Campus de São Carlos

**ESTIMAÇÃO DE PARÂMETROS DE MOVIMENTO
A PARTIR DE SEQUÊNCIA DE IMAGENS ESTÉREO**

Autor: Rafael do Espírito-Santo

Orientadora: Prof.^{te} Dr.^{te} Vilma Alves de Oliveira

UNIVERSIDADE DE SÃO PAULO

**ESCOLA DE ENGENHARIA
DE SÃO CARLOS**

ESTIMAÇÃO DE PARÂMETROS DE MOVIMENTO A PARTIR DE SEQUÊNCIA DE IMAGENS ESTÉREO

Rafael do Espírito-Santo

**Dissertação apresentada à Escola de
Engenharia de São Carlos da Universidade
de São Paulo, como parte dos requisitos para
obtenção de título de Mestre em Engenharia
Elétrica**

Orientadora: Prof^a. Dr^a. Vilma Alves Oliveira

**São Carlos
1999**

Class.	TESE - EESC
Outt.	2000
Tombo	0139/99

S/S 1063333

21.05

Ficha catalográfica preparada pela Seção de Tratamento
da Informação do Serviço de Biblioteca - EESC/USP

E77e

Espírito-Santo, Rafael do

Estimação de parâmetros de movimento a partir de
seqüência de imagens estéreo / Rafael do Espírito-
Santo. -- São Carlos, 1999.

Dissertação (Mestrado) -- Escola de Engenharia de
São Carlos-Universidade de São Paulo, 1999.

Área: Engenharia Elétrica.

Orientador: Profa. Dra. Vilma Alves Oliveira.

1. Visão computacional. 2. Imagem stereo.
3. Estimação de parâmetros recursiva

FOLHA DE APROVAÇÃO

Candidato: Engenheiro **RAFAEL DO ESPÍRITO SANTO**

Dissertação defendida e aprovada em 02.10.1998
pela Comissão Julgadora:

Prof. Dra. **VILMA ALVES DE OLIVEIRA (Orientadora)**
(Escola de Engenharia de São Carlos - Universidade de São Paulo)

Prof. Dr. **ELDER MOREIRA HEMERLY**
(Instituto Tecnológico de Aeronáutica - ITA)

Prof. Dr. **VALENTIM OBAC RODA**
(Escola de Engenharia de São Carlos - Universidade de São Paulo)

Prof. Doutor **DENIS VINICIUS COURY**
Coordenador da Área de Engenharia Elétrica

JOSÉ CARLOS A. CINTRA
Presidente da Comissão de Pós-Graduação da EESC

Aos meus pais
Nilton e Almerinda

AGRADECIMENTOS

À Prof^a. Dr^a. Vilma Alves Oliveira, pela orientação, dedicação, incentivo e colaboração na realização deste trabalho, especialmente no esclarecimento de pontos nebulosos.

Ao Prof. Dr. Elder Moreira Hemerly, pela orientação inicialmente dada a este trabalho no Instituto Tecnológico da Aeronáutica – ITA, incentivo e colaboração.

Ao Prof. Dr. Flávio Matsuyama pelas opiniões e sugestões. Aos amigos Gerson Antonio Rubin e Roberto Toshihiko Nakane pela colaboração e sugestões dadas. A Rubens Gedraite pelo incentivo.

A Cláudio Camargo pelo apoio e sugestões dadas na parte experimental. Aos amigos Roberto Limão e Luiz Fernando de Andrade pelas opiniões dadas durante a realização dos experimentos.

Ao CTM-SP, Centro tecnológico da Marinha em São Paulo e em especial, ao chefe do Departamento de Tecnologia de Reatores – CF José Carlos Pires Ferreira - pelo incentivo que tem dado aos programas de pós-graduação de seus subordinados.

À Adriana de Oliveira Bueno.

À Patrícia Andréia Paesani.

À Rosane Aranda.

A vocês, pai e mãe.

SUMÁRIO

INTRODUÇÃO.....	01
I.1 Motivação	01
I.2. Conteúdo do trabalho	02
I.3. Ambiente de visão estéreo e componentes.....	04
I.3.1 Condições de realização do experimento com imagens reais.....	04
I.3.2 Condições de realização do experimento com imagens sintéticas.....	05
I.3.3 Interferência das diferentes fontes de ruído no experimento com imagens reais.....	06
I.3.4 Valores máximos e mínimos dos parâmetros que descrevem o movimento do objeto	06
I.3.5 Especificação dos elementos sensores (câmeras) utilizados no experimento com imagens reais.....	07
I.4. Organização do trabalho	07
 PRIMEIROS TRABALHOS EM VISÃO COMPUTACIONAL DESTINADOS À ESTIMAÇÃO DOS MOVIMENTOS CONTIDOS EM UMA CENA.....	 10
II.1. Utilização de pequena sequência de imagens monocular.....	10
II.1.1. Algoritmos que utilizam o fluxo óptico de campo	10
II.1.2. Algoritmos que fazem uso de grandezas geométricas.....	13
II.2. Utilização de longa sequência monocular.....	15
II.2.1. Estimativa de uma sequência de imagens monocular	15
II.3. Estimação do movimento de objetos usando a teoria	

de sistemas de controle em malha fechada.....	18
II.4. Calibração automática de sistema de visão estéreo.....	18
UTILIZAÇÃO DE SEQUÊNCIA DE IMAGENS ESTÉREO RUIDOSA.....	20
III.1. Comentários iniciais.....	20
III.2. Modelos considerados.....	22
III.2.1. Modelo para projeção ortogonal nos planos-imagens das câmeras	Erro!
Indicador não definido.	22
III.2.2. Modelo para observação de pontos ruidosos	31
III.2.3. Modelo da estrutura do objeto	38
III.2.4. Modelos que descrevem os movimentos do objeto	43
III.3. Formulação do problema de estimação recursiva.....	48
III.3.1. Considerações iniciais	48
III.3.2. Sistema de equações dinâmicas do movimento do objeto.....	50
III.3.3. Equação da planta.....	50
III.3.4. Equação de medida.....	55
III.3.4. Equações do IEKF	64
SISTEMA DE AQUISIÇÃO, DIGITALIZAÇÃO E	
PROCESSAMENTO DE IMAGENS	68
IV.1. Comentários iniciais.....	68
IV.2. Montagem do SADI	69
IV.3. Primeiro método de detecção de cantos	71
IV.3.1. Descrição do algoritmo.....	72
IV.4. Segundo método de detecção de cantos	74
IV.4.1. Formulação do problema de otimização	74
IV.4.2. Resolução do problema de otimização a partir do modelo	
gráfico de um canto.....	76
IV.4.3. Generalização do modelo.....	79
IV.5. Técnicas de verificação da correspondência entre imagens adquiridas por	
câmeras distintas	79

IV.5.1. Primeiro método de verificação da correspondência	80
IV.5.2. Segundo método de verificação da correspondência.....	83
IV.6. Erro de digitalização no experimento com imagens reais	85
IV.6.1. Obtenção de R.....	85
IV.7. Seleção da matriz de covariância do ruído de estado Q	87
IV.8. Ambiente do dispositivo de visão estéreo de um manipulador robótico	92
RESULTADOS DE EXPERIMENTOS.....	93
V.1. Comentários iniciais	93
V.2. Avaliação dos métodos de detecção considerados	93
V.3. Avaliação das técnicas de verificação da correspondência entre imagens.....	100
V.3.1. Avaliação da técnica pertencente à classe Estereoscopia Baseada em Área	101
V.3.2. Avaliação da técnica pertencente à classe Estereoscopia Baseada em Atributos.....	102
V.4. Resultados experimentais com modelo de obtenção de pontos ruidosos.....	103
V.5. Resultados experimentais com sequência de imagens reais.....	104
V.5.1. Movimento do objeto de teste no espaço tridimensional	105
V.5.2. Processo de aquisição de sequência de imagens estéreo	107
V.5.3. Processo de calibração da câmara utilizada no SADI	107
V.5.4. Estimação paramétrica	111
CONCLUSÃO	131
REFERÊNCIAS BIBLIOGRÁFICAS.....	135
APÊNDICE	

LISTA DE FIGURAS

Fig. 3.1. Observação de um corpo rígido em movimento por duas câmeras	24
Fig. 3.2. Localização de um ponto em dois sistemas coordenados paralelos	26
Fig. 3.3. Posição dos eixos após rotação	29
Fig. 3.4. Visão estéreo	35
Fig. 3.5. Vetores diretores de S	40
Fig. 3.6. O sistema fixo no objeto	41
Fig. 4.1. Montagem para aquisição de imagens	70
Fig. 4.2. Modelo gráfico de um canto	76
Fig. 4.3. Configuração do SADI	90
Fig. 4.4. Execução sequencial do procedimento de estimação paramétrica	91
Fig. 4.5 Ambiente proposto para visão estéreo do manipulador	92
Fig. 5.1. Imagens resultantes após a segmentação feita pelo método do gradiente. (1) é a imagem triangular segmentada e (2) é a imagem retangular segmentada. No processo de detecção os limiares utilizados em cada caso foram respectivamente 0.002 e 0.07 valores menores do que estes não produziam melhorias nos desempenhos das detecções	97
Fig. 5.2. Imagens resultantes após a segmentação feita pelo detetor ótimo. (1) é a imagem triangular segmentada e (2)	

é a imagem retangular segmentada.....	97
Fig. 5.3. Quadros utilizados no estudo da correspondência . Atributos não pertencentes ao objeto de teste são os reflexos da fonte de luz na mesa de teste. O objeto de teste encontra-se mais próximo da câmera da esquerda.....	101
Fig. 5.4. Diagrama do modelo generalizado da câmera	108
Fig. 5.5. Diagrama do fator de escala	110
Fig. 5.6. Gráfico da posição do centro de massa do objeto. Os componentes x e y são idênticos, pois o objeto seguiu uma trajetória diagonal, cruzando o eixo ótico da câmera. O componente z foi sempre igual a 1,15cm, pois não ouve movimento na direção o_z	115
Fig. 5.7. Gráfico da velocidade do centro de massa do objeto. Os componentes v_x e v_y são idênticos, pois o objeto seguiu uma trajetória diagonal, cruzando o eixo ótico da câmera. O componente v_z foi sempre igual a 0, pois não houve movimento na direção o_z	115
Fig. 5.8. Gráfico da aceleração do centro de massa do objeto. Os componentes a_x , a_y e a_z são idênticos e iguais a 0,2cm/min ² . Estados medidos, estimados e o erro cometido nas estimativas	116
Fig. 5.9. Gráfico da velocidade angular do objeto de teste. Os componentes w_x , w_y e w_z são idênticos e iguais a 10,0graus/min	116
Fig. 5.10. Gráficos do desvio entre as estimativas dos componentes da posição do centro de massa e os valores desses componentes observados experimentalmente (1º caso).....	117
Fig. 5.11. Gráficos do desvio entre as estimativas dos componentes	

da velocidade do centro de massa e os valores desses componentes observados experimentalmente (1º caso).....	117
Fig. 5.12. Gráficos do desvio entre as estimativas dos componentes da aceleração do centro de massa e os valores desses componentes observados experimentalmente (1º caso).....	118
Fig. 5.13. Gráficos do desvio entre as estimativas da velocidade angular de três pontos característicos (e_{c1} , e_{c2} , e_{c3}) em relação ao eixo OB (1º caso)	118
Fig. 5.14. Gráficos do desvio entre as estimativas dos componentes da posição do centro de massa e os valores desses componentes observados experimentalmente (2º caso).....	120
Fig. 5.15. Gráficos do desvio entre as estimativas dos componentes da velocidade linear do centro de massa e os valores desses componentes observados experimentalmente (2º caso).....	120
Fig. 5.16. Gráficos do desvio entre as estimativas dos componentes da aceleração linear do centro de massa e os valores desses componentes observados experimentalmente (2º caso).....	121
Fig. 5.17. Gráficos do desvio entre as estimativas dos componentes da velocidade angular de três pontos característicos do objeto de teste (e_{c1} , e_{c2} , e_{c3}) e os valores desses componentes observados experimentalmente (2º caso).....	121
Fig. 5.18. Gráficos do desvio entre as estimativas das componentes da posição do centro de massa e os valores desses componentes observados experimentalmente (3º caso).....	123
Fig. 5.19. Gráficos do desvio entre as estimativas dos componentes da velocidade linear do centro de massa e os valores desses componentes observados experimentalmente	

(3º caso).....	123
----------------	-----

Fig. 5.20. Gráficos do desvio entre as estimativas dos componentes da aceleração linear do centro de massa e os valores desses componentes observados experimentalmente (3º caso).....	124
--	-----

Fig. 5.21. Gráficos do desvio entre as estimativas dos componentes da velocidade angular de três pontos característicos do objeto de teste (e_{c1} , e_{c2} , e_{c3}) e os valores desses componentes observados experimentalmente (3º caso).....	124
--	-----

Fig. 5.22. Gráficos do desvio entre as estimativas dos componentes da posição do centro de massa e os valores desses componentes observados experimentalmente (4º caso).....	126
---	-----

Fig. 5.23. Gráficos do desvio entre as estimativas dos componentes da velocidade linear do centro de massa e os valores desses componentes observados experimentalmente (4º caso).....	126
---	-----

Fig. 5.24. Gráficos do desvio entre as estimativas dos componentes da aceleração linear do centro de massa e os valores desses componentes observados experimentalmente (4º caso).....	127
---	-----

Fig. 5.25. Gráficos do desvio entre as estimativas dos componentes da velocidade angular de três pontos característicos do objeto de teste (e_{c1} , e_{c2} , e_{c3}) e os valores desses componentes observados experimentalmente (4º caso).....	127
---	-----

LISTA DE TABELAS

Tabela 5.1. Imagens sintéticas usadas no experimento	94
Tabela 5.2. Tempo de execução dos algoritmos de detecção de cantos e o casamento.....	96
Tabela 5.3. Comparação entre os desempenhos dos diferentes MOPR's empregados na recuperação de informações tridimensionais	104
Tabela 5.4. Trajetória seguida pelo centro de massa do objeto teste.....	106

LISTA DE SÍMBOLOS

- a - Vetor aceleração.
- c_m - Centro de massa do objeto.
- $\hat{c}_m$ - Centro de massa estimado.
- e_x, e_y, e_z - Vetores diretores de S.
- f - Foco da câmera.
- I_n - Matriz identidade.
- n - Ruído gaussiano.
- $q(t)$ - Vetor dos quatêrnios.
- r_j - Pontos característicos do objeto.
- R_c - *Range* associado ao ponto de imagem.
- S_j - Sistema de três eixos das coordenadas fixo no objeto.
- s_j - Ponto pertencente a S.
- $\hat{s}_j$ - Estimativa do ponto s_j .
- t_i - Intervalo de tempo entre medições sucessivas.
- w - Vetor velocidade angular.
- $v(t)$ - Vetor velocidade linear.
- $x(t)$ - Estados a serem estimados.
- $\hat{x}(t)$ - Estimativas do estado $x(t)$.
- x_{op} - Ponto de operação.
- X, Y, Z - Eixos de coordenadas no espaço referencial inercial.
- X_c, Y_c, Z_c - Eixos de coordenadas no referencial da câmera.
- OX, OY, OZ - Eixos orientados nas direções x, y e z respectivamente.
- $f(w)$ - Matriz de transição de $\Omega[w]$.
- I_j - Autovalores associados a uma matriz.

