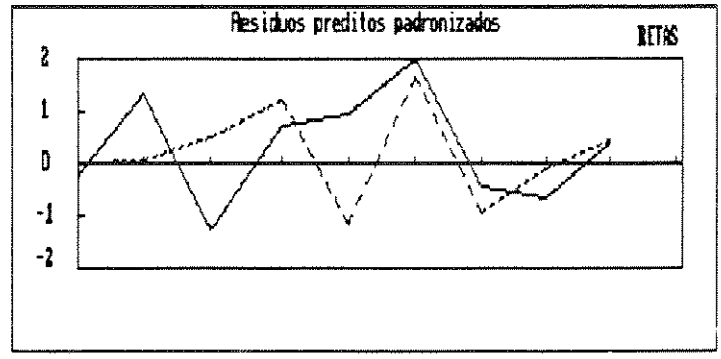
A imagem mostrada na figura 4.3.5, tomada com o cubo em movimento, produziu os resultados dispostos na tabela 4.3.4. Verifica-se, claramente, que os erros finais são maiores que os obtidos com o cubo estático. Isto é causado pela perda de nitidez das bordas, o que causa um aumento na variância das observações. Estimou-se um erro máximo de 15' para as rotações, e de 4.2mm para as translações. Não foi observada divergência do filtro, neste caso.

Tabela 4.3.4 Resultados com o cubo em movimento.

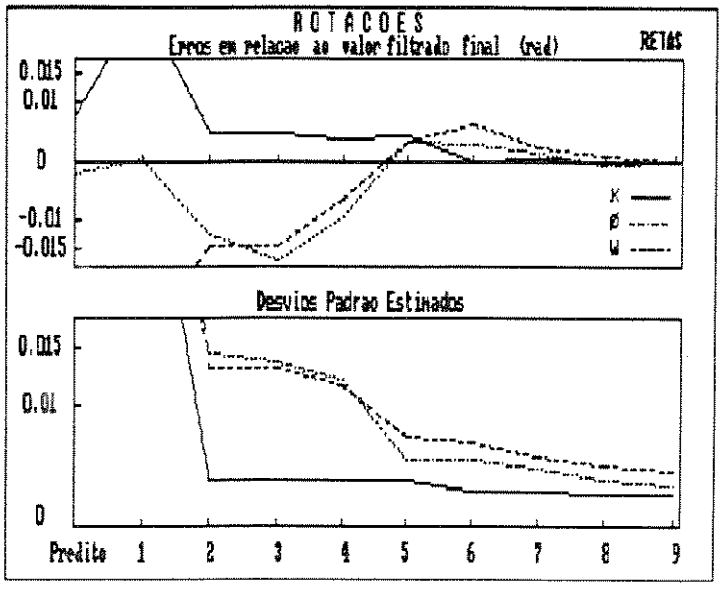
	VALORES PREDITOS		VALORES FILTRADOS	
	Estado Predito	Desvio Padrão	Estado Filtrado	Desvio Padrão
κ	2.6449	0.0860	2.63718	0.00259
ϕ rad	0.4725	0.0860	0.47470	0.00319
ω	-0.7399	0.0860	-0.69825	0.00456
X_c	568.00	20.00	552.201	3.418
Y_c mm	669.00	20.00	627.137	3.206
Z_c	783.00	20.00	785.410	4.248

A convergência das estimativas em relação ao valor filtrado final, para o experimento com o cubo em movimento, é apresentada na figura 4.3.6.

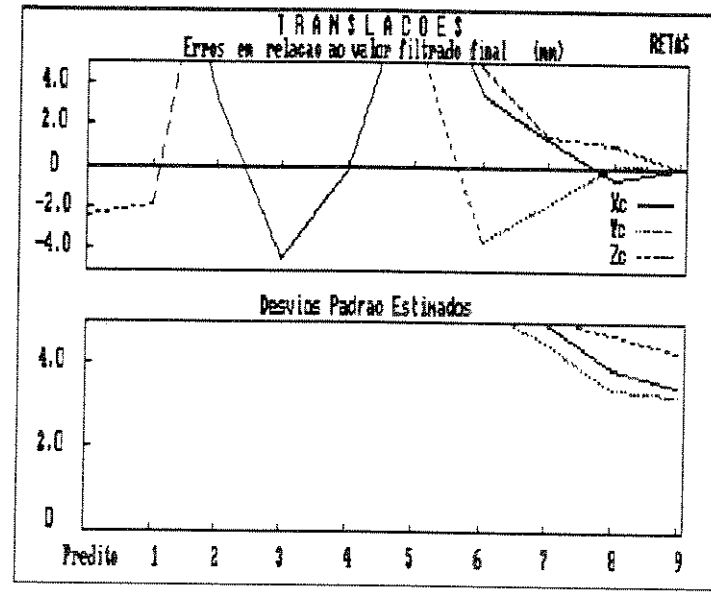
Neste caso, é possível verificar que as rotações convergiram para um valor próximo ao valor filtrado final (fig. 4.3.6.(b)), já com a introdução da 5ª reta. Com as translações (fig. 4.3.6.(c)), entretanto, a convergência foi mais lenta, atingindo um valor próximo ao valor final, somente com a 7ª reta. Mesmo neste caso, em que a qualidade da imagem foi deteriorada, o sistema foi capaz de encontrar uma solução confiável. A análise dos resíduos preditos padronizados (figura 4.3.6.(a)) mostra que não houve divergência do filtro.



(a)



(b)



(c)

Figura 4.3.6 Convergência da estimativa para o cubo em movimento.

Da análise deste experimento, conclui-se que o processo recursivo proposto consegue tratar mesmo imagens de qualidade inferior, apenas resultando em parâmetros de localização menos precisos.

4.3.4 AVALIAÇÃO DA REPETIBILIDADE DOS EXPERIMENTOS COM UMA ÚNICA IMAGEM.

O objetivo desta seção é apresentar os resultados da avaliação da repetibilidade de um mesmo experimento. Para avaliar esta repetibilidade foram coletadas três imagens, com um intervalo de 1 minuto, sem mover o objeto e a câmara, mas com uma pequena alteração na posição da fonte de iluminação.

Inicalmente, serão apresentados os resultados de um dos experimentos

isoladamente, de modo a permitir uma avaliação de algumas características particulares, como o valor numérico das observações e suas variâncias. Posteriormente serão apresentados gráficos com os desvios dos parâmetros da câmara para cada experimento, em relação à média, e os desvios padrão estimados.

Neste experimento, o cubo foi posicionado a 1,13m da câmara. A imagem de bordas resultante (pixels com magnitude do gradiente acima de um limiar) é mostrada na figura 4.3.7. As outras duas imagens são semelhantes.