α - Ângulo de rotação do eixo da câmera.

ABSTRACT

This work presents a procedure for the estimation of object motion parameters, namely, position, linear velocity, angular velocity and linear acceleration, using computer vision techniques based on manipulation of a sequence of noisy stereo images. The estimation involves corner detection, 3-D noisy point observation, image acquisition, digital image processing and state estimates. Special attention was given to the first two techniques mentioned above because the acquisition of the object features during the image processing was based on them. The process of recovering three-dimensional points was simulated from the 2-D information, using two methods based on noisy points observation. One of these methods formulates the problem of determination of the three-dimensional points as a linear system equation problem via least square techniques. The other solves the same problem by using the definition of the range of image points and the correspondence that exists between a projected point on one camera and the correspondent conjugate point on the other camera. We still made in this work a study of techniques dedicated to solve the problem of stereo images correspondence. Two approaches are examined. In one of them, the correspondence is made from segmented images. In the other one, the acquired images are used directly.

An experiment using the parameter estimate procedure was performed. The experiment consisted in observing a moving object on a test table and in acquiring a stereo image sequence. The estimates were compared with the motion parameters experimentally observed.

RESUMO

Este trabalho apresenta um procedimento de estimação dos parâmetros de movimento de um corpo rígido, quais sejam, posição, velocidade linear, velocidade angular e aceleração linear, empregando técnicas de visão computacional fundamentadas na manipulação de seqüências de imagens estéreo ruidosas. A estimação envolve o uso de detetores de cantos, observação de pontos ruidosos tridimensionais, aquisição e processamento de imagens e estimação de estados. Atenção especial é dada às duas primeiras técnicas pois é a partir delas que os atributos geométricos do objeto são adquiridos durante o processamento de imagens digitalizadas. O processo de recuperação de pontos tridimensionais foi simulado a partir de informações bidimensionais, usando dois métodos de observação de pontos ruidosos. Um desses métodos reduz o problema de determinação dos pontos tridimensionais à resolução de um sistema de equações lineares pelo método dos mínimos quadrados. O outro utiliza informação do *range* de um ponto imagem e da correspondência entre um ponto projetado em uma câmera e o correspondente ponto conjugado, localizado na outra câmera. Ainda neste trabalho, técnicas aplicáveis na resolução do problema da correspondência entre imagens estéreos foram estudadas. Duas abordagens são consideradas. Em uma delas, o estudo da correspondência é feito a partir da manipulação de imagens segmentadas e na outra utiliza-se as imagens adquiridas.

Um experimento empregando o procedimento de estimação dos parâmetros obtido foi realizado. O experimento consistiu no deslocamento de um corpo rígido em uma mesa e na aquisição de uma seqüência de imagens estéreo. As estimativas obtidas foram comparadas com os valores dos parâmetros de movimento observados experimentalmente.

CAPÍTULO I

INTRODUÇÃO

I.1 Motivação

O interesse pelo tema abordado neste trabalho surgiu da necessidade de se dotar um manipulador robótico com recursos visuais que lhe permita a manipulação de objetos móveis. Esta necessidade deve-se ao fato da crescente mecanização dos meios de produção e exploração verificada nos últimos anos. Percebe-se aí, um forte estímulo para automatizar certas tarefas realizadas pelo ser humano que, com o passar do tempo foram ficando tediosas, dado a frequência de repetição. Além disso, há situações, tais como exploração submarina em águas profundas, coleta de materiais radioativos, exploração de planetas, etc, em que não é possível a presença humana, realizando tais tarefas. Então, se, a um manipulador robótico for acrescentado um dispositivo capaz de interpretar uma cena filmada, detectando seus atributos, quais sejam, aspectos geométricos (volumes, superfícies, curvas, regiões, etc), condição de iluminação da cena e etc, poderão ser executadas tarefas repetitivas em atmosfera nociva à presença humana.

I.2 Conteúdo do trabalho

As estimativas dos parâmetros de um objeto (corpo rígido) em movimento, isto é, sua forma geométrica, velocidade linear e angular, aceleração e posição, feitas a partir da manipulação de quadros de imagens consecutivos, são elementos fundamentais no processamento de imagens e visão computacional. Existe um grande número de trabalhos dedicados ao assunto. Veja, por exemplo, [1] para estimativas feitas usando-se seqüência estéreo de imagens e [2] - [4] para estimativas feitas a partir de uma seqüência monocular de imagens. Os trabalhos que utilizam seqüência monocular apresentam baixa robustez, por usarem um número pequeno de quadros, tipicamente 2 ou 3, conforme em [5] - [7], e são numericamente instáveis quando trabalham com imagens fortemente ruidosas.

Neste trabalho será apresentado um procedimento dedicado à obtenção dos parâmetros de movimento de um corpo rígido, usando uma longa seqüência de quadros de imagem ruidosa. Trabalhos recentes, usando a mesma filosofia podem ser encontrados em [2], [3] e [1]. A referência [1] forneceu a maior parte dos métodos utilizados neste trabalho, pois nessa referência é delineada toda a teoria necessária à manipulação de seqüência de imagens estéreo e estimação de parâmetros. Alguns desses métodos foram modificados ou redirecionados para que os objetivos deste trabalho fossem alcançados. Também neste trabalho, métodos não abordados em [1], quais sejam, métodos de detecção de cantos, métodos de observação de pontos ruidosos e o tratamento da correspondência entre imagens estéreo, foram considerados, com a finalidade de melhorar a robustez da estimação paramétrica. Portanto, será mostrado aqui o detalhamento, a especificação de componentes e a validação, via resultados experimentais, objetivando aplicações em tempo real.

O uso de uma seqüência longa de imagens aumenta a robustez da estimação paramétrica, bem como reduz o problema da instabilidade numérica. O processo de detecção de cantos e o modelo para observação de pontos ruidosos (MOPR) são relevantes no desempenho da abordagem ora analisada e no tratamento da correspondência de imagens estéreo.

O MOPR proposto em [1] determina as posições ocupadas por um objeto em movimento no espaço tridimensional, ou seja, os componentes x , y e z do vetor posição. Isto é feito a partir de informações contidas nos planos-imagem das câmeras, utilizadas como sensores do movimento do objeto. Essas informações são as projeções de duas desses três componentes em relação ao terceiro componente, utilizado como elemento normalizador. A deficiência desse modelo está na determinação (recuperação) do terceiro componente. Neste trabalho será utilizado o modelo de observação de pontos ruidosos descrito em [8], em substituição ao MOPR descrito em [1], que supera esse problema a partir da definição de um parâmetro normalizador (*range*) comum aos três componentes.

Os recentes trabalhos realizados em detecção de cantos, por exemplo, [9] e [10], necessitam da definição matemática de um canto e de máscaras de convolução. Nesses trabalhos há a necessidade de configurar máscaras com números (níveis de cinza) cuja disposição nas máscaras seja semelhante à forma do objeto. Percebe-se, portanto, que no processo de detecção, relatado em [9] e [10], é de fundamental importância o prévio conhecimento da forma geométrica do objeto. Porém, a possibilidade desse processo ser usado em sequência de imagens não foi abordada em [9] e [10]. Neste trabalho serão estudados dois métodos de detecção. Um deles é completamente independente do conhecimento antecipado das características geométricas do objeto e pode ser utilizado na detecção de cantos contidos em sequência de imagens, o outro necessita de informações relativas à forma geométrica do objeto.

Assim, as contribuições deste trabalho em relação ao realizado em [1] são:

- a) Melhoria da qualidade das estimativas dos parâmetros, por uso de um método de obtenção de pontos ruidosos mais eficiente que o proposto por [1] (método apresentado no Capítulo III);
- b) Estudo de dois métodos de detecção de cantos, analisando a influência de cada um deles no procedimento de estimação paramétrica;

- c) Proposta de um ambiente de um dispositivo de visão estéreo para aplicação em tempo real;
- d) Proposta de métodos de verificação da correspondência entre pares de imagens estéreo, definindo um critério de aceitação de imagens correspondentes.

I.3 Ambiente de visão estéreo e componentes

Nesta seção é apresentada a especificação do ambiente de visão estéreo e componentes utilizados neste trabalho. O objeto aqui é a caracterização da situação investigada neste trabalho como um problema de engenharia, evidenciando o contexto onde o trabalho encontra-se inserido. Neste contexto, os valores limites (máximos e mínimos) dos parâmetros a serem estimados são definidos, as condições de realização dos experimentos com imagens sintéticas ou reais são apresentadas, a maneira como as diversas fontes de ruído influencia no modelo integrado de estimação paramétrica e é apresentada a especificação dos sensores (câmeras) empregados nos experimentos. A caracterização do problema é necessária, pois é a partir dela que as magnitudes dos parâmetros estimados poderão ser avaliados à luz do contexto estabelecido.

I.3.1 Condições de realização do experimento com imagens reais

Este item é de extrema importância na elaboração deste trabalho, pois aqui é especificado todo o cenário de realização do experimento com imagem real. Este experimento tem a finalidade de fornecer os dados necessários à implementação de um procedimento de estimação paramétrica apresentado, bem como determinar (através de métodos diretos ou indiretos) os valores de referências das estimativas que serão obtidas, isto é, os valores que permitirão julgar a consistência das estimativas dos parâmetros. A especificação do cenário compreende a descrição física do local de realização do experimento e o planejamento da trajetória de movimentação de objetos.

O cenário ser caracterizado da seguinte forma:

- Um espaço tri-dimensional com referencial inercial, representado fisicamente por um sistema de eixos coordenados O_x , O_y , e O_z , montados em uma mesa de testes;
- Os eixos O_x , O_y possuem 250.0cm de comprimento e eixo O_z 150cm;
- Um objeto de teste (de cor branca) contrastando com a mesa de testes (de cor preta), possibilitando a aquisição de imagens preto e branco (binária);
- Cena iluminada (na área de movimentação do objeto) com luz difusa, já que o contraste mesa-objeto de teste garante silhuetas bem definidas do objeto;
- Movimento do objeto de teste regidos pelas leis de movimento da cinemática;
- Aceleração do movimento constante;
- Velocidade angular também constante.

I.3.2 Condições de realização do experimento com imagens sintéticas

O experimento com imagens sintéticas tem por finalidade principal testar as diferentes técnicas de segmentação de imagens. Este experimento será realizado com imagens na seguinte condição:

- Imagens sintéticas de figuras geométricas semelhantes às projeções do objeto de teste a ser usado no experimento com imagem real.

I.3.3 Interferência das diferentes fontes de ruído no experimento com imagens reais

A interferência das diferentes fontes de ruído no experimento com imagens reais será representada por um modelo fundamentado nas estatísticas (média e variância) do ruído associado ao processo de digitalização de imagens no local de realização do experimento.

I.3.4 Valores máximos e mínimos dos parâmetros que descrevem o movimento do objeto

Neste trabalho os parâmetros que descrevem os movimentos são: velocidade linear, velocidade angular, aceleração linear e posição linear.

Os valores dos parâmetros podem ser determinados experimentalmente por métodos direto e indireto. No método direto, os parâmetros são determinados por simples medição de suas magnitudes durante a realização do experimento. É o que acontece na determinação da posição linear. No método indireto, medições também são feitas, mas a determinação efetiva dos valores dos parâmetros somente se verifica após a utilização dessas medições em fórmulas ou leis da cinemática. O método indireto é aplicado na determinação dos três primeiros parâmetros enumerados anteriormente.