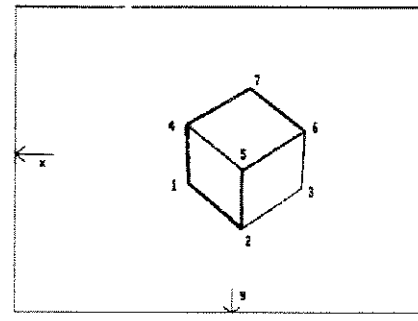


Figura 4.3.7 Imagem de bordas do cubo a 1,13m da câmara.

Tabela 4.3.5 Valores medidos na imagem (parâmetros das retas);
. σ_0^2 é o desvio padrão a posteriori no ajuste das retas;
. ptos indica o número de pontos usados no ajuste;
. σ_a^2 e σ_b^2 são as variâncias das observações;

Reta	a (a*)	b (b*)	$\hat{\sigma}_0$	ptos	σ_a^2	σ_b^2
67 g	-0.76387	-1.1117	0.0024	28	2.50e-05	2.94e-05
74 g	0.58592	-0.9105	0.0019	24	2.50e-05	2.64e-05
45 g	-0.82548	0.0353	0.0045	29	2.50e-05	3.91e-05
56 g	0.63740	0.1259	0.0051	12	4.65e-05	5.36e-05
52 f	0.00963	-0.0506	0.0030	38	2.50e-05	2.99e-05
23 g	0.65911	0.8694	0.0029	26	2.50e-05	2.94e-05
36 f	0.05439	-0.8601	0.0031	37	2.50e-05	2.90e-05
41 f	-0.02817	0.6629	0.0032	37	2.50e-05	2.90e-05
12 g	-0.86678	0.8159	0.0027	30	2.50e-05	3.07e-05

Na tabela 4.3.5 são apresentados os valores obtidos para os parâmetros das retas, suas respectivas variâncias, o número de pontos usados no ajuste da reta e o desvio padrão a posteriori. A variância a posteriori é um indicativo

da dispersão dos pontos componentes da reta em relação ao valor ajustado. As variâncias das observações foram obtidas calculando-se a matriz covariância dos parâmetros da reta e introduzindo, adicionalmente, os efeitos dos erros nos parâmetros intrínsecos, através de propagação de covariâncias.

Entretanto, observou-se que, mesmo com a introdução dos erros nos parâmetros intrínsecos, algumas variâncias permaneciam muito pequenas, o que introduziria um peso maior para a observação correspondente, no processo de Filtragem. Alguns autores (Jazwinski, 1970) recomendam aumentar a variância da observação, nestes casos, para prevenir a divergência do filtro.

Como exemplo, as variâncias obtidas no ajuste da reta 67 são:

σ_a^2 = 4.89e-06 e σ_b^2 = 1.51e-06;

Isto indicaria um erro de apenas 1.2μm em b. Como estes valores obtidos no ajuste da reta pelo MMQ são indicativos de dispersão, é necessário considerar, ainda, o efeito dos erros sistemáticos não modelados. Mesmo com a consideração dos erros nos parâmetros intrínsecos, σ_a^2 permaneceria pequena, pois o parâmetro a não sofre a influência de Cx e Cy.

Experimentalmente, chegou-se à conclusão que um valor mínimo de 2.5e-05 faria com que os resíduos preditos tivessem o comportamento esperado, evitando a divergência do filtro. Este valor de 2.5e-05 corresponde a $(1/200)^2$, ou seja, espera-se um erro de 1 pixel em 200.

Na tabela 4.3.6 são apresentados os resultados obtidos para este mesmo experimento. Pode-se notar os desvios padrão estimados para as rotações da ordem de 7' (0.002 rad) e para as translações da ordem de 2,5 mm.

Tabela 4.3.6 Valores preditos e filtrados, obtidos com o IEKF, para um cubo a 1,12m da câmara.

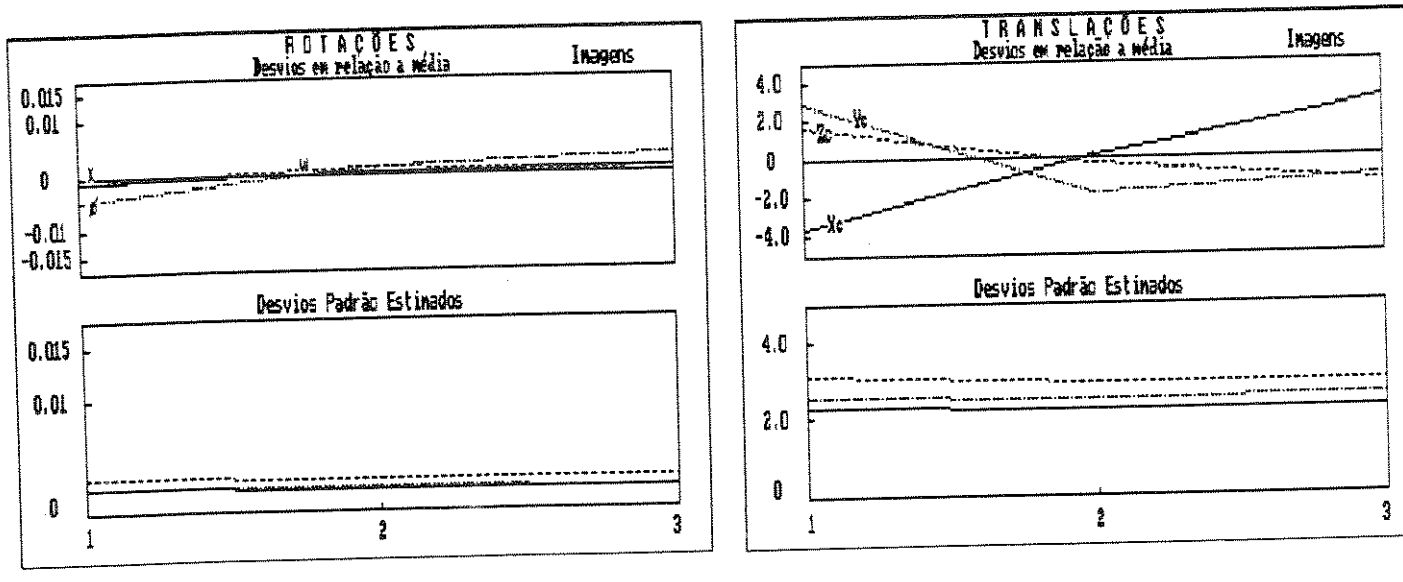
	VALORES PREDITOS		VALORES FILTRADOS	
	Estado Predito	Desvio Padrão	Estado Filtrado	Desvio Padrão
κ	2.5810	0.0871	2.584700	0.0020
ϕ rad	0.4750	0.0871	0.481402	0.0020
ω	-0.6330	0.0871	-0.632504	0.0031
Xc	533.00	20.00	537.61	2.23
Yc mm	599.00	20.00	594.33	2.54
Zc	799.00	20.00	795.14	3.05

Comparando-se estes resultados com os apresentados na seção 4.3.2 (Tabela 4.3.2) verifica-se que, embora a distância seja a mesma, os experimentos apresentados na seção 4.3.2 apresentaram melhores resultados. Isto se deve, principalmente, à qualidade da imagem naquele caso, que foi obtida com o ponto de iluminação em uma posição mais adequada, permitindo um melhor contraste entre as faces do objeto. Isto produziu bordas mais nítidas que as produzidas no experimento desta seção.

Isto mostra que a melhoria da qualidade da imagem leva a obtenção de resultados mais precisos. Para aumentar a precisão do processo de localização podem ser usados métodos de deteção de bordas e de vetorização mais robustos e parâmetros intrísecos de melhor qualidade.

Para avaliar a repetibilidade deste experimento outras duas imagens foram tomadas em condições idênticas, com pequenas variações na iluminação; o processo recursivo foi executado para cada imagem, sendo extraida a média dos parâmetros da câmara para os três experimentos.

Na figura 4.3.8 são mostrados os resultados obtidos com as três imagens semelhantes, tomadas sequencialmente com pequenas variações na iluminação. No gráfico (a) são apresentados os desvios em relação à media para os elementos de rotação, bem como os desvios padrão estimados. No gráfico (b) elementos análogos são apresentados para as rotações.



(a) (b)

Figura 4.3.8 Resultados para 3 imagens, tomadas na mesma posição.

Verifica-se, da análise destes gráficos, que os desvios dos elementos do vetor estado, em relação à media, são de mesma ordem de grandeza dos valores estimados pela matriz covariância. Isto indica uma boa avaliação da precisão obtida. Contudo, não é possível avaliar se existe uma tendência nas estimativas, pois não são conhecidos os valores verdadeiros do vetor estado.

Em todos os três experimentos os resíduos preditos padronizados apresentaram, em sua grande maioria, valores inferiores a 1. Para estes mesmos experimentos, foram calculados os elementos do vetor estado e sua matriz covariância, usando o filtro em batch (MMQ); foram observados valôres de mesma ordem de grandeza com variações, em relação ao IEKF, da ordem de 1' para as rotações de 0.3mm para as translações.

4.3.5 Resultados com uma sequência de imagens.

O experimento a ser discutido neste ítem procura simular o movimento de um cubo em uma esteira, com a câmara fixa coletando as imagens. Este tipo de problema aparece tanto em Visão "eye-in-hand", quanto em Visão "eye-off-hand". No caso "eye-in-hand" o garra do manipulador poderia se mover e imagear um objeto fixo; eventualmente, a garra poderia estar em repouso aguardando a resposta do sistema de visão sobre a localização de um determinado objeto. No caso "eye-off-hand" a câmara mantém-se fixa e busca-se monitorar o movimento dos objetos.

No experimento em questão, o cubo foi colocado sobre um mesmo plano, em quatro posições sucessivas, aproximadamente alinhadas em X, com uma diferença de 70 mm em Y, entre as posições sucessivas. Como o cubo foi deslocado manualmente, existe um erro de estado que simularia uma perturbação de estado de um ambiente real. Pode-se supor, a priori que o erro de posicionamento do objeto é de, aproximadamente, 2° nas rotações e de 5 mm nas translações. Entretanto, para evitar a possibilidade de divergência do filtro admitiu-se que o erro de transição de estado, no processo recursivo é de 2° para as rotações e 10 mm para as translações.

As imagens de bordas desta sequência de imagens são mostradas na figura 4.3.9. Observa-se, nesta figura, que nas duas primeiras imagens duas arestas não foram detectadas. Isto ocorreu devido à iluminação inadequada, que produziu um baixo contraste entre as duas regiões que definem a borda. Como foi utilizado um limiar único para toda a imagem, a magnitude do gradiente nestas duas retas foi muito menor que nas demais. Se este limiar fosse reduzido, ocorreria a presença de ruído em outras regiões.

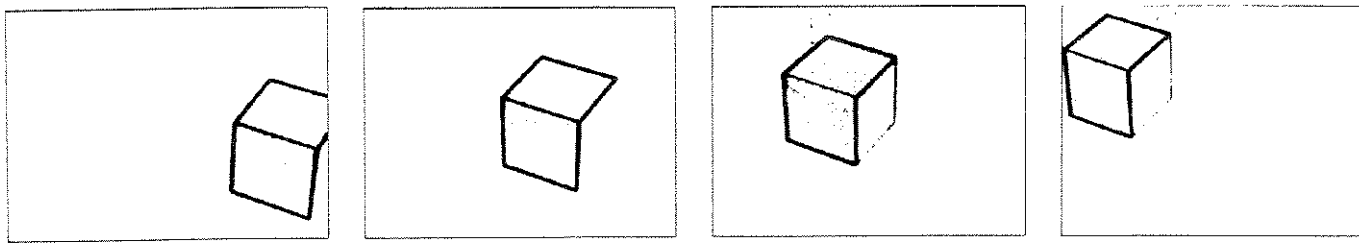


Figura 4.3.9 Sequência de imagens de bordas simulando o cubo em movimento.

Para dar início ao processo recursivo foram utilizadas estimativas preditas com erros de 3° para as rotações e de 10mm para as translações. Na tabela 4.3.7 são apresentados os valores do estado filtrado para cada uma das imagens e seus respectivos desvios padrão estimados, para cada imagem. É importante lembrar que o estado filtrado serve para estabelecer o estado predito da próxima imagem. A matriz covariância predita é calculada através de propagação de covariâncias, considerando-se a matriz covariância filtrada da imagem anterior e os erros de transição de estado.

Tabela 4.3.7 Resultados do processo recursivo para uma sequência de imagens.

	IMAGEM 1		IMAGEM 2		IMAGEM 3		IMAGEM 4	
	Estado Filtrado	Desvio Padrão	Estado Filtrado	Desvio Padrão	Estado Filtrado	Desvio Padrão	Estado Filtrado	Desvio Padrão
κ	2.30021	0.00431	2.31051	0.00364	2.33190	0.00234	2.31191	0.00218
ϕ	0.77688	0.00580	0.77991	0.00473	0.77385	0.00386	0.79644	0.00334
ω	-0.55384	0.00489	-0.56842	0.00436	-0.57741	0.00315	-0.57211	0.00296
X_c	621.534	4.735	619.001	3.930	614.039	2.954	628.627	2.987
Y_c	245.242	2.540	321.555	2.522	397.334	2.382	460.840	2.603
Z_c	524.301	4.174	516.031	3.464	516.407	3.081	505.913	2.708

Analisando a tabela 4.3.7 é possível verificar a redução lenta dos desvios padrão estimados à medida que as imagens são processadas. Não é possível verificar o erro verdadeiro, pois os valores reais da localização do cubo não são conhecidos.

Para avaliar as discrepâncias entre as posições foram calculadas as diferenças do vetor estado entre imagens vizinhas. Como o cubo descreveu um movimento aproximadamente linear seriam esperadas diferenças nulas exceto em Y, que seria de 70mm. Na tabela 4.3.8 são mostradas estas diferenças. Com relação às rotações, observam-se diferenças da ordem de 10' a 1°15', o que e' compatível com o erro cometido no posicionamento manual do cubo. Com relação às translações, constatam-se diferenças desde 0.4mm até 14.5mm, o que merece algumas considerações adicionais.