A especificação dos valores máximos e mínimos dos parâmetros está intimamente ligada com as dimensões físicas do arranjo experimental utilizado (item 1.3.1), com a base de tempo considerada nos movimentos contidos nas cenas filmadas e com as imposições feitas a esses movimentos no item 1.3.1. Sendo assim, tem-se as seguintes especificações:

- Base de tempo: 2min;
- Valores máximos e mínimos da velocidade linear: 7.6 ± 0.5 e 0.4 ± 0.5 cm/min, respectivamente;
- Valor da velocidade angular: 10.0 ± 0.5 graus/min;
- Valor da aceleração linear: 0.2 cm/min²;

- Valores máximos e mínimos da posição linear: 224.40 ± 0.05 e 10.00 ± 0.05 cm/min, respectivamente.

I.3.5 Especificação dos elementos sensores (câmeras) utilizados no experimento com imagens reais

A especificação das câmeras utilizadas neste trabalho será feita a partir das características das imagens adquiridas e da resolução necessária ao processo de digitalização das mesmas.

I.4 Organização do trabalho

O trabalho está dividido em seis capítulos como segue.

O Capítulo II contém um histórico dos trabalhos que se dedicaram à obtenção de estimativas de parâmetros a partir de uma seqüência de imagens ruidosa. O capítulo está dividido em quatro seções. A primeira seção comenta sobre a utilização de pequena seqüência monocular de imagem. Duas abordagens são descritas: a) algoritmos que usam o fluxo óptico de campo e b) algoritmos fundamentados em grandezas geométricas. Qualquer das duas abordagens possui baixa robustez diante da presença de ruído e os valores determinados podem não ser os únicos. A segunda seção dedica-se à discussão das estimativas feitas a partir do uso de uma longa seqüência monocular de imagem. São, então, apresentados alguns métodos que vêm sendo desenvolvidos por vários pesquisadores nos últimos anos, entre eles, os métodos que utilizam os quartêrnios para resolver o problema da propagação do movimento de rotação. A terceira seção relata sobre o emprego de visão em malha fechada na estimação do movimento de objetos. E, finalmente, a quarta seção faz comentários acerca de manipuladores robóticos que utilizam sistema de visão estéreo descalibrado.

No Capítulo III é feito um estudo sobre a utilização de seqüência de imagem estéreo ruidosa na estimação dos parâmetros de um objeto em movimento. O capítulo será dividido em 3 seções. Na primeira são feitas considerações iniciais acerca dos trabalhos recentes dedicados à estimativa de parâmetros a partir de seqüência de

imagem estéreo. Na segunda seção tem-se o início do equacionamento dos modelos necessários ao estimador recursivo – Filtro de Kalman Estendido com Iterações (IEKF). São considerados dois modelos matemáticos para os algoritmos de observação de pontos ruidosos (MOPR).

O Capítulo IV descreve o Sistema de Aquisição e Digitalização de Imagem (SADI) e o estudo de dois métodos de detecção de cantos. Neste capítulo, um canto é definido a partir de modelagem matemática. Serão apresentadas as análises matemáticas e expressões analíticas para o cálculo do canto. Será também mostrado neste capítulo que as máscaras utilizadas no cálculo do gradiente não necessitam do prévio conhecimento da forma geométrica do objeto. O Capítulo IV ainda descreve os métodos de avaliação da correspondência entre imagens estéreo. Duas abordagens são consideradas. Em uma delas, a correspondência é determinada a partir da manipulação das imagens originais (digitalizadas) ao passo que na outra abordagem, exige-se a prévia segmentação das imagens antes da determinação da correspondência. Ainda no Capítulo IV, é feita a modelagem das fontes de ruído presentes no processo de digitalização de imagens.

O Capítulo V mostra os resultados obtidos nas simulações com sequência de imagens estéreo simulada e real, bem como os resultados encontrados com os diferentes métodos de cálculo da correspondência abordados no Capítulo IV. O capítulo é dividido em três seções. A primeira tem um caráter introdutório, fazendo considerações iniciais. A segunda seção mostra os resultados experimentais obtidos com os métodos de detecção de cantos considerados neste trabalho, estabelecendo um quadro comparativo entre ambos. Na terceira seção, examina-se os resultados obtidos com os experimentos relativos às técnicas de verificação da correspondência entre um par de imagens adquiridas por câmeras distintas. Nesta seção, considera-se tanto as imagens originais, quanto as imagens segmentadas, contendo exclusivamente propriedades geométricas da cena filmada: os chamados atributos. A quarta seção exhibe os resultados experimentais com as diferentes abordagens dos MOPR's considerados. A quinta seção mostra os resultados alcançados com simulações feitas com sequência estéreo de imagens reais. São exibidos gráficos dos valores dos estados determinados experimentalmente (medidos), gráficos contendo os desvios existente entre as

estimativas dos estados e os valores medidos (quadro a quadro) ao longo das simulações.

Finalmente o Capítulo VI apresenta a conclusão e sugestões para trabalhos futuros.

CAPÍTULO II

PRIMEIROS TRABALHOS EM VISÃO COMPUTACIONAL DESTINADOS À ESTIMAÇÃO DOS MOVIMENTOS CONTIDOS EM UMA CENA

II.1 Utilização de pequena sequência de imagens monocular

Vários trabalhos em visão computacional têm sido desenvolvido com o objetivo de determinar as estimativas dos parâmetros tridimensionais que definem o movimento e a forma geométrica de um objeto. Os elementos que possibilitam essas estimativas estão contidos em seqüências de imagens monoculares. Basicamente, são duas essas técnicas. Uma delas tem seus fundamentos no fluxo ótico de campo, enquanto que a outra consiste no uso de grandezas geométricas, tais como ponto, aresta e bordas.

II.1.1 Algoritmos que utilizam o fluxo ótico de campo

Nas técnicas que usam o fluxo ótico de campo, o procedimento mais comum requer dois passos. Inicialmente faz-se o cálculo do fluxo ótico de campo e a seguir interpreta-se o campo calculado. O fluxo ótico pode ser produzido pelo movimento de um sensor através de um ambiente ou por movimentos independentes em um campo visual. A interpretação da informação originária desse fluxo consiste no resgate dos parâmetros móveis em cada movimento do objeto e a determinação de sua forma geométrica. Consequentemente, a presença de ruído durante este processo tende a dificultar o processamento do fluxo ótico.

O fluxo ótico tanto pode representar campo de velocidade, quanto deslocamento de vetores que representam o movimento dos elementos contidos nas imagens (quadro

após quadro). Quanto mais lenta for a evolução dessas imagens, mais preciso será o cálculo do fluxo ótico de campo e, conseqüentemente, melhor será a interpretação das informações contidas no campo visual. A presença de ruídos interfere no reconhecimento do fluxo ótico, tornando-o praticamente incorreto. Algoritmos que fazem uso desse princípio tendem a falhar nestas circunstâncias. Existe, no entanto, técnicas, relatadas em [11], que, por utilizarem todas as informações contidas no campo visual, podem apresentar robustez. Contudo, a escolha de métodos de otimização inadequados acabam por limitar o desempenho dessas técnicas.

Muitos dos algoritmos que fazem uso do fluxo ótico têm sido objeto de estudos. Poucos são bem sucedidos na presença de níveis altos de ruídos. Trabalhos como [4] - [6], [12] e [13] fazem o uso de equações não-lineares nas quais os parâmetros relacionados com o movimento são desconhecidos. A resolução dessas equações é feita geralmente por métodos iterativos, cuja convergência depende das condições iniciais. As técnicas exploradas em [5], [6] e [14] fazem considerações sobre a sensibilidade desses algoritmos ao ruído. Examinando os procedimentos dados, fica claro que os resultados podem ser sensivelmente melhorados aumentando-se o número de fluxo ótico de campo, o que implica aumentar o número de medidas e ampliação da região que contém o fluxo ótico, contanto que não sejam considerados movimentos independentes nas imagens. Assim, admitindo que os objetos em uma cena atendem a esta última condição, as técnicas anteriormente citadas dão bons resultados quando são aplicadas em toda a imagem (todo campo visual).

Os métodos desenvolvidos em [15] e [16] definem técnicas de resolução de equações de sistemas de equações não-lineares. Problemas como a unicidade das soluções são abordados por esses trabalhos, onde a principal dificuldade é a sensibilidade ao ruído.

As técnicas abordadas em [17] e [7] são recomendadas para otimização dos algoritmos mencionadas acima. Nessas técnicas o critério de mínimos quadrados é utilizado para minimizar a diferença entre as medidas feitas e os valores calculados, isto é, parâmetros móveis e estrutura do objeto. Em [17] e [7] os sistemas de equações não-lineares são calculados por métodos numéricos. Considerando-se apenas o movimento de translação, todo o fluxo ótico é orientado para fora do plano de

ou para um ponto no plano. Este ponto é o foco de expansão (FOE) e é o que define a direção da translação. As técnicas relacionadas a seguir utilizam o FOE.

De acordo com as considerações feitas acima, ao sensor só é permitido o movimento de translação. O ambiente é admitido somente possuir superfícies planas em qualquer das orientações definidas anteriormente. Então, o algoritmo consiste nas relações existentes entre o FOE e as distâncias das superfícies na cena. Estudos sobre robustez no uso do FOE são feitos nos trabalhos desenvolvidos em [18] e [2]. Nesses trabalhos, não há restrições quanto à forma das superfícies, mas os mesmos só são usados para movimento de translação. Outra característica marcante desses trabalhos é a correspondência entre o erro de medida amostrado e a posição do FOE. Com isto consegue-se determinar a exata localização do valor mínimo. Trata-se, portanto, de um algoritmo de minimização. Existem técnicas que dependem do campo de velocidade dos componentes de translação e rotação. Em [2] é descrita uma delas. Considerando componentes de rotação hipotéticas, essas técnicas podem calcular o FOE do correspondente movimento de translação e o erro de medida associado. Assim, a função erro de rotação tridimensional é obtida e a solução pode ser determinada por minimização dessa função.

Existe uma quantidade relativamente grande de trabalhos baseados na análise local do fluxo de campo. Em [10] é possível a presença de movimentos independentes, condição bastante crítica nos trabalhos apresentados até então, e pequenas explorações de conjunto de fluxos adjacentes. Se existe uma única interpretação nesse conjunto de fluxos, então ela é considerada correta e os vetores que representam esse fluxo são agrupados. Entretanto, por ser esta análise essencialmente local, existe muita sensibilidade aos erros no fluxo de campo.

As técnicas desenvolvidas em [18], [2] e [3], definem equações para o cálculo de parâmetros móveis e forma geométrica usando análise local. São calculados por esses métodos a estrutura e a rotação em um dado ponto do ambiente de fluxo ótico, para uma dada cena filmada. Também são calculadas as derivadas primeiras, isto é, velocidades dos correspondentes pontos na imagem. Se a cena consiste de vários objetos em movimento, então o método é aplicado em cada um deles. O cálculo da segunda derivada é impreciso na presença de ruído, segundo a referência [19]. Das três abordagens referidas acima, somente a desenvolvida em [3] consegue bons resultados

na determinação dos parâmetros móveis e estrutura da superfície em ambiente com ruído.

Tudo que foi exposto nos parágrafos anteriores ilustra as dificuldades e restrições existentes na estimação de parâmetros móveis e da forma geométrica de objetos (estrutura), quando a técnica considerada é fundamentada no fluxo ótico de campo. A maioria delas são bastantes sensíveis ao ruído e de difícil aplicação em objetos com movimentos independentes. Em alguns casos ocorre ambigüidade na interpretação de parâmetro tridimensional, originário de fluxo de campo ruidoso. Contudo, é possível conseguir bons resultados quando são empregadas todas as informações disponíveis e a utilização dos critérios de otimização adequados, como por exemplo, a transformada de Hough generalizada [20] e [21].

II.1.2 Algoritmos que fazem uso de grandezas geométricas

Os dois problemas fundamentais das técnicas que usam grandezas geométricas são: a) encontrar os pontos de canto em um par de imagens consecutivas, e b) resolver um conjunto de equações lineares para determinar os parâmetros móveis na cena. Somente são permitidos movimentos de translação. Rotação ou outros tipos de movimentos não são considerados. Então, somente a velocidade linear ou posição do objeto pode ser determinada por estes métodos.

A primeira dessas técnicas é apresentada em [22]. Nessa técnica, os parâmetros móveis tridimensionais de uma cena podem ser determinados pelo cálculo de uma matriz 3×3 que contém os chamados oito parâmetros puros (VDS). O problema é determinar as coordenadas bidimensionais espaço-imagem de um ponto na cena. Para tanto a imagem é projetada no plano da câmera, base geométrica do problema, nos diferentes intervalos de tempo. Somente dois quadros, (t_1, t_2) , são usados nessa técnica. Assim, tem-se (x, y) correspondendo ao tempo t_1 e (x', y') relacionado com o t_2 . Para se determinar (x', y') oito coeficientes são necessários - oito parâmetros puros. Em [23] eles são calculados via resolução de um sistema de equações lineares. Com esse procedimento, o movimento de um ponto objeto-espaco é estimado. Se houver mais de um ponto-espaco na cena, então o movimento deles pode ser determinados por aplicações sucessivas do método.

A segunda técnica, descrita em [23], utiliza um algoritmo capaz de determinar a posição das bordas e os cantos de um objeto. Esse algoritmo é baseado nos estudos feitos em [12], manipula pixels de imagem originária das câmeras de baixa resolução para estimar os parâmetros móveis tridimensionais. Nessa técnica a precisão da medida das coordenadas de um par de pontos tem grande significado na resolução das equações de estimativa de movimento. Contudo, o uso das câmeras de baixa resolução não é recomendado, a menos que os pixels pudessem ser determinados com altíssima precisão. Os trabalhos desenvolvidos em [12] recomenda o uso de câmeras com melhor resolução, mas deixa claro que esta opção dificulta as estimativas dos parâmetros, por aumentar o tempo de processamento.