Tabela 4.3.8 Diferenças entre os vetores estado para imagens vizinhas.

	2 - 1	3 - 2	4 - 3
$\Delta\kappa$	0.010306	0.021393	-0.019996
$\Delta\phi$ rad	0.003027	-0.006060	0.022586
$\Delta\omega$	-0.014580	-0.008992	0.005294
ΔX_c	-2.533	-4.962	14.588
ΔY_c mm	76.313	75.779	63.506
ΔZ_c	-8.270	0.376	-10.495

As maiores diferenças, tanto nas rotações quanto nas translações, foram verificadas entre as imagens 3 e 4. Na quarta imagem, como pode-se observar na figura 4.3.8, o cubo está localizado no canto superior direito do frame; como foi verificado nas simulações estas configurações são, potencialmente, problemáticas. Outra explicação que pode ser dada para erros em translação é um possível erro na distância focal calibrada da câmara. As simulações mostraram que o erro na distância focal causa um efeito sistemático nas translações, mas pouco afeta as rotações. Além disto, nesta sequência, as imagens estão ligeiramente desfocadas, o que não permite um nível de precisão semelhante aos experimentos dos ítens anteriores.

Contudo, não deve-se confundir estas diferenças com o erro verdadeiro de posicionamento do cubo em cada instante. Os desvios padrão apresentados na tabela 4.3.7 procuram indicar a precisão do vetor estado para aquela determinada posição, e não para as diferenças. As diferenças podem ter um efeito acumulativo.

Uma conclusão importante deste experimento, é que o processo recursivo proposto atinge seus objetivos, conseguindo um tratamento dinâmico para o problema de localização. O processo é suficientemente robusto para tratar

imagens com ruído e ausência de algumas retas na imagem. Imagens de melhor qualidade e ambientes mais estruturados podem levar a um aumento na precisão.

4.3.6 SUMÁRIO DAS CONCLUSÕES OBTIDAS NOS EXPERIMENTOS COM DADOS REAIS.

Através dos experimentos com dados reais, procurou-se simular um problema protótipo, que ocorre com frequência nas aplicações de Visão.

Verificou-se que a precisão obtida é um pouco inferior àquela verificada em experimentos análogos com simulações. Como não são conhecidos os valores verdadeiros dos parâmetros em experimentos reais, não foi possível avaliar a exatidão dos resultados, isto é, seu desvio em relação aos valores verdadeiros (erro verdadeiro).

Em imagens de boa qualidade foi possível obter desvios estimados nas rotações da ordem de 5' para as rotações e de 1,5 mm para as translações. Em imagens desfocadas ou em movimento, esta qualidade diminui, devido ao baixo contraste das bordas, o que aumenta a dispersão no ajuste das retas.

A utilização do método de agrupamento θ-ρ permitiu uma considerável eficiência na detecção das retas. O ajuste final pelo MMQ, com verificação de pontos espúrios e sua eliminação, reduziu sensivelmente a possibilidade de erros grosseiros nas observações. Esta possibilidade acontece quando existem duas retas semelhantes na janela de busca; a análise do resíduo padronizado permite detectar estes enganos. Devido a estes procedimentos, algumas observações com variâncias super-estimadas foram rejeitadas, em alguns casos. O aumento da variância das observações, evita estas rejeições incorretas.

Devido ao modelo não linear e à utilização do IEKF, a sequência em que as retas são processadas influi na solução. Para um mesmo conjunto de dados obtém-se soluções diferentes, alterando a ordem de extração das retas. A magnitude destas diferenças dependerá da qualidade da estimativa predita, mas é de grandeza inferior aos desvios estimados.

No processamento de uma sequência de imagens, verificou-se a necessidade de introduzir erros de estado maiores que os reais, para prevenir a divergência do filtro. Isto ocorre porque as variâncias das translações são subestimadas, especialmente nas situações em que o cubo aparece nos cantos da imagem. Este tipo de problema foi verificado nas simulações, causado, possivelmente, por erros de calibração, especialmente na distância focal.

4.4 REQUISITOS DE PRECISÃO EM APLICAÇÕES DE VISÃO DE MÁQUINA.

Uma vez verificada a exequibilidade do processo recursivo e a sua robustez, é necessário analisar a sua aplicabilidade a problemas reais, onde o problema de localização deva ser resolvido. Esta aplicabilidade deve ser verificada com relação aos requisitos de precisão da aplicação e do tempo de resposta exigidos do sistema de Visão.

Na tabela 4.4.1 são apresentadas aplicações do processo de localização a alguns problemas práticos. É possível verificar que o processo recursivo, proposto neste trabalho, atende aos requisitos de precisão destas aplicações. Somente o 4º exemplo exige um nível de precisão que não foi alcançado nos experimentos práticos. Na aplicação relatada (Gruen, 1992), entretanto, foram utilizadas 4 câmaras para atingir o nível de precisão requerida.

Tabela 4.4.1 Características de algumas aplicações envolvendo localização.

APLICAÇÃO	PRECISÃO REQUERIDA		Tempo de resposta	Distância da câmara	Tamanho do objeto
	Posição	Orient.			
Localização de componentes para montagem de PCIs (Capson, 1990)	1mm	1°	4 s	120mm	30x30x0.76mm
Localização de unidades de deflexão (Browne, 1986)	1mm	1°	0.5s	---	---
Localização de junta de carburador (Feddema, 1991)	2mm	2°	100ms	215mm	75x75mm
Localização de estruturas de carros para aplicação de selagem (4 câmaras). (Gruen, 1992)	0.4mm	--	5s	1.5-2.5m	carro
Localização de satélite para o manipulador do "Space Shuttle" (Gruen, 1992)	5mm	0.1°	47ms	---	Satélite

Utilizando o processo recursivo proposto, e adotando estratégias de melhoria da precisão como, por exemplo, aumentar o número de feições, usar várias câmaras e melhorar a qualidade das medidas na imagem, é possível atingir o nível de precisão requerido pela aplicação mencionada. Com relação ao tempo de resposta, seria necessário acelerar o processamento, utilizando máquinas mais eficientes, melhorando os algoritmos ou recorrendo ao processamento paralelo.

CAPÍTULO V

5. CONCLUSÕES

5.1 SÍNTESE DO TRABALHO DESENVOLVIDO

Neste trabalho foi proposto um processo recursivo para a solução do problema de localização, baseado em Filtragem Kalman, usando como apoio linhas retas, cujas equações paramétricas são conhecidas no espaço objeto, e cujas projeções são medidas na imagem.

Um modelo matemático explícito foi desenvolvido, permitindo o cálculo simultâneo dos seis parâmetros de localização. Este modelo difere da maioria dos modelos referenciados na literatura para este tipo de feições que, normalmente, são do tipo implícito e com os parâmetros de localização calculados em mais de um estágio.

O tratamento deste modelo usando filtragem Kalman (IEKF) permite a obtenção de estimativas sequenciais, que podem ser utilizadas para realimentar o processo de extração de feições, pela redução da janela de busca da feição na imagem. Este conceito de redução gradativa da janela de busca, também não é mencionado na literatura.

Foram implementados programas em linguagem C, para todas as etapas do processo. Apenas a detecção de bordas das imagens foi realizada "off-line", embora já existam soluções que permitam esta operação em tempo real.

O procedimento proposto foi testado com dados simulados e com dados reais. Os resultados foram analisados, permitindo uma avaliação global do modelo, dos níveis de precisão alcançados, bem como das situações potencialmente problemáticas.

5.2 CONCLUSÕES

Considerando o modelo matemático, verificou-se a sua adequação às aplicações em Visão de Máquina, e suas vantagens em relação aos demais modelos existentes:

- o modelo é do tipo explícito, isto é, as observações são dadas como uma função explícita dos parâmetros de localização. Não foi encontrado modelo semelhante na literatura;
- o modelo é simultâneo, isto é, todos os parâmetros extrínsecos são calculados em uma única etapa, evitando a propagação de erros entre estágios, que ocorre na maioria dos métodos existentes;
- entretanto, o fato do modelo ser não linear, exige estimativas preditas de boa qualidade;

A natureza explícita do modelo permite o emprego direto da Filtragem Kalman, possibilitando a utilização mesmo em configurações deficientes, conferindo robustez ao sistema. Estas configurações ocorrem com frequência em ambientes industriais, e um sistema robusto deve estar preparado para trata-las adequadamente.

Além disto, problemas com características dinâmicas recebem uma solução adequada, devido ao emprego da Filtragem Kalman e do modelo proposto.

A característica recursiva da Filtragem Kalman permitiu desenvolver um processo iterativo, no qual as estimativas filtradas geradas pelo processamento sequencial das observações, são usadas para realimentar a etapa de extração de feições, reduzindo o custo computacional desta etapa.

Com relação ao tempo de processamento, verificou-se que a fase crítica no processo é a vetorização. O custo computacional do processo de estimação de parâmetros é pequeno, quando comparado ao da vetorização. Isto justifica fortemente o emprego da abordagem proposta; conclui-se, que o emprego de técnicas de estimação de parâmetros, além de emprestar confiabilidade ao processo, é instrumento de redução no custo computacional. Para atingir o processamento em tempo real, ou seja, fornecer respostas na mesma escala de tempo da amostragem, algumas estratégias adicionais poderiam ser implementadas:

- utilizar máquinas de alto desempenho e processamento paralelo, nas etapas em que isto for possível;

. aumentar a eficiência da implementação, reduzindo as operações em ponto flutuante, introduzindo tabelas pré-calculadas para funções trigonométricas, etc...;
. utilizar métodos de Filtragem Kalman que empreguem formas fatoradas da matriz covariância (Bierman, 1977);

A existência de estimativas preditas iniciais de boa qualidade merece algumas considerações:
. do ponto de vista da Filtragem, mesmo com estimativas preditas com erros de grande magnitude, o processo converge. Entretanto:
. estimativas preditas muito grosseiras definiriam uma janela de busca inicial muito grande, englobando toda a imagem; isto eliminaria as vantagens introduzidas pela realimentação permitida pela filtragem;

As estimativas preditas são obtidas, dependendo do problema a ser tratado, das seguintes maneiras:
. no problema de localização de objetos, a etapa de reconhecimento fornece uma estimativa de localização; heurísticas adicionais podem ser utilizadas, como as posições de estabilidade dos objetos possíveis;
. no problema de localização da câmara, e mais especificamente no caso de aplicação em Robótica, na configuração "eye-in-hand", uma boa estimativa "a priori" é obtida a partir do modelo cinemático do manipulador;

A abordagem desenvolvida encontra várias aplicações, principalmente nos Sistemas industriais de Visão de Máquina, em veículos auto-dirigidos (AGV's) e em Fotogrametria e Cartografia.

A análise dos resultados obtidos, tanto com dados reais quanto com dados simulados, permite concluir que o procedimento proposto funciona adequadamente, atingindo níveis de precisão compatíveis com a maioria das aplicações citadas.

Como exemplo, num experimento com a câmara posicionada a 1,13 m do objeto foram obtidos erros nas rotações de aproximadamente 5' e de 1,5 mm nas translações. Estes níveis de precisão podem ser aumentados, caso a aplicação o exija, tomando-se alguns cuidados:
. melhorar a qualidade das medidas que definem o modelo do objeto, isto é, as equações das retas no espaço objeto;
. utilizar câmaras de melhor qualidade, com alta estabilidade;

. restringir a localização de objetos para profundidades de campo que evitem imagens excessivamente desfocalizadas;
. utilizar as 400 linhas de resolução (nos experimentos com dados reais foram usadas apenas linhas alternadas, para otimizar o uso de memória e acelerar o processamento);
. adotar métodos de detecção de bordas mais precisos que os operadores de Sobel, embora, possivelmente, a custo computacional maior;
. aumentar a confiabilidade dos parâmetros intrínsecos, através da utilização de um número maior de pontos e de imagens no processo de calibração; verificou-se que alguns erros nas translações podem ser causados por erros na distância focal, o que pode, eventualmente, ocasionar a divergência do filtro no processamento de uma sequência de imagens;

Nos experimentos conduzidos, a possibilidade de divergência do filtro foi verificada em algumas situações, sendo que as mais importantes são:
. estimativa incorreta da variância das observações;
. estimativa a priori com tendência;
. erros grosseiros nas observações;
. parâmetros intrínsecos pouco confiáveis;
. erros de transição de estado sub-estimados;

5.3 Recomendações

Alguns pontos abordados neste trabalho merecem estudos adicionais, embora o processo funcione a contento na disposição atual.

Recomenda-se estudar em detalhes as seguintes questões:

. a obtenção de estimativas preditas de boa qualidade através de integração com outros sensores;
. métodos de detecção de bordas e vetorização, exeqüíveis em tempo real, que explorem as características do procedimento recursivo, como, por exemplo, a existência de uma estimativa "a priori" da localização do objeto;
. a possibilidade de integrar o processo recursivo proposto com o reconhecimento de objetos; isto seria feito extraindo-se um conjunto mínimo de primitivas de indexação, estabelecendo-se uma hipótese de localização e calculando-se esta localização usando o modelo simultâneo e a Filtragem em

Batch; seria, então, verificada a hipótese estabelecida, usando as projeções das demais retas com os parâmetros calculados;

- a questão da sensibilidade do modelo matemático desenvolvido em relação aos erros nos parâmetros intrínsecos, uma vez que a primeira equação do modelo não é afetada nem pelos erros no centro da imagem nem pelos erros na distância focal;
- a possível extensão do modelo matemático proposto a um método de determinação dos parâmetros intrínsecos (calibração);
- a utilização de uma abordagem de filtragem que leve em conta a incerteza em alguns parâmetros, até agora considerados invariantes, como, por exemplo, a distância focal, e verificar a qualidade dos resultados e o aumento do custo computacional decorrente disto;

APÊNDICE 1

1.1 TRANSFORMAÇÕES

1.1.1 TRANSLAÇÕES

A transformação do espaço objeto para o espaço imagem pode ser realizada em três passos. O primeiro passo leva em conta a diferença espacial entre as origens, e compensa esta diferença por meio de três translações.