Como se pode observar, a maioria das técnicas que utilizam grandezas geométricas apresentam limites quanto aos movimentos, sensibilidade ao ruído e baixa robustez. Os resultados obtidos não são muito precisos, ficando a eficiência muito dependente da resolução da câmera utilizada. A rotação não pode ser muito grande, tipicamente ângulos próximos de 50° . Pode-se, contudo obter bons resultados se o movimento de translação nas direções x e y (Dx e Dy) forem pequenos ou se houver grandes deslocamentos na direção de z .

Conforme se pode concluir do exposto acima, uma quantidade razoável de trabalhos foram desenvolvidos baseados em uma pequena seqüência de imagens, tipicamente dois ou três quadros de imagem ruidosa. Muitos dos métodos existentes consideram que o movimento de um objeto é, a partir de uma seqüência de imagens, a leitura da posição do objeto na primeira imagem e a sua localização na segunda. Essa transformação pode ser representada por uma translação seguida de rotação em relação a um eixo. Já que o tempo entre uma imagem e outra pode ser medido, o resultado é a sensação de movimento (tipicamente translação). Contudo, esses métodos, por usarem pouca quantidade de quadros, apresentam baixa robustez, falta de estabilidade numérica e geralmente apresentam problemas de unicidade dos valores determinados. Usando as técnicas desenvolvidas em [20] - [22], por exemplo, e o uso de uma longa seqüência de quadros, pode-se melhorar o desempenho dessas técnicas.

II.2 Utilização de longa seqüência monocular

Na seção anterior foram examinados os primeiros trabalhos feitos em visão computacional com o objetivo de estimar os parâmetros que descrevem o movimento de um objeto e a sua forma geométrica (estrutura). Naquela seção ficou claro que problemas tais como baixa robustez na presença de ruído, falta de estabilidade numérica, etc, são os que limitam a eficiência dos algoritmos mencionados. No final daquela seção foi sugerido que a utilização de técnicas que fazem uso de um maior número de quadros melhoraria as estimativas feita por aqueles algoritmos. Então, passa-se a analisá-las na próxima seção.

II.2.1 Estimativa de uma seqüência de imagens monocular

A situação abordada por essa técnica objetiva o estabelecimento de um algoritmo recursivo capaz de, a partir de uma seqüência de imagens monoculares, estimar a forma geométrica e a cinética de movimentação de um objeto contido na seqüência de imagens. A estrutura do mesmo não varia com o tempo, condição que favorece a utilização dos argumentos da dinâmica dos corpos rígidos e a suavidade dos movimentos permite possíveis linearizações de trajetória. O movimento de translação é retilíneo e uniformemente variante no tempo, enquanto que o movimento de rotação, sem limite de ângulo, é estudado com base no conceito de quatêrnios. Equações dinâmicas que contêm informações de movimento e da estrutura são deduzidas para a propagação no tempo dos estados estimados. Trata-se, portanto, de uma técnica recursiva.

Pontos de uma seqüência de imagem são extraídos quadro após quadro. Eles são usados na estimativa da forma geométrica do objeto. São possíveis situações com imagem livre de ruído e com imagem ruidosa. A primeira é hipotética, difícil de ocorrer na prática. A segunda corresponde à situação usual, na qual o ruído é originário do processo de medição ou da condição de iluminação da cena. Portanto, níveis de ruído são permitidos nessa técnica e a forma geométrica do objeto vai sendo estimada com o tempo, quando uma longa e arbitrária quantidade de quadros é utilizada. Os resultados das estimativas da forma geométrica são, então, usados no

estimador recursivo de parâmetros, geralmente um (IEKF), para obter as estimativas da cinética do objeto.

A seguir são relacionados alguns trabalhos que fazem o uso de longas seqüências de quadros.

Em [17] são desenvolvidos sistemas de equações não-lineares que contêm informações retiradas das imagens. Essas informações são as coordenadas do objeto e a posição da câmera. Técnicas numéricas são usadas para resolver essas equações e o sucesso do método depende da boa escolha das condições iniciais. As técnicas abordadas em [23] - [24] analisam o problema da estimação usando métodos teóricos e numéricos.

As técnicas discutidas em [23] exploram as propriedades da projeção central de imagens e a elaboração de modelos. Algumas informações acerca da estrutura do objeto são usadas nas simulações dos dados. Métodos recursivos foram utilizados na manipulação de imagens envolvendo somente translação e estrutura desconhecida. Em [4] o caso geral de objeto em movimento com estrutura desconhecida foi explorado.

A tarefa essencial da estimação monocular é desenvolver modelos que representem o movimento do objeto, a correspondente observação no tempo do mesmo, para que seja possível a formulação das estimativas de estado recursiva. O uso de representação por série de Taylor permite a utilização de quantas derivadas forem necessárias à modelagem. São permitidos, de acordo com o exposto acima, modelos de translação e rotação. As equações dos modelos de translação são lineares, pois esse movimento é uniformemente variante no tempo. Contudo, a rotação é mais complexa, suas equações são quase sempre não-lineares e para resolver o problema de sua propagação são usados os quatêrnios.

Na situação física do problema, um objeto está no sistema de coordenadas $OXYZ$, onde O é a origem. As distâncias entre os pontos (X,Y,Z) definem as suas dimensões. O movimento de translação é para longe do sistema $OXYZ$ (progressivo) ou de aproximação (retrógrado), enquanto que a rotação pode ser ao redor de qualquer eixo que passe na origem O . O que interessa no objeto são os pontos característicos, isto é, os pontos de cantos. Esses pontos formam um conjunto constante no tempo, pois o objeto é um corpo rígido. Uma câmera, suposta inercial, colocada na origem

observa os movimentos do objeto. Então, os pontos característicos são projetados no seu plano imagem.

Finalmente, considerando todos os quadros de imagens ruidosas e um número finito de pontos característicos em cada quadro, são estimados os parâmetros desconhecidos (forma geométrica e parâmetros que representam o movimento). A equação de estado é baseada nas equações da dinâmica de corpos rígidos e é usada para descrever a evolução no tempo dos estados. As características não-lineares dessa equação requer anterior linearização, antes do uso do IEKF [1]. O desempenho do IEKF depende das condições iniciais dos estados estimados, da precisão dos algoritmos de linearização e da intensidade do ruído presente.

O uso de uma longa seqüência de imagem ruidosa na determinação das estimativas de parâmetros apresenta numerosas vantagens sobre os outros métodos mencionados nas seções anteriores. Trata-se de uma técnica recursiva que pode ser aplicada em tempo real. De acordo com [4], ela pode ser usada na observação de veículos, visão de braços mecânicos, visão de teleguiados, etc.

Outras vantagens são a liberdade na escolha do número de pontos-imagens, independência dos algoritmos do número de imagens na seqüência e a relativa robustez em relação à presença de ruído. O principal objetivo dessa técnica é estabelecer a correspondência entre os pontos-imagens originário da câmera e a correspondente localização da estrutura do objeto no espaço tridimensional. Resolver esse problema, [4], requer um número elevado de quadros, mas esse procedimento acaba por aumentar o tempo de convergência dos algoritmos e, conseqüentemente, dificulta a aplicação dessa técnica em tempo real. Entretanto, o tempo requerido para a manipulação dos pontos-imagens (tempo necessário para os algoritmos convergirem) de um quadro a outro pode ser extremamente reduzido pelo uso do IEKF [1]. Mas esse filtro pode divergir por má escolha das condições iniciais ou por baixa robustez dos métodos numéricos.

As características dos métodos computacionais que fazem uso de uma única câmera, visão monocular, já foram examinadas neste trabalho. A maioria deles apresentam baixa robustez em relação aos dados utilizados. É crítico o problema da unicidade dos resultados, pois geralmente esses métodos tem que resolver sistemas de equação não-lineares para determinar a estrutura dos objetos. Existem, entretanto,

técnicas que dão bons resultados na obtenção das estimativas da cinética e da forma geométrica do objeto existente em uma cena (uso de sequência monocular de imagem ruidosa) [4]. Seus algoritmos são robustos e existe a possibilidade de serem usados em tempo real, contanto que um IEKF seja utilizado. Mas, mesmo esse sofisticado filtro pode divergir em certas condições. Contudo, a referência [4] sugere que se forem usadas seqüências distintas de duas câmeras (seqüência binocular estéreo), em muito seria aumentada a eficiência do IEKF, desde que seja resolvido o problema da não-linearidade existente nas equações dinâmicas. O filtro convergiria rapidamente e as condições iniciais não seriam mais um problema.

II.3 Estimação do movimento de objetos usando a teoria de sistemas de controle em malha fechada

A interpretação geométrica e cinemática de uma região do espaço tridimensional são os requisitos básicos para um ser vivo se locomover e reconhecer objetos nessa região. Embora as primeiras formulações de problemas que relacionam o movimento de seres ou objetos com dispositivos de visão tenham sido iniciadas no começo deste século [24], somente nos últimos anos as ferramentas da teoria de controle e estimação de sistemas, particularmente a teoria dos sistemas em malha fechada, [25] – [27] têm sido empregada em aplicações tradicionalmente difíceis de serem realizadas. Tem-se como exemplo dessas aplicações: servo-manipuladores, rastreamento de objetos e visão ativa [28]. Visão em malha fechada pode ser configurada a partir de duas abordagens. Em uma delas, a malha fechada é obtida, distribuindo-se um grande número de sensores no sistema de visão. Na outra abordagem, o sistema em malha fechada é representado pelo processamento dos sinais originários desses sensores, objetivando a análise e controle das magnitudes dos mesmos, caracterizando uma nova classe de sistemas dinâmicos [29].

II.4 Calibração automática de sistema de visão estéreo

Neste trabalho, os componentes utilizados com sensores (câmeras) são calibrados previamente, antes de serem empregados em experimentos com imagens reais, isto é, as câmeras apresentam calibração fixa. Contudo, há situações em que os

sensores são calibrados durante a filmagem de um cena. Trata-se de uma abordagem em que os sensores são calibrados automaticamente. Nessa abordagem o processo de calibração de um dispositivo de visão, constituído de duas câmeras não calibradas, é feito a partir do controle da posição de um manipulador robótico em uma região de trabalho predefinida [30]. O mesmo problema também foi considerado em [31], onde o sistema constituído do manipulador e das câmeras foi modelado como se fosse um sistema linear em que os coeficientes desconhecidos da matriz associada variam lentamente no tempo. Por conta dessa condição, os resultados obtidos em [31] são válidos somente se o manipulador executa movimentos lentos.

Como uma maneira de lidar com essa limitação, em [30] é realizado um estudo, considerando-se, na modelagem das câmeras, as reais características não-lineares existentes. Por causa dessas características, o desenvolvimento de um algoritmo capaz de fazer a calibração *em tempo real*, usando os dados reais originários das câmeras (imagens) no tempo de execução da trajetória do manipulador, não é uma tarefa das mais fáceis. O manipulador pode, nessa condição, perder a capacidade de rastrear. A falta dessa capacidade é resolvida a partir da manipulação de um grande número, mas finito, de possíveis modelos do sistema manipulador-câmera.

Mesmo com um número finito de candidatos a modelos, o problema da auto calibração ainda persiste, pois o número desses modelos candidatos é elevado e cada um deles apresenta não-linearidade. A abordagem adotada em [31] é fundamentada em estimação de parâmetros supervisionada, associada a um controle de equivalência a certeza [32]. Também nessa abordagem é usada a “consistência epipolar” [33] com o objetivo de evitar que o supervisor consuma um tempo elevado com modelos candidatos pouco representativos do sistema câmera-manipulador.

CAPÍTULO III

UTILIZAÇÃO DE SEQUÊNCIA DE IMAGENS ESTÉREO RUIDOSA

III.1 Considerações iniciais

Os primeiros trabalhos em visão computacional trataram a análise de dados visuais, feita a partir de imagem estéreo estática e sequência monocular de imagens, de maneira distinta para estimação da estrutura de objetos. Cada uma dessas abordagens tinha que enfrentar obstáculos peculiares. Enquanto a análise feita por imagem estéreo estática tinha que resolver problemas como a ocorrência de bordas oclusas, a fundamentada em sequências monoculares de imagem examinava a solução de equações não-lineares e o estabelecimento de um vínculo entre a interpretação tridimensional e um fator de escala arbitrário [4]. Mas, a partir da análise do processo da visão humana (processo que manipula canais de visão distintos), passou a existir forte evidência de que os dois métodos poderiam ser interpretados juntos. Em [4], resultados que usam esses argumentos podem ser encontrados. Neste trabalho é utilizado um modelo que é a fusão dessas duas abordagens.

Um outro trabalho dedicado ao estudo desse modelo é o realizado em [1]. Nesse trabalho foi analisada a possibilidade de obtenção das estimativas da estrutura de objetos usando análise estéreo estática e manipulação de sequências monoculares de imagem. Nas técnicas utilizadas na referência [34] ficou claro que é relevante o uso de grandes quantidades de medidas instantâneas ao longo do tempo. Os estudos feitos em [35] consideram uma técnica que usa o fluxo relativo de imagens de duas câmeras que têm movimento entre si conhecidos. A grande contribuição desses trabalhos à visão

computacional foi, sem dúvida, a possibilidade do uso de fluxo de imagens binoculares. Tem-se assim, o surgimento de uma técnica que manipula seqüência de imagens binoculares.