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} - \begin{bmatrix} Xc \\ Yc \\ Zc \end{bmatrix} \qquad (1.1.1)$$

1.1.2 ROTAÇÕES

O segundo passo da transformação procura compensar as diferenças angulares entre os sistemas imagem e sistema objeto transladado, podendo ser expressa por:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = M \begin{bmatrix} U \\ V \\ W \end{bmatrix} \qquad (1.1.2)$$

A matriz M de rotação pode ter várias formas, dependendo da escolha da sequência das rotações. Uma das formas mais usuais é a sequência $Mz(\kappa)$ $My(\phi)$

$Mx(\omega)$, isto é, aplica-se a rotação primária ω em torno do eixo X, gerando um novo referencial. Sobre o eixo Y deste novo referencial aplica-se a rotação ϕ e, finalmente, sobre o eixo Z resultante, aplica-se a rotação κ.

Considerando as matrizes de rotação:

$$Mx(\omega) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & \operatorname{sen}\omega \\ 0 & -\operatorname{sen}\omega & \cos\omega \end{bmatrix} \qquad (1.1.3)$$

$$My(\phi) = \begin{bmatrix} \cos\phi & 0 & -\operatorname{sen}\phi \\ 0 & 1 & 0 \\ \operatorname{sen}\phi & 0 & \cos\phi \end{bmatrix} \qquad (1.1.4)$$

$$Mz(\kappa) = \begin{bmatrix} \cos\kappa & \operatorname{sen}\kappa & 0 \\ -\operatorname{sen}\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (1.1.5)$$

Assumindo a sequência $Mz(\kappa)\ My(\phi)\ Mx(\omega)$, a matriz de rotação final será:

$$\begin{bmatrix} \cos\phi\cos\kappa & \cos\omega\operatorname{sen}\kappa + \operatorname{sen}\omega\operatorname{sen}\phi\cos\kappa & \operatorname{sen}\omega\operatorname{sen}\kappa - \cos\omega\operatorname{sen}\phi\cos\kappa \\ -\cos\phi\operatorname{sen}\kappa & \cos\omega\cos\kappa - \operatorname{sen}\omega\operatorname{sen}\phi\operatorname{sen}\kappa & \operatorname{sen}\omega\cos\kappa + \cos\omega\operatorname{sen}\phi\operatorname{sen}\kappa \\ \operatorname{sen}\phi & -\operatorname{sen}\omega\cos\phi & \cos\omega\cos\phi \end{bmatrix} \qquad (1.1.6)$$

1.1.3 ESCALA

O terceiro e último passo na transformação de coordenadas imagem para coordenadas objeto é considerar a diferença de escala entre os referenciais que, após os passos anteriores tem a mesma origem e orientação. Isto é feito através do fator de escala λ.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} \qquad (1.1.7)$$

Substituindo, a forma final das transformações será:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \; M \begin{bmatrix} X - Xc \\ Y - Yc \\ Z - Zc \end{bmatrix} \qquad (1.1.8)$$

1.2 EQUAÇÕES PROJETIVAS

Dividindo x e y por z obtém-se duas equações onde o fator de escala λ é eliminado. Estas equações, conhecidas como equações de colinearidade ou equações projetivas podem ser obtidas por semelhança de triângulos considerando a colinearidade entre um ponto no espaço objeto, seu homólogo na imagem e o centro perspectivo. Considera-se, ainda, que a coordenada z no plano imagem, é constante e igual à distância focal da câmara (f).

$$x_i = f \; \frac{(X - Xc)\, m_{11} + (Y - Yc)\, m_{12} + (Z - Zc)\, m_{13}}{(X - Xc)\, m_{31} + (Y - Yc)\, m_{32} + (Z - Zc)\, m_{33}} = f \, \frac{m}{q} \qquad (1.1.9)$$

$$y_i = f \; \frac{(X - Xc)\, m_{21} + (Y - Yc)\, m_{22} + (Z - Zc)\, m_{23}}{(X - Xc)\, m_{31} + (Y - Yc)\, m_{32} + (Z - Zc)\, m_{33}} = f \, \frac{n}{q}$$

onde os m_{ij} são os elementos da matriz de rotação.

Existe a possibilidade de se calcular um valor para a coordenada z, após a transformação projetiva, que é usada na remoção de linhas escondidas. O desenvolvimento a sequir é adaptado de Newman e Sproull (1973, p. 268).

A figura A.1.1 mostra uma lente de distância focal f, posicionada na origem do referencial imagem. A imagem do ponto P é um ponto P', cujas coordenadas podem ser obtidas por semelhança de triângulos. Os elementos x_i e y_i já foram apresentados através das equações (1.1.9).

Através de semelhança de triângulos estabelece-se:

$$x_i = f.\frac{x}{z} \qquad y_i = f.\frac{y}{z} \qquad \text{e}$$

$$\frac{y}{f} = \frac{y_i}{(z_i - f)} \qquad (1.1.10)$$

Substituindo:

$$\frac{y}{f} = \frac{f\frac{y}{z}}{(z_i - f)} \qquad (1.1.11)$$

Finalmente, após algumas manipulações algébricas:

$$z_i = f - \frac{f^2}{z} \qquad (1.1.12)$$

A coordenada z_i, associada a x_i e y_i, define um volume deformado pela perspectiva.

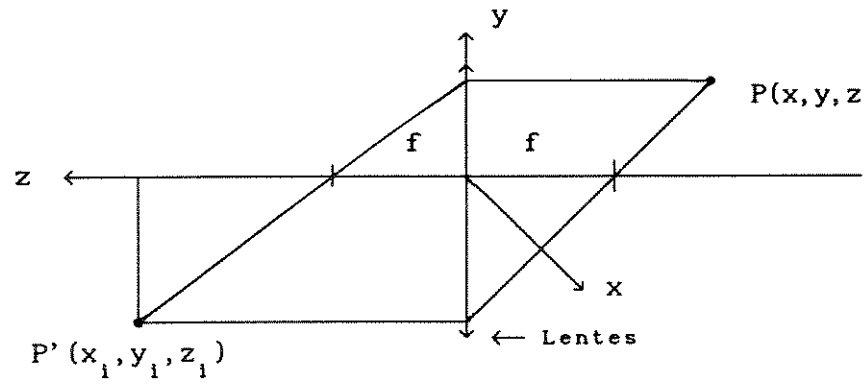


Figura A.1.1 Projeção perspectiva.