As vantagens da técnica que manipula seqüência de imagens binoculares sobre as técnicas de imagens estéreo estáticas e seqüência de imagens monoculares são:

- a) Detecção da profundidade dos objetos; e
- b) Capacidade de refinar as estimativas por fazer uso de duas câmeras cuja posição entre si é conhecida.

Também, a manipulação de seqüência de imagem binocular tem a vantagem de permitir o cálculo da estrutura do objeto em movimento, sem o problema de ambigüidade nos resultados. Isto acontece por não haver a necessidade de resolver equações não-lineares. Resta saber de que maneira esta técnica pode ser aplicada como um algoritmo de visão capaz de identificar os parâmetros que descrevem o movimento e a estrutura de um objeto no espaço tridimensional.

Neste capítulo será apresentado o estudo de métodos e técnicas utilizados em visão computacional (propostos em [1]) que representam um aprimoramento dos métodos vistos no Capítulo II, cujas formulações matemáticas podem ser encontradas nas referências [4], [1] e [35], que fazem uso de modelos robustos e de uma longa seqüência ruidosa de imagens estéreo para observação do movimento de um objeto. Por causa da sua natureza recursiva, os mesmos necessitam de um grande número de quadros e de alta robustez numérica para que seja eficiente a estimação dos parâmetros relacionados com o movimento e a estrutura do objeto.

Na caracterização dos referidos modelos, são considerados os fundamentos da dinâmica de corpos rígidos, projeções geométricas e triangularização de imagens. O objeto em questão encontra-se em um espaço referencial inercial I , podendo apresentar os movimentos de rotação e translação. Durante a evolução dos quadros, o mesmo é observado por duas câmeras que apresentam posição entre si conhecidas. O movimento de translação tem dependência quadrática com o tempo. O movimento de rotação é propagado usando os quatêrnios.

As informações contidas nas imagens estéreo permitem a separação dos parâmetros que definem a estrutura do objeto dos relacionados com o movimento. Esses últimos são estimados pelo IEKF, após anterior definição dos modelos pertinentes.

III.2 Modelos considerados

Serão analisados os quatro modelos utilizados na implementação do algoritmo estimador de parâmetros. Esses são:

- a) Modelo para projeção ortogonal nos planos-imagens das câmeras;
- b) Modelos para observação ruidosa;
- c) Modelo da estrutura do objeto;
- d) Modelos de movimentos.

Após a descrição desses modelos, os mesmos são configurados em conjunto com um filtro de Kalman [38] e reunidos com outros modelos descritos no Capítulo IV. O resultado dessa união compõe o procedimento de estimação apresentado. O objetivo é estimar os parâmetros (de movimentos e estrutura) do objeto enquanto o mesmo movimentava-se numa mesa de teste.

A seção seguinte apresenta a análise detalhada dos modelos citados anteriormente. Os fundamentos físicos, matemáticos ou computacionais essenciais à consistência desses modelos são delineados nas referências [4], [1], [36] - [38].

III.2.1 Modelo para projeção ortogonal nos planos-imagens das câmeras

O objetivo desta seção é analisar as equações de recorrência dos pontos que representam a estrutura do objeto.

Informações tridimensionais de uma cena podem ser obtidas a partir de seqüências de imagem estéreo. Isto é possível porque a cena é projetada nos planos das câmeras com diferentes localizações. Portanto, mesmo que as projeções da imagem

implique em perda de uma das dimensões, é possível recuperar a informação tridimensional por combinação das distintas projeções nos planos das câmeras.

Considere o caso simples de um ponto $p(x, y, z)$ sendo observado por duas câmeras distintas. Se a sua imagem pode ser projetada nos planos de cada uma delas, com a posição entre elas conhecida, então a coordenada tridimensional de p pode ser calculada por simples triangulação. Para o caso de um corpo rígido que possui infinitos pontos p , deve ser definida uma maneira de calculá-los. Isto é feito em [1].

Sejam duas câmeras, CAM_L e CAM_R , colocadas originalmente na origem do sistema inercial I com linhas de visadas alinhadas com a direção z de I . Desloca-se CAM_L ao longo da direção x de I (o deslocamento poderia ser na direção y) por uma quantidade l_c . Após esse movimento de translação, a câmera CAM_L sofre uma rotação em volta da direção y por um ângulo θ , de modo que a nova linha de visada z' e a original, conforme mostrado a seguir, formem esse ângulo.

Fig. . 3.1 - Observação de um corpo rígido em movimento por duas câmeras.

Conseqüentemente, tem-se:

$$(xl_c, yl_c, zl_c)^T = \begin{bmatrix} \cos\theta_l & 0 & -\sin\theta_l \\ 0 & 1 & 0 \\ \sin\theta_l & 0 & \cos\theta_l \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} l_c \cos\theta_l \\ 0 \\ l_c \sin\theta_l \end{bmatrix} \quad (3.1)$$

onde $(x, y, z)^T$ representa as coordenadas dos pontos no espaço tridimensional inercial e $(x_l, y_l, z_l)^T$ são as coordenadas do ponto P no sistema fixo das câmeras. Usando os mesmos argumentos considerados para a câmera esquerda, tem-se a seguinte expressão para a câmera direita:

$$(xl_r, yl_r, zl_r)^T = \begin{bmatrix} \cos\theta_r & 0 & -\sin\theta_r \\ 0 & 1 & 0 \\ \sin\theta_r & 0 & \cos\theta_r \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} l_c \cos\theta_r \\ 0 \\ l_c \sin\theta_r \end{bmatrix} \quad (3.2)$$

Na expressão (3.1) ou (3.2) pode ser observada a matriz 3x3 com senos e cossenos zeros e uns. Esta é a matriz de rotação. Ela pode ser definida como uma transformação que opera um vetor posição no espaço tridimensional Euclidiano, mapeando suas coordenadas para outras que fazem ângulos θ com a posição original. Esta operação é denominada rotação e para realizá-la é necessário um eixo de referência. Este é o eixo de rotação. Após a rotação obtém-se um novo vetor com posição angular diferente da original.

De onde vem as matrizes 3x3 em (3.1) ou (3.2)? Existem outras configurações? Para responder a essas perguntas, considere o diagrama mostrado na figura 3.2. Nela vê-se dois sistemas de eixos orientados. Um é o O_{XYZ} com seus O_x , O_y e O_z . O outro é O_{UVW} e os respectivos eixos O_U , O_V e O_W . Também pode ser notado que O_{XYZ} é alinhado com O_{UVW} e a origem dos dois são coincidentes no ponto O .

Fig. 3.2 - Localização de ponto em dois sistemas coordenados paralelos.

O sistema coordenado O_{XYZ} está fixo no sistema tridimensional inercial e é tomado como sistema referência. As coordenadas de O_{UVW} podem girar em relação a O_{XYZ} . Consideremos (i_x, j_x, k_x) e (I_u, J_v, K_w) como sendo vetores ortonormais que geram os espaços O_{XYZ} e O_{UVW} respectivamente. O ponto P na figura pode ser representado por suas coordenadas em ambos os sistemas. Então, o ponto P pode ser representado por suas coordenadas nos sistemas O_{UVW} e O_{XYZ} , respectivamente, como,

$$P_{uvw} = \begin{bmatrix} P_u \\ P_v \\ P_w \end{bmatrix}$$

$$P_{xyz} = \begin{bmatrix} P_x \\ P_y \\ P_z \end{bmatrix} \quad (3.3)$$

onde P_{uvw} e P_{xyz} representam o mesmo ponto P em diferentes sistemas coordenados.

A partir desses argumentos, o objetivo é determinar a matriz R_r , tal que seja possível escrever uma transformação que passe as coordenadas de P_{uvw} para as coordenadas expressas em relação a O_{xyz} . Esta transformação é:

$$P_{xyz} = R_r P_{uvw} \quad (3.4)$$

onde o produto de R_r por P_{uvw} significa que o ponto P_{uvw} sofreu rotação, bem como o sistema O_{uvw} . Voltando a (3.3), podemos escrever P_{uvw} como:

$$P_{uvw} = p_{uiu} + p_{yiv} + p_{wk_w} \quad (3.5)$$

e P_x , P_y e P_z podem ser calculados pela projeção de P nos respectivos eixos. Então, usando a definição de produto escalar e $P = i_u P_u + j_v P_v + k_w P_w$, o que é possível, pois o sistema O_{uvw} contém o ponto P , tem-se:

$$\begin{aligned} P_x &= i_x P = i_x i_u P_u + i_x j_v P_v + i_x k_w P_w \\ P_y &= j_y P = j_y i_u P_u + j_y j_v P_v + j_y k_w P_w \\ P_z &= k_z P = k_z i_u P_u + k_z j_v P_v + k_z k_w P_w \end{aligned} \quad (3.6)$$

e na forma matricial tem-se,

$$\begin{bmatrix} P_x \\ P_y \\ P_z \end{bmatrix} = \begin{bmatrix} i_x i_u & i_x j_v & i_x k_w \\ j_y i_u & j_y j_v & j_y i_w \\ k_z i_u & k_z j_v & k_z i_w \end{bmatrix} \begin{bmatrix} P_u \\ P_v \\ P_w \end{bmatrix} \quad (3.7)$$

Então, se a matriz em (3.7) for chamada de R_r , tem-se,

$$R = \begin{bmatrix} i_x i_u & i_x j_v & i_x k_w \\ j_y i_u & j_y j_v & j_y i_w \\ k_z i_u & k_z j_v & k_z i_w \end{bmatrix} \quad (3.8)$$

A relação em (3.2) é conhecida como transformação ortogonal, já que, como se pode ver em (3.7), a matriz R_r representa a rotação das coordenadas do sistema O_{uvw} em relação aos eixos do sistema de referência O_{xyz} . Se for selecionado um eixo, por exemplo OX , e um ângulo qualquer α , pelo qual se deseja fazer a rotação, então o ponto O_{xyz} de (3.3) pode ser calculado a partir de P_{uvw} (3.3) usando (3.4), (3.7) (i_u substituído por i_x) e a Fig. (3.3). A matriz R_r passa a R_x , α - matriz de rotação em relação ao eixo OX por um ângulo α . Assim, tem-se,

$$R_x, \alpha = \begin{bmatrix} i_x i_u & i_x j_v & i_x k_w \\ j_y i_u & j_y j_v & j_y i_w \\ k_z i_u & k_z j_v & k_z i_w \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \quad (3.9)$$

De maneira análogos, se o eixo escolhido fosse o OY ou OZ , então, poderiam ser definidas R_y, ϕ e R_z, θ que são respectivamente matrizes de rotação em relação aos eixos Y e Z com ângulos ϕ e θ .

Fig.3.3 - Posição dos eixos após rotação.

$$R_{y,\phi} = \begin{bmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{bmatrix}$$

(3.10)

$$R_{z,\theta} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

(3.11)

O segundo termo da expressão (3.1) ou (3.2) é definido como o elemento de translação. Ele representa a nova coordenada da câmera, antes representado no espaço tridimensional de referência, após o movimento de rotação.

Passa-se agora a considerar os aspectos da projeção ortogonal de uma imagem nos planos das câmeras. Para tanto, será considerado que uma imagem é um conjunto de pontos bidimensionais que podem ser calculados pelos métodos expostos em [1].

A projeção central dos pontos do espaço tridimensional fixo na câmera (ponto-imagem), nos seus respectivos planos, pode ser representada pela seguinte transformação:

$$H:I_c \rightarrow M$$

onde

$$r = (x_c, y_c, z_c)^T \in I = \{(x_c, y_c, z_c)^T \in R^3, z > 0\}$$

e

$$M = (X_{cam}, Y_{cam})^T \in R^2$$

é o modelo de projeção central daqueles pontos nos planos das câmeras. Este modelo pode ser definido da seguinte forma [4]:

$$(X_{cam}, Y_{cam})^T = \begin{pmatrix} \frac{f_{xc}}{z_c} \\ \frac{f_{yc}}{z_c} \end{pmatrix} + \begin{pmatrix} n_x \\ n_y \end{pmatrix} = h[r] + n, \quad (3.12)$$

onde f é a distância focal e (X_{cam}, Y_{cam}) são as coordenadas dos pontos $r(x_c, y_c, z_c)$ projetadas no plano da câmera. Esses pontos que são as coordenadas (x, y, z) no sistema referencial inercial fixo na câmera são obtidos por utilização das expressões (3.1) ou (3.2). Define-se ainda: $(X_{cam}, Y_{cam})_L$ e $(X_{cam}, Y_{cam})_R$, com sendo as projeções na câmera direita e esquerda, respectivamente.

$$(X_{cam}, Y_{cam})_L^T = \begin{pmatrix} \frac{f_{xc_L}}{z_{c_L}} \\ \frac{f_{yc_L}}{z_{c_L}} \end{pmatrix} + \begin{pmatrix} n_x \\ n_y \end{pmatrix} \quad (3.13)$$

$$(X_{cam}, Y_{cam})_R^T = \begin{pmatrix} \frac{f_{xc_R}}{z_{c_R}} \\ \frac{f_{yc_R}}{z_{c_R}} \end{pmatrix} + \begin{pmatrix} n_x \\ n_y \end{pmatrix} \quad (3.14)$$

onde X_{c_L} , Y_{c_L} , Z_{c_L} , X_{c_R} , Y_{c_R} e Z_{c_R} são calculados usando (3.1) e (3.2).

III.2.2 Modelo para observação de pontos ruidosos

O objetivo do modelo de observação de pontos ruidosos é encontrar, a partir de informações bidimensionais originárias de duas câmeras, a fórmula de recorrência dos pontos tridimensionais do objeto em movimento. As imagens das cenas são digitalizadas e projetadas nos planos das câmeras, conforme visto na Seção III.2.1. Isto caracteriza o processo de medição. Com as medidas feitas e fazendo-se manipulações algébricas, utilizando-se o método dos mínimos quadrados[1], chega-se a uma expressão analítica que permite calcular os pontos (x, y, z) . Então, a partir de (3.1) e (3.13) tem-se,

$$\begin{cases} X_{cl} = \cos \alpha_l (x + l) - z \sin \alpha_l \\ Y_{cl} = Y \\ Z_{cl} = \sin \alpha_l (x + l) + z \cos \alpha_l \end{cases} \quad e \quad \begin{cases} X_{cam_l} Z_{cl} = f X_{cl} \\ Y_{cam_l} Z_{cl} = f Y_{cl} \end{cases}$$

ou

$$\begin{cases} X_{cam_l} \frac{Z_{cl}}{f} = X \cos \alpha_l - Z \sin \alpha_l + l \cos \alpha_l \\ Y_{cam_l} \frac{Z_{cl}}{f} = Y \\ Z_{cl} = X \cos \alpha_l + l \sin \alpha_l - Z \cos \alpha_l \end{cases} \quad (3.15)$$

Aplicando Z_{CL} nas duas primeiras linhas de (3.15), tem-se,

$$\begin{cases} \frac{X_{cam_l}}{f} (X \sin \alpha_l + l \sin \alpha_l + Z \cos \alpha_l) = X \cos \alpha_l - Z \sin \alpha_l + l \cos \alpha_l \\ \frac{Y_{cam_l}}{f} (X \sin \alpha_l + l \sin \alpha_l + Z \cos \alpha_l) = Y \end{cases} \quad (3.16)$$

onde f e l são respectivamente a distância focal e o deslocamento em relação à origem.

Escrevendo (3.16) de uma outra forma, tem-se,

$$\begin{cases} \frac{-X_{cam_l}}{f} l \sin \alpha_l + l \cos \alpha_l = (-\cos \alpha_l + \frac{X_{cam_l}}{f} \sin \alpha_l) X + (\frac{-X_{cam_l}}{f} \cos \alpha_l + \sin \alpha_l) Z \\ \frac{-Y_{cam_l}}{f} l \sin \alpha_l = \frac{Y_{cam_l}}{f} X \sin \alpha_l - Y + \frac{Y_{cam_l}}{f} Z \cos \alpha_l \end{cases}$$

ou ainda

$$\begin{bmatrix} -\cos \alpha_l + \frac{X_{cam_l}}{f} \sin \alpha_l & 0 & \frac{X_{cam_l}}{f} \cos \alpha_l + \sin \alpha_l \\ \frac{Y_{cam_l}}{f} \sin \alpha_l & -1 & \frac{Y_{cam_l}}{f} \cos \alpha_l \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} \frac{-X_{cam_l}}{f} l \sin \alpha_l + l \cos \alpha_l \\ \frac{Y_{cam_l}}{f} l \sin \alpha_l \end{bmatrix}$$

Finalmente, considerando as duas câmeras, tem-se,

$$\begin{bmatrix} -\cos\alpha_l + \frac{X_{cam\,l}}{f}\sin\alpha_l & 0 & \frac{X_{cam\,l}}{f}\cos\alpha_l + \sin\alpha_l \\ \frac{X_{cam\,l}}{f}\sin\alpha_l & -1 & \frac{Y_{cam\,l}}{f}Z\sin\alpha_l \\ -\cos\alpha_r + \frac{X_{cam\,r}}{f}\sin\alpha_r & 0 & \frac{X_{cam\,r}}{f}\cos\alpha_r + \sin\alpha_r \\ \frac{X_{cam\,r}}{f}\sin\alpha_r & -1 & \frac{Y_{cam\,r}}{f}Z\sin\alpha_r \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} -\frac{X_{cam\,l}}{f}l\sin\alpha_l + l\cos\alpha_l \\ -\frac{Y_{cam\,l}}{f}l\sin\alpha_l \\ \frac{X_{cam\,r}}{f}l\sin\alpha_r - l\cos\alpha_r \\ \frac{Y_{cam\,r}}{f}l\sin\alpha_r \end{bmatrix}$$

$$HX = Z$$

(3.17)

A matriz em (3.17), H , é não quadrada, conseqüentemente não possui inversa, o que impossibilita a obtenção da solução única do sistema $HX = Z$. Mas se aos dois membros de (3.17) forem multiplicados por H^T , matriz transposta de H , obtém-se o sistema mostrado em (3.18). Já que a matriz $(H^T * H)$ desse novo sistema é inversível e (x, y, z) pode, então, ser unicamente determinado. A técnica utilizada acima é conhecida como Método dos Mínimos Quadrados aplicado à resolução de sistemas lineares.

$$(H^T H)X = H^T Z$$

(3.18)

$$X = (H^T H)^{-1} H^T Z$$

(3.19)

Modelo alternativo para observação de pontos ruidosos

A necessidade de utilizar o modelo descrito nessa seção surgiu durante a fase de testes com o modelo proposto em [1], usando-se seqüência de imagem simulada. Naquela fase foi verificado que os erros cometidos na recuperação dos componentes

associados ao eixo de projeção dos outros dois componentes eram elevados, ou seja, esses componentes não eram recuperados corretamente, empregando-se o procedimento descrito em [1]. Assim, consultando a referência [8], analisando-a e fazendo os mesmos testes feitos em [1], chegou-se a conclusão de que o método lá delineado poderia ser empregado na observação de pontos ruidosos, em substituição ao método descrito em [1].

A Figura 3.4 mostra a geometria necessária à compreensão desse método de observação de pontos ruidosos. Nessa figura a linha de visada da câmera coincide com o eixo x . $V_1 = (v_1, v_2, v_3)$ é o vetor separação entre as câmeras. Seja v_{c1} e v_{c2} os vetores com origem no centro de projeção da câmeras c_1 e c_2 , respectivamente, e extremidade no ponto de cena $X(x, y, z)$. A relação entre esses três últimos vetores é:

$$v_{c2} = v_{c1} - V_1 \quad (3.20)$$

O método consiste em determinar o *range*, definido como sendo a projeção de v_{c_l} na direção de N_p . O *range* tem a seguinte expressão:

$$R_c = v_{c_l} N_p \quad (3.22)$$

onde R_c é o *range* associado ao ponto de imagem na imagem 1 (y_{c_l}, z_{c_l}).

O vetor v_{c_l} , por ser colinear com $(f_{c_l}, y_{c_l}, z_{c_l})$, pode se definido como segue:

$$v_{c_l} = \beta(f_{c_l}, y_{c_l}, z_{c_l}) \quad (3.23)$$

Combinando (3.22) com (3.23) tem-se:

$$\beta = \frac{R_c}{n_1 f_{c_l} + n_2 y_{c_l} + n_3 z_{c_l}} \quad (3.24)$$

onde $N_p = (n_1, n_2, n_3)$

Dividindo (3.21) por (3.23) tem-se

$$\frac{\alpha}{\beta} \begin{bmatrix} f_{c_2} \\ y_{c_2} \\ z_{c_2} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} f_{c_1} \\ y_{c_1} \\ z_{c_1} \end{bmatrix} \quad (3.25)$$

Dividindo-se as linhas 2 e 3 pela linha 1 em (3.25), tem-se:

$$\frac{y_{c_2}}{f_{c_2}} = \frac{A_{c_2}}{A_{c_1}} \quad e \quad \frac{z_{c_2}}{f_{c_2}} = \frac{A_{c_3}}{A_{c_1}} \quad (3.26)$$

onde

$$A_{c_i} = r_{i1} f_{c_1} + r_{i2} y_{c_1} + r_{i3} z_{c_1}$$

A equação (3.25) pode, ainda, ser escrita da seguinte forma:

$$\begin{aligned} R_c &= (R_{11}y_{c2} - R_{12}f_{c2}) = T_1y_{c2} - T_2f_{c2} \\ R_c &= (R_{11}z_{c2} - R_{13}f_{c2}) = T_1z_{c2} - T_3f_{c2} \end{aligned} \quad (3.27)$$

com

$$\begin{aligned} R_i &= (r_{i1}f_{c1} + r_{i2}y_{c1} + r_{i3}z_{c1}) \\ T_i &= V_i(n_1f_{c1} + n_2y_{c1} + n_3z_{c1}) \end{aligned} \quad (3.28)$$

O sistema definido em (3.27) permite escrever a seguinte equação matricial:

$$A_c R_c = B_c \quad (3.29)$$

onde

$$A_c = \begin{bmatrix} R_{r1}y_{c1} - R_{r2}f_{c1} \\ R_{r1}z_{c1} - R_{r3}f_{c1} \end{bmatrix} \quad \text{e} \quad B_c = \begin{bmatrix} T_1y_{c1} - R_2f_{c1} \\ R_1z_{c1} - R_3f_{c1} \end{bmatrix} \quad (3.30)$$

Aplicando mínimos quadrados na resolução do sistema (3.30), tem-se:

$$R_c = (A_c^T A_c)^{-1} A_c^T B_c = (A_{c1} B_{c1} + A_{c2} B_{c2})(A_{c1}^2 + A_{c2}^2)^{-1} \quad (3.31)$$

Substituindo (3.31) em (3.24), pode-se obter as coordenadas do ponto X em relação à câmera $c1$ empregando a expressão (3.23).

Assim, com as informações de duas câmeras, sequência de imagem estéreo, é possível compor ponto a ponto uma cena filmada por elas, sem perda de qualquer uma das três dimensões. As medidas são tomadas em tempos discretos t_j e X contém as coordenadas de cada ponto em relação ao sistema referencial inercial I . Se j for a quantidade de pontos necessária, então pode-se escrever:

$$r_j(t_i) = X_j(t_i) \quad \text{ou}$$

$$r_j(t_i) = (x_j(t_i), y_j(t_i), z_j(t_i))$$

(3.32)

onde $r_j(t_i)$ representa as coordenadas dos pontos livres de ruído. Se no processo de medição os pontos forem ruidosos, então $r_j(t_i)$ é calculado segundo a expressão mostrada abaixo,

$$r_{nj}(t_i) = r_j(t_i) + n_j(t_i)$$

(3.33)

onde $n_j(t_i) = (n_{x_j}(t_i), n_{y_j}(t_i), n_{z_j}(t_i))$ é o ruído gaussiano com média zero, conforme considerações feitas em [1].

III.2.3 Modelo da estrutura do objeto

Nesta seção será feita a análise de uma expressão que permite estimar a estrutura de um objeto em movimento no espaço referencial inercial I . O objeto é suposto ser um corpo rígido e cada nova posição ocupada pelo mesmo é registrada por um par de câmeras que fazem, independentes uma da outra, a projeção da imagem do objeto. Um sistema de coordenadas fixo no objeto e contido no sistema inercial I será definido. Serão examinadas duas abordagens. De início será analisada a expressão sem incerteza (livre de ruído) e em seguida será feita uma modificação nesta última com a finalidade de obter uma expressão que possibilite a obtenção das estimativas da estrutura do objeto na presença de ruído.

Seja um sistema de coordenadas tridimensional fixo no objeto e com origem no ponto O_s . Chamemos este sistema de S . Seja s qualquer ponto tridimensional pertencente a S . Dos infinitos pontos que formam o objeto são suficientes três deles para se estimar a cinética do mesmo. Não há uma razão especial que justifique o uso

vetores diretores de S . O ponto s pode ser representado por suas coordenadas em S . Então, tem-se,

$$s_j = \begin{bmatrix} x_{sj} \\ y_{sj} \\ z_{sj} \end{bmatrix} \quad (3.34)$$

onde j individualiza um ponto na imagem.

Essas coordenadas representam a estrutura do objeto e não são funções do tempo, pois o objeto é assumido ser um corpo rígido.

Para se encontrar o sistema coordenado S que representa a estrutura do objeto é necessária a definição de três eixos coordenados, fixos no objeto, como sendo os vetores diretores do sistema e uma origem.

Qualquer ponto pode ser origem de S . Então, é razoável admitir que o centro de massa seja a origem. Então, tem-se,

$$c_m(t_i) = \frac{1}{m} \sum_{j=1}^m r_j(t_i) \quad (3.35)$$

onde m é o número de pontos utilizados.

O centro de massa calculado por (3.36) é um valor exato, pois o seu cálculo depende de $r(t)$, que de acordo com os argumentos anteriores não possui ruído.

Para se calcular os vetores ortonormais, é necessário que os três pontos sejam não colineares. Então a partir da Figura (3.35), tem-se,

i) Cálculo do vetor e_x

$$e_x(t_i) = \frac{(r_2(t_i) - r_1(t_i))}{\|(r_2(t_i) - r_1(t_i))\|} \quad (3.36)$$

ii) Cálculo do vetor e_y

$$e_y(t_i) = \frac{e_x(t_i)(r_3(t_i) - r_1(t_i))}{\|e_x(t_i)(r_3(t_i) - r_1(t_i))\|} \quad (3.37)$$

iii) Cálculo do vetor e_z

$$e_z(t_i) = e_x(t_i) \times e_y(t_i) \quad (3.38)$$

Fig. 3.5 - Vetores diretores de S.