5. REFERÊNCIAS BIBLIOGRÁFICAS

ABDEL-AZIZ, Y.I.; KARARA, H. M. Photogrammetric potential of non-metric cameras. **Univ. Illinois at Urbana-Champaign, Civil Eng. Studies,** 1974, Photogrammetry Series, 36.

ALOIMONOS, J. Visual shape computation. **Proceedings of the IEEE**, v. 76, n. 8, p. 899-916, 1988.

BALLARD, D.H.; BROWN, C.M. **Computer Vision.** New Jersey: Prentice-Hall, 1982. 522 p.

BARNARD, S. T. Interpreting perspectives images. **Artificial Intelligence**, v. 21, n. 4, p. 435-462, 1983.

BERGEVIN, R.; LEVINE, M.D. Extraction of line drawing features for object recognition. **Pattern Recognition**, v. 25, n. 3, p. 319-334, 1992.

BIERMAN, G. J. **Factorization methods for discrete sequential estimation.** San Diego, CA: Academic Press, 1977. 241 p.

BROIDA, T.J.; CHELLAPPA, R. Estimation of object motion parameters from noisy images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 8, n. 1, p. 90-98, 1986.

CAPRILE, B.; TORRE, V. Using vanishing points for camera calibration. **International Journal of Computer Vision**, n. 4, p. 127-140, 1990.

CHEN, S.Y.; TSAI, W.H. Determination of robot locations by common object shapes. **IEEE Transactions on Robotics and Automation**, v. 7, n. 1, p. 149-156, 1991.

CHEN,Z.; TSENG, D.; LIN, J. A simple vision algorithm for 3-d position determination using a single calibration object. **Pattern Recognition**, v. 22, n. 2, p. 173-187, 1989.

CHEN, W.; JIANG, B. C. 3-D camera calibration using vanishing point concept. **Pattern Recognition**, v. 24, n. 1, p. 57-67, 1991.

DHOME, M.; RICHETIN, M.; LAPRESTÉ, J.; RIVES, G. Determination of the attitude of 3-D objects from a single perspective view. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 11, n. 12, p. 1265-1278, 1989.

DICKINSON, S.J. From volumes to views: an approach to 3-D object

recognition. **Computer Vision, Graphics, and Image Processing; Image Understanding**, v. 55, n. 2, p. 130-154, 1992.

DUDANI, S.A.; LUK, A.L. Locating straight-line edge segments on outdoor scenes. **Pattern Recognition**, v. 10, p. 145-157, 1978.

ECHIGO, T. A camera calibration technique using three sets of parallel lines. **Machine Vision and Applications**, n. 3, p. 159-167, 1990.

FAIG, W.; MONIWA, H. Convergent photos for close range. **Photogrammetric Engineering**, v. 39, n. 6, 1973.

FEDDEMA, J. T.; LEE, C. S. G. Adaptive image feature prediction and control for visual tracking with a hand-eye coordinated camera. **Transactions on Systems, Man, and Cybernetics**, v. 20, n. 5, p. 1173-1183, 1990.

FEDDEMA, J.T.; LEE, C.S.G.; MITCHELL, O.R. Weighted selection of image features for resolved rate visual feedback control. **IEEE Transactions on Robotics and Automation**, v. 7, n. 1, p. 31-47, 1991.

FISCHLER, A. M.; BOLLES, R. C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. **Communications of the ACM**, v. 24, n. 6, p. 381-395, 1981.

GOLDGOF, D.B.; LEE, H.; HUANG. T.S. Matching and motion estimation of three-dimensional point and line sets using eigenstruture without correspondences. **Pattern Recognition**, v. 25, n. 3, p. 271-286, 1992.

GRUEN, A. W. Towards real-time photogrammetry. **Photogrammetria**, n. 42, p. 209-244, 1988.

HALARICK, R. M. Using perspective transformations in scene analysis. **Computer Graphics and Image Processing**, v. 13, p. 191-221, 1980.

HALARICK, R. M. Determining camera parameters from the perspective projection of a rectangle. **Pattern Recognition**, v. 32, n. 3, p. 225-230, 1989.

HALARICK, R. M. et al Pose estimation from corresponding point data. **Transactions on Systems, Man, and Cybernetics**, v. 19, n. 6, p. 1246-1446, 1989.

HAN M.H.; RHEE, S. Camera calibration for three-dimensional measurement. **Pattern Recognition**, v. 25, n. 2, p. 155-164, 1992.

HOLT, R. J.; NETRAVALI, A. N. Camera calibration problem: some new results. **Computer Vision, Graphics, and Image Processing: Image Understanding**, v. 54, n. 3, p. 368-383, 1991.

JAZWINSKI, A.H. **Stochastic Processes and Filtering Theory**. New York: Academic Press, 1970. 376 p.

KAMGAR-PARSI, B.; KANGAR-PARSI, B. Evaluation of quantization error in computer vision. **IEEE Transactions on Pattern Analysis and Machine**

Intelligence, v. 11, n. 9, p. 929-940, 1989.

KITTLER, J.; EGGLETON, T.; ILLINGWORTH, J.; PALER, K. An averaging edge detector. **Pattern Recognition Letters**, v. 6, n. 1, p. 27-32, 1987.

KRIEGMAN, D. J. Computing stable poses of piecewise smooth objects. **Computer Vision, Graphics, and Image Processing; Image Understanding**, v. 55, n. 2, p. 109-118, 1992.

LEE, R.; LU, P.C; TSAI, W.H. Robot location using single views of retangular shapes. **Photogrammetric Engineering and Remote Sensing**, v. 56, n. 2, p. 231-238, 1990.

LENZ, R. K.; TSAI, R. Y. Calibrating a cartesian robot with eye-on-hand configuration independent of eye-to-hand relationship. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 11, n. 9, p. 916-928, 1989.

LIANG, P.; CHANG, Y. L.; HACKWOOD, S. Adaptive self-calibration of vision-based robot systems. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 19, n. 4, p. 811-824, 1989.

LINNAINMAA, S.; HARWOOD, D.; DAVIS, L. S. Pose determination of a three-dimensional object using triangle pairs. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 10, n. 5, p. 634-647, 1988.

LIU, Y.; HUANG, T. S. Estimation of rigid body motion using straight line correspondences. **Computer Vision, Graphics and Image Processing**, v. 43, n. 1, p. 37-52, 1988-a.

LIU, Y.; HUANG, T. S. A linear algorithm for motion estimation using straight line correspondences. **Computer Vision, Graphics and Image Processing**, v. 44, n. 1, p. 35-37, 1988-b.

LIU, Y.; HUANG, T. S. Determining straight line correspondences from intensity images. **Pattern Recognition**, v. 24, n. 6, p. 489-504, 1991.