Como se pode ver em (3.37), (3.38) e (3.39), os vetores e_x , e_y e e_z são avaliados no tempo de medição t_i .

Tendo definido o centro de massa como sendo a origem do sistema coordenado S e as expressões dos vetores diretores, pode-se, então, calcular os pontos s_j pertencentes a S . Esses pontos representam a estrutura do objeto em S e quanto maior for a

quantidade deles utilizada, melhor será a definição da forma do objeto. Contudo, o uso de três pontos s_j , j inteiro e $0 \leq j \leq 3$) é suficiente para a estimação dos parâmetros relacionados com os movimentos do objeto. Então, da Figura (3.6), tem-se,

Fig.3.6 - O sistema S fixo no objeto.

$$x_{sj} = e_x(t_i) (r_j(t_i) - c_m(t_i))$$

$$y_{sj} = e_y(t_i) (r_j(t_i) - c_m(t_i))$$

$$z_{sj} = e_z(t_i) (r_j(t_i) - c_m(t_i))$$

e

$$s_j = \begin{bmatrix} x_{sj} \\ y_{sj} \\ z_{sj} \end{bmatrix} \quad (3.40)$$

A expressão (3.40) é a fórmula de recorrência de todos os pontos que formam a estrutura do objeto na ausência de ruído. Ela não depende do tempo porque, conforme considerações feitas no início da seção o objeto é um corpo rígido. O índice t_i nas expressões diz respeito ao tempo em que um valor de s_j é estimado. As variações nos valores desse índice não significa que está ocorrendo modificações dos componentes s_j em relação ao centro de massa.

Se as coordenadas tridimensionais, $r_j(t_i)$, são contaminadas por ruído, o que faz com que seja usada a expressão (3.39) ao invés da (3.38) nos cálculos, então o centro de massa passa a ser calculado por uma expressão, (3.42) que é uma modificação feita em (3.35) para incluir o ruído de medição.

$$\hat{c}_m = \frac{1}{m} \sum_{i=1}^m r_{nj}(t_i) \quad (3.41)$$

A exemplo de $\hat{c}_m(t_i)$, e_x , e_y e e_z , que dependem de $r_n(t_i)$, são substituídos por $\hat{e}_x$, $\hat{e}_y$ e $\hat{e}_z$ com a mesma finalidade.

Com essas modificações, $s(t)$ não mais é um valor exato, pois possui incerteza. Contudo, ele pode ser estimado pelo uso da expressão (3.42), uma expressão recursiva no tempo de estimativas, em que o valor provável é a média de $s_j(t_i)$ ou seja, $\hat{s}(t_i)$.

$$\hat{s}_j(t_i) = \frac{1}{i+1} \sum_{k=0}^i \begin{bmatrix} e_x(t_k)(r_{nj}(t_k) - \hat{c}_m(t_k)) \\ e_y(t_k)(r_{nj}(t_k) - \hat{c}_m(t_k)) \\ e_z(t_k)(r_{nj}(t_k) - \hat{c}_m(t_k)) \end{bmatrix} \quad (3.42)$$

Assim, com as expressões desenvolvidas nesta seção pode-se calcular a estrutura do objeto. Quando não há ruído no processo de medida, a estrutura é encontrada por (3.40). Porém, se existe incerteza nas medidas, devido ao fato delas serem ruidosas, então a estrutura é definida por (3.42).

III.2.4 Modelos que descrevem os movimentos do objeto

Esta seção será dividida em duas subseções. A primeira desenvolverá as equações do movimento de translação com aceleração constante e a segunda apresentará as deduções das equações do movimento de rotação com velocidade angular constante.

Modelo de translação

O objetivo desta subseção é definir as equações do movimento de translação do centro de rotação do objeto - o ponto OB . Este ponto coincide com o centro de massa e é fixo em relação ao sistema de coordenadas S definido na seção anterior ou seja, de todos os pontos que formam o objeto, OB é o único que não realiza movimento de rotação. Pode, contudo, movimentar-se com velocidade variante no tempo e aceleração constante.

Diante dos argumentos feitos no parágrafo anterior, o movimento de translação será representado pelas equações da cinemática que descrevem o movimento uniformemente variante no tempo de um ponto material (OB) em relação ao espaço inercial I . Então, tem-se,

$$\begin{cases} \frac{dr}{dt} = v(t) \\ \frac{dv}{dt} = a(t) \\ \frac{da}{dt} = 0 \end{cases} \quad (3.44)$$

Resolvendo-se o sistema de equações em (3.44) por integração direta, tem-se

$$\begin{cases} r(t_2) = r(t_1) + (t_2 - t_1) v(t_1) + (t_2 - t_1)^2 \frac{a(t_1)}{2} \\ v(t_2) = v(t_1) + (t_2 - t_1) a(t_1) \\ a(t_2) = a(t_1) \end{cases} \quad (3.45)$$

Assim, os vetores $r(t) = (r_x(t), r_y(t), r_z(t))$, $v(t) = (v_x(t), v_y(t), v_z(t))^T$ e $a = (a_x, a_y, a_z)$ são respectivamente a posição, a velocidade e a aceleração do centro de rotação OB em relação ao sistema referencial inercial I e $t_2 - t_1$ é tempo de amostragem τ .

Modelo de rotação

Conforme mencionado anteriormente, as equações que representam o movimento de rotação fará uso dos quatêrnios. Deu-se preferência pela utilização desta abordagem por motivos que ficarão evidentes nas seções subseqüentes.

Seja o vetor $w = (w_x, w_y, w_z)^T$ a velocidade angular em relação ao centro de rotação OB com coordenadas referidas ao sistema inercial I . A velocidade w é constante no tempo, conforme considerações feitas no início da seção. Então o vetor de quatro coordenadas, expressão mostrada a seguir, (os quatro quatêrnios) é usado para propagar o movimento de rotação do objeto em volta do centro de rotação OB ,

$$q(t) = (q_1(t), q_2(t), q_3(t), q_4(t)) \quad (3.46)$$

Os quatérnios formam um vetor unitário que está relacionado com os elementos de rotação, isto é, ângulo de rotação e eixo de referência. A expressão a seguir mostra esta relação:

$$q(t) = (n_1 \cos(-\frac{\alpha}{2}), n_2 \sin(-\frac{\alpha}{2}), n_3 \sin(-\frac{\alpha}{2}), \cos(-\frac{\alpha}{2}))^T$$

onde o vetor $n = (n_1, n_2, n_3)$ é a direção do eixo de rotação e α o ângulo pelo qual é feito o movimento de rotação em volta do vetor n . A equação a seguir, cujo detalhes de sua expressão podem ser encontrados na referência [4], é a equação de propagação no tempo do vetor $q(t)$, isto é,

$$\begin{aligned} \dot{q}(t) &= \Omega[w] q(t_o) \\ q(t_o) &= q_o \end{aligned}$$

onde

$$\Omega[w] = \frac{1}{2} \begin{bmatrix} 0 & w_z & -w_y & w_x \\ -w_z & 0 & w_x & w_y \\ w_y & -w_x & 0 & w_z \\ -w_x & -w_y & -w_z & 0 \end{bmatrix} \quad (3.47)$$

Como se pode observar, (3.47) é um sistema linear invariante no tempo da forma $dx/dt = Ax$, onde $A = \Omega(w)$, pois $\Omega(w)$ não depende do tempo. Assim, podem ser empregados na resolução de (3.47) os métodos bastantes utilizados em teoria de sistemas lineares [42].

Conforme teoria desenvolvida em [39], pode-se resolver a equação diferencial (3.47) a partir do cálculo da matriz de transição de $\Omega(w)$. A seguir será descrito, resumidamente, como isto é pode ser feito.

i) Inicialmente define-se a matriz de transição de $\Omega(w)$, isto é,

$$\phi(w) = e^{\Omega \tau}$$

onde $\tau = t_{i+1} - t_i$

ii) Calcula-se a solução de (3.47), ou seja,

$$q(t) = q(t_0) e^{\Omega \tau}$$

iii) Calcula-se o valor de $f(\Omega(w)) = \exp[\Omega(w) \tau]$

iii.1) Inicialmente calcula-se os autovalores de $\Omega(w)$

$$0 = |\lambda I_4 - \Omega| = -\frac{1}{2} \begin{vmatrix} \lambda & w_z & -w_y & w_x \\ -w_z & \lambda & w_x & w_y \\ w_y & -w_x & \lambda & w_z \\ -w_x & -w_y & -w_z & \lambda \end{vmatrix}$$

Com $\lambda_0 = j|\Omega(w)|$; $\lambda_1 = -j|\Omega(w)|$; $\lambda_2 = 0$ e $\lambda_3 = 0$

iii.2) Define-se o polinômio $g(\lambda) = a_0 + a_1 \lambda + \dots$

iii.3) Usa-se $f(\lambda) = e[\lambda \tau]$

iii.4) Faz-se $f(\lambda) = g(\lambda)$ com o objetivo de calcular os coeficientes de $g(\lambda)$ do item iii.2.

$$a_0 = \cos\left(\|w\| \frac{\tau}{2}\right)$$

$$a_1 = \frac{2}{\|w\|} \cos\left(\|w\| \frac{\tau}{2}\right)$$

equações marcadas em azul devem ser passadas para *itálico*.

iv) Calcula-se $f(\Omega(w)) = e^{[\Omega(w)\lambda]}$ utilizando os valores de a_0 e a_1 .

$$f(\Omega[w]) = \cos\left(\|w\| \frac{\tau}{2}\right) I_4 + \frac{2}{\|w\|} \cos\left(\|w\| \frac{\tau}{2}\right) \Omega$$

v) Finalmente, substitui-se esse último valor de $f(\Omega(w))$ no item ii, por conta do item iii, para definir a solução da equação diferencial (3.48).

$$q(t) = q(t_0) \left(\cos\left(\frac{\|w\|\tau}{2}\right) I_4 + \frac{2}{\|w\|} \cos\left(\frac{\|w\|\tau}{2}\right) \Omega \right) \quad (3.49)$$

Portanto, o movimento de rotação com velocidade angular $w = (w_x, w_y, w_z)$ é perfeitamente propagado pela equação diferencial (3.47) que tem solução mostrada em (3.49). Esta equação, por revelar a relação linear existente entre o vetor w e os quatêrnios, pode ser resolvida à luz da teoria dos sistemas lineares [39], onde o conceito de matriz de transição, cálculo dos autovalores de matrizes, etc, são recursos que permitem a resolução de equações diferenciais como a (3.47). Aspectos como estes e outros a serem exibidos na Seção III.3.4 neste trabalho confirmam as vantagens de usar os quatêrnios na propagação do movimento de rotação.

III.3 Formulação do problema de estimação recursiva

III.3.1 Considerações iniciais

O algoritmo recursivo que será analisado nesta seção será usado para observar o movimento de um objeto em um espaço referencial inercial. A característica fundamental desse tipo de aproximação é que o método geralmente necessita de muito menos computação para cada novo quadro medido por usar, em cada cálculo, as informações adquirida no passo imediatamente anterior. O algoritmo continuamente estima e propaga os estados no tempo, tendo como base informações obtidas no passado, para auxiliar no processamento do próximo conjunto de dados. As estimações e propagações continuam até que seja atingida uma condição em que as estimativas obtidas seja aceitáveis à luz de um critério de aceitação ou rejeição.

Um outro aspecto dos algoritmos recursivos é que aproximações, tais como a natureza não-linear da planta e/ou do processo de medição são requeridas. Isto significa, por um lado, que os estados estimados não são valores suficientes para os dados passados, como seriam nas aproximações lineares, o que torna ineficiente a utilização dos dados passados e, conseqüentemente, degradação dos estados estimados. Por outro lado, se os modelos utilizados na descrição da situação são apenas aproximações da verdadeira planta e do processo de medição, mas com pequenos desvios dos verdadeiros valores, então, segundo [4], os requisitos de memória não são críticos nas implementações de algoritmos recursivos.

Um filtro recursivo será usado para estimar os parâmetros móveis, isto é, a posição, a velocidade, a aceleração e a velocidade angular, a partir de medições feitas em quadros de imagens de uma seqüência de imagens etéreo. Nesses quadros, atributos, pontos de cantos existentes nas imagens, são capturados. Esses pontos são parâmetros de entrada das equações que representam o filtro recursivo

O estimador IEKF [38], bastante utilizado em sistemas não-lineares, é de natureza recursiva. Essa característica é desejada na implementação de algoritmos destinados à observação do movimento do objeto, desde que os parâmetros

móveis podem ser propagados sempre que uma nova medição for realizada. As medidas são feitas em intervalos de tempo t , oportunidade em que pontos são lidos dos quadros de imagens. A estrutura do objeto é, então, estimada pelo algoritmo desenvolvido na Seção III.2.3. O IEKF estima os parâmetros móveis do objeto, considerando as estimativas passadas, devido a sua natureza recursiva, até que uma nova medida seja feita, reiniciando todo o processo.

Nesta seção os parâmetros estimados serão representados pelo vetor de estado mostrado a seguir, isto é,

$$x(t) = (r^T(t), v^T(t), a^T, q^T(t), w^T) \quad (3.50)$$

onde

$$r(t) = (r_x(t), r_y(t), r_z(t)) \quad (3.51)$$

$$v(t) = (v_x(t), v_y(t), v_z(t)) \quad (3.52)$$

$$a = (a_x, a_y, a_z) \quad (3.53)$$

$$q(t) = (q_1(t), q_2(t), q_3(t), q_4(t)) \quad (3.54)$$

$$w = (w_x, w_y, w_z) \quad (3.55)$$

A estimativa do vetor de estado será feita pelo IEKF.