LIU, Y.; HUANG, T. S.; FAUGERAS, O. D. Determination of camera location from 2-D to 3-D line and point correspondences. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 12, n. 1, p. 28-37, 1990.

LOWE, D. G. Three-dimensional object recognition from single two-dimensional images. **Artificial Intelligence**, v. 31, p. 355-395, 1987.

LUGNANI, J.B. Using digital entities as control. Fredericton, 1980, 159 p., PhD Thesis, Department of Surveying Engineering, The University of New Brunswick.

MARR, D. **Vision**. San Francisco: Freeman, 1982.

MASRY, S. E. Digital mapping using entities: A New Concept. **Photogrammetric Engineering and Remote Sensing**, v. 48, n. 11, p. 1561-1565, 1981.

MITICHE, A.; HABELRIH. Interpretation of straight line correspondences using

angular relations. **Pattern Recognition**, v. 22, n. 3, p. 299-308, 1989.

MITICHE, A.; FAUGERAS, O.; AGGARWAL, J. K. Counting straight lines. **Computer Vision, Graphics and Image Processing**, v. 47, p. 353-360, 1989.

MULAWA, D. C.; MIKHAIL, E. M. Photogrammetric treatment of linear features. In: **Proceedings of the International Society for Photogrammetry and Remote Sensing**, 1988, Kyoto, Commission III, p. 383-393.

NEWMAN, W.M.; SPROULL, R.F. **Principles of interactive computer graphics**. 1ª ed., New York: McGraw-Hill, 1973.

PAUL, R. P. **Robot manipulators: mathematics, programming and control**. Cambridge, MA: MIT Press, 1981.

PENNA, M. Determining camera parameters from the perspective projection of a quadrilateral. **Pattern Recognition**, v. 24, n. 6, p. 533-541, 1991.

ROBERTS, L. G. Recognition of three-dimensional objects. In: **The Handbook of Artificial Intelligence**, Reading, MA: Addison-Wesley, 1982, v. III, p. 139-142.

ROGERS, D. F. **Procedural Elements for Computer Graphics**. Singapore: McGraw-Hill, 1985. 433 p.

SALARI, E.; JONG, C. A method to calculate the structure and motion parameters from line correspondences. **Pattern Recognition**, v. 23, n. 6, p. 553-561, 1990.

SPETSAKIS, M.E.; ALOIMONOS, J. Structure from motion using line correspondences. **International Journal of Computer Vision**, v. 4, p. 171-183, 1990.

TEUNISSEN, P.J.G.; SALZMANN, M.A. A recursive slippage test for use in state-space filtering. **Manuscripta Geodaetica**, n. 14, p. 383-390, 1989.

TOMMASELLI, A.M.G.; LUGNANI, J. B. An alternative mathematical model to the collinearity equation using straight features. In: **Proceedings of the International Society for Photogrammetry and Remote Sensing**, 1988, Kyoto, Commission III, p. 765-774.

TOMMASELLI, A.M.G.; TOZZI, C.L. Aplicações industriais de visão de máquina. **Faculdade de Engenharia Elétrica, Universidade Estadual de Campinas (Unicamp)**, 1990, 60 p., (Publicação Interna), n. RT-DCA-019/90.

TOMMASELLI, A.M.G.; TOZZI, C.L. Um estudo prático sobre calibração de câmaras. **Faculdade de Engenharia Elétrica, Universidade Estadual de Campinas (Unicamp)**, 1991, 83 p., (Publicação Interna), n. RT-DCA-019/91.

TOMMASELLI, A.M.G.; TOZZI, C. L. A filtering based approach to eye-in-hand robot vision. In: **Proceedings of the International Society for Photogrammetry and Remote Sensing**, 1992, Washington Commission V.

TSAI, ROGER Y. A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. **IBM Research Report RC 51342**, 1985.

TSAI, R. Y.; LENZ, R. K. A new technique for fully autonomous and efficient 3D robotics hand/eye calibration. **IEEE Transactions on Robotics and Automation**, v. 5, n. 3, p. 345-358, 1989.

VERESS, S.A.; YOUCAI, H. A method for improving the efficiency of the sequential estimation procedure in photogrammetry. **Photogrammetric Engineering and Remote Sensing**, v. 53, n. 6, p. 613-616, 1987.

WANG, C.C. Extrinsic calibration of a vision sensor mounted on a robot. **IEEE Transactions on Robotics and Automation**, v. 8, n. 2, p. 161-175, 1992.

WANG, L.L.; TSAI, W.H. Computing camera parameters using vanishing-line information from a rectangular parallelepiped. **Machine Vision and Applications**, v. 3, p. 129-141, 1990.

WANG, Y.F.; KARANDIKAR, N.; AGGARWAL, J.K. Analysis of video image sequences using point and line correspondences. **Pattern Recognition**, v. 24, n. 11, p. 1065-1084, 1991.

WENG, J.; HUANG, T.; AHUJA, N. Motion and structure from line correspondences: closed-form solution, uniqueness, and optimization. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 14, n. 3, p. 318-336, 1992.

WONG, E.K. Model matching in robot vision by subgraph isomorphism. **Pattern Recognition**, v. 25, n. 3, p. 287-303, 1992.

ZENG, Z.; WANG, X. A general solution of a closed-form space resection. **Photogrammetric Engineering and Remote Sensing**, v. 58, n. 3, p. 327-338, 1992.

6. BIBLIOGRAFIA

BEYER, H. A. Linejitter and geometric calibration of CCD-cameras. **ISPRS Journal of Photogrammetry and Remote Sensing**, n. 45, p. 17-32, 1990.

BRADY, M. Computacional approaches to image understanding. **Computing Surveys**, v. 14, n. 1, p. 3-71, 1982.

CRAIG, J. J. **Introduction to robotics: mechanics and control.** Reading, MA: Addison Wesley, 1986. 303 p.

FORSTNER, W. Reliability analysis of parameter estimation in linear models with applications to mensurations problems in computer vision. **Computer Vision, Graphics, and Image Processing**, v. 40, p. 273,310, 1987.

GONZALEZ, R. C.; WINTZ, P. **Digital image processing**. Reading: Addison-Wesley, 1987.

HORN, B. K. P. **Robot vision**. Cambridge, MA: MIT Press and McGraw-Hill, 1986. 509 p.

MIKHAIL, E. M.; ACKERMANN, F. **Observations and least squares**. New York: IEP, 1976. 497 p.

MILLER, K.S.; LESKIW, D.M. **An introduction to Kalman filtering with applications**. Malabar, Florida: Robert E. Krieger Publishing Company, 1987. 113 p.

ROSENFELD, A. Computer vision: basic principles. **Proceedings of the IEEE**, v. 76, n. 8, p. 863-868, 1988.

WEISS, L. E.; SANDERSON, A. C.; NEUMAN, C. P. Dynamic sensor-based control of robots with visual feedback. **IEEE Transactions on Robotics and Automation**, v. RA-3, n. 5, p. 404-417, 1987.