III.3.2 Sistema de equações dinâmicas do movimento do objeto

O comportamento no tempo dos estados, bem como o processo de medição serão representados pelo sistema de equações não-lineares mostrado em (3.56). Este sistema contém o modelo da planta e o modelo do processo de medição. Como se pode perceber em (3.56), a não-linearidade deve-se ao fato de h ser função não linear do vetor posição $X(t)$ ou $r(t)$ (expressão (3.19)) e f ser função do estado $x(t)$. O ruído de processo, $w(t)$ e o de medição, $v_n(t)$, são gaussianos com média zero e covariância definida, respectivamente, por (3.57) e (3.58).

$$\begin{cases} \frac{dx}{dt} = f[x(t); t] + G(t)w(t) \\ z(t_i) = h[x(t_i); t] + v_n(t_i) \end{cases} \quad (3.56)$$

$$E\{w(t_i)w^T(t+\tau)\} = Q(t)\delta(\tau) \quad (3.57)$$

$$E\{v_n(t_i)v_i^T(t_j)\} = \begin{cases} R_v(t_i), & t_i = t_j \\ 0, & t_i \neq t_j \end{cases} \quad (3.58)$$

III.3.3 Equação da planta

Nesta seção será definida a equação da planta que permite a propagação do vetor de estado $x(t)$ definido em (3.50). Já são conhecidas as equações de propagação da posição, da velocidade e da aceleração (3.44) e dos quatêrnios (3.47). A equação da planta é, portanto, um sistema que contém todas essas equações de propagação. Assim, tem-se,

$$\frac{d}{dt} \begin{bmatrix} r(t) \\ v(t) \\ a \\ q(t) \\ w \end{bmatrix} = \begin{bmatrix} 0_{3 \times 3} & I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 4} & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & I_{3 \times 3} & 0_{3 \times 4} & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 4} & 0_{3 \times 3} \\ 0_{4 \times 3} & 0_{4 \times 3} & 0_{4 \times 3} & \Omega[w] & 0_{4 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 4} & 0_{3 \times 3} \end{bmatrix} \begin{bmatrix} r(t) \\ v(t) \\ a \\ q(t) \\ w \end{bmatrix} \quad (3.59)$$

onde $\Omega[w]$ é definida em (3.48).

Observando o sistema (3.59), percebe-se que o mesmo é não-linear. Um dos estados, o vetor w , tem seus componentes como elemento da matriz do sistema (eles aparecem nas linhas e nas colunas de $\Omega[w]$). Diante de tal não-linearidade, deve ser definido um procedimento de linearização do sistema em um ponto de operação bem definido.

A seguir será feita a linearização do sistema (3.59) no ponto de operação x_{op} , oportunidade em que será calculada a matriz $\frac{df[x(t)]}{dt} = F$, uma aproximação linear de $f[x(t)]$, de acordo com o critério de linearização de funções por desenvolvimento em série de Taylor [41].

Seja $x(t)$ o vetor de estado definido em (3.50). Este vetor pode se rescrito de uma outra forma em que são mostrados todas as suas componentes. Então, tem-se:

$$x(t) = [x_0 \quad x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7 \quad x_8 \quad x_9 \quad x_{10} \quad x_{11} \quad x_{12} \quad x_{13} \quad x_{14} \quad x_{15}]^T \quad (3.60)$$

Considerando-se (3.60) e (3.59), pode-se escrever $f[x(t)]$ como,

$$f[x(t)] = \dot{x}(t) = \begin{bmatrix} x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ 0 \\ 0 \\ 0 \\ 0.5(x_{15}x_{10} - x_{14}x_{11} + x_{13}x_{13}x_{12}) \\ -0.5(x_{15}x_9 - x_{13}x_{11} + x_{13}x_{14}x_{12}) \\ 0.5(x_{14}x_9 - x_{13}x_{10} + x_{13}x_{15}x_{12}) \\ -0.5(x_{13}x_9 - x_{14}x_{10} + x_{13}x_{15}x_{12}) \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (3.61)$$

então a derivada de $f[x(t)]$ em relação a $x(t)$ é,

$$F = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x_{15} & -x_{14} & x_{13} & x_{12} & -x_{11} & x_{10} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -x_{15} & 0 & x_{13} & x_{14} & x_{11} & x_{12} & -x_{10} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x_{14} & -x_{13} & 0 & x_{15} & -x_{10} & x_9 & x_{12} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -x_{13} & -x_{14} & x_{15} & 0 & -x_9 & -x_{10} & x_{11} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(3.62)

Uma vez calculada a matriz F , pode-se rescrever (3.59) como um sistema linear invariante no tempo. Então,

$$\frac{dx}{dt} = Fx + G(t)w(t) \quad (3.63)$$

Esta expressão permite escrever as seguintes equações:

$$\frac{dr}{dt} = I_{3 \times 3} r(t) \quad (3.64)$$

$$\frac{dv}{dt} = I_{3 \times 3} v(t) \quad (3.65)$$

$$\frac{dq}{dt} = \Omega[w]_{4 \times 4} q(t) \quad (3.66)$$

$$\frac{da}{dt} = 0_{3 \times 3} a(t) \quad (3.67)$$

$$\frac{dw}{dt} = 0_{3 \times 3} w(t) \quad (3.68)$$

Pode-se escrever:

$$\begin{bmatrix} r(t_{i+1}^-) \\ v(t_{i+1}^-) \\ a(t_{i+1}^-) \\ q(t_{i+1}^-) \\ w(t_{i+1}^-) \end{bmatrix} = \begin{bmatrix} I_{3 \times 3} & \tau I_{3 \times 3} & \frac{\tau^2}{2} I_{3 \times 3} & 0_{3 \times 4} & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & \tau I_{3 \times 3} & 0_{3 \times 4} & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 4} & 0_{3 \times 3} \\ 0_{4 \times 3} & 0_{4 \times 3} & 0_{4 \times 3} & \Omega[w] & 0_{4 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 4} & 0_{3 \times 3} \end{bmatrix} \begin{bmatrix} r(t_i^+) \\ v(t_i^+) \\ a(t_i^+) \\ q(t_i^+) \\ w(t_i^+) \end{bmatrix} \quad (3.69)$$

onde

$$\tau = t_{i+1}^- - t_i^+$$

é o tempo de medida entre quadros consecutivos e $\phi(w)$ a matriz de transição do sistema linear (3.66) dada por

$$\phi(w) = \cos(|w| \frac{\tau}{2}) I_{4 \times 4} + \frac{2}{|w|} \cos(|w| \frac{\tau}{2}) \Omega \quad (3.70)$$

A matriz de transição do sistema (3.69) foi obtida para propagar o vetor de estado $x(t)$. É interessante notar que, com o uso de (3.69), não é necessário integração numérica para resolver a equação diferencial (3.59). A única operação utilizada para se calcular $x(k+1)$ a partir de $x(k)$ é um produto matricial.

III.3.4 Equação de medida

A equação de medida do sistema de equações de estado (3.56) será definida nesta seção. Nesta equação $r(t)$ é o estado que contém informação da posição do centro de rotação OB em relação ao sistema referencial inercial I , determinada a partir do modelo de observação de pontos ruidosos (Seção III.2.1.1), $s_j(t)$ são as coordenadas dos pontos que formam a estrutura do objeto, calculadas empregando-se o algoritmo analisado na Seção III.2.3 ((3.40) ou (3.42)) e $R_r(t_i)$ é a matriz de rotação de coordenadas que permite alinhar os eixos do sistema S com os do sistema I . Assim, a equação de medida é uma transformação de coordenadas, que consiste de uma rotação seguida de uma translação.

A matriz de alinhamento $R_r(t)$ é bastante utilizada, por exemplo, no cálculo de posição de veículos em um sistema referencial [37]. Podem ser usadas várias configurações dessa matriz nesse cálculo, mas as mais freqüentes são a matriz dos ângulos de Euler [36], matriz dos cossenos diretores [36] e a matriz dos quatêrnios [36] e [1]. Nesta seção será dada preferência pela matriz dos quatêrnios na construção da equação do processo de medição. Esta escolha é justificada pelo fato de que tanto a matriz de Euler como a matriz dos cossenos diretores possuem termos com expressões trigonométricas. Essas expressões, por não ser fácil a manipulação algébrica de produto de senos e cossenos, dificulta o cálculo da primeira derivada, fundamental no processo de linearização de sistemas, e podem apresentar problemas numéricos [37] na avaliação dos senos e cossenos.

Os quatêrnios foram, segundo [36], desenvolvidos por Hamilton em 1843, mas não eram utilizados em dinâmica de corpos rígidos até bem pouco tempo atrás. Eles formam um conjunto de quatro elementos, (3.45), que estão relacionados com os elementos de rotação entre sistemas coordenados pela expressão (3.46). Segundo comentários feitos em [37], é o quarto termo adicional em (3.45) que melhora os cálculos numéricos feitos na propagação do movimento de rotação, onde ocorre multiplicações termo a termo das matrizes envolvidas, pois com eles os produtos entre os termos das matrizes são mais simples, mesmo que a sua presença implique o

aumento da quantidade destas operações (aumenta de 3x3, quando $R_T(t)$ é a matriz de Euler, para 4x4, quando $R_T(t)$ é a matriz dos quatêrnios).

A partir de (3.46) e a matriz de Euler, pode-se encontrar uma expressão de R_T que depende apenas dos quatêrnios, isto é,

$$R_T[q] = \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 + q_3q_4) & 2(q_1q_3 - q_2q_4) \\ 2(q_1q_2 - q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 + q_1q_4) \\ 2(q_1q_3 + q_2q_4) & 2(q_2q_3 - q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix} \quad (3.71)$$

Com a definição da expressão de $R_T[q(t)]$, é possível escrever a equação do processo de medição de uma outra forma, em que $R_T(t)$ é substituída por $R_T[q(t)]$.

Então, tem-se: $z(t_i) = h[x(t_i)] + v_n(t_i)$ ou

$$\begin{bmatrix} x_{n1}(t_i) \\ y_{n1}(t_i) \\ z_{n1}(t_i) \\ \cdot \\ \cdot \\ \cdot \\ x_{n3}(t_i) \\ x_{n3}(t_i) \\ x_{n3}(t_i) \end{bmatrix} = \begin{bmatrix} r_x(t_i) \\ r_y(t_i) \\ r_z(t_i) \\ \cdot \\ \cdot \\ \cdot \\ r_x(t_i) \\ r_y(t_i) \\ r_z(t_i) \end{bmatrix} + \begin{bmatrix} R_T[q(t_i)] \\ \cdot \\ \cdot \\ \cdot \\ R_T[q(t_i)] \end{bmatrix} \begin{bmatrix} \hat{s}_{x1}(t_i) \\ \hat{s}_{y1}(t_i) \\ \hat{s}_{z1}(t_i) \\ \cdot \\ \cdot \\ \cdot \\ \hat{s}_{x3}(t_i) \\ \hat{s}_{y3}(t_i) \\ \hat{s}_{z3}(t_i) \end{bmatrix} + \begin{bmatrix} n_{x1}(t_i) \\ n_{y1}(t_i) \\ n_{z1}(t_i) \\ \cdot \\ \cdot \\ \cdot \\ n_{x3}(t_i) \\ n_{y3}(t_i) \\ n_{z3}(t_i) \end{bmatrix}$$

(3.72)

A expressão (3.72) é a equação de medida para o IEKF. Nesta expressão três medidas tridimensionais são requeridas (3 cantos), isto é, $j = 3$, e a estrutura do objeto é calculada pelo estimador da Seção III.2.3, expressão (3.42). Por causa da não linearidade existente em (3.72), $z(t)$ é função não-linear do estado $r(t)$, nesse caso deve

ser feita a linearização de $z(t)$ no ponto de operação, definido aqui como sendo as estimativas de $x(t)$. Sejam $H = \frac{dh[x(t)]}{dx}$, $x(t) = (r(t)^T, v(t)^T, a^T, q(t)^T, w^T)$ e $z(t)$ tendo a expressão definida em (3.73). Então $\partial h / \partial x$ é representado por:

$$\frac{\partial h}{\partial x} = \begin{bmatrix} \frac{\partial h_{11}}{\partial r(t)} & \frac{\partial h_{11}}{\partial v(t)} & \frac{\partial h_{11}}{\partial a} & \frac{\partial h_{11}}{\partial q(t)} & \frac{\partial h_{11}}{\partial w} \\ \frac{\partial h_{21}}{\partial r(t)} & \frac{\partial h_{21}}{\partial v(t)} & \frac{\partial h_{21}}{\partial a} & \frac{\partial h_{21}}{\partial q(t)} & \frac{\partial h_{21}}{\partial w} \\ \frac{\partial h_{31}}{\partial r(t)} & \frac{\partial h_{31}}{\partial v(t)} & \frac{\partial h_{31}}{\partial a} & \frac{\partial h_{31}}{\partial q} & \frac{\partial h_{31}}{\partial w} \end{bmatrix} \quad (3.73)$$

onde os elementos de (3.73) são definidos nos itens i a v a seguir.

i) Derivada de $h(t)$ em relação ao estado

$$r(t) = (x(t), y(t), z(t)).$$

$$\frac{\partial h_{11}}{\partial r(t)} \triangleq \begin{bmatrix} \frac{\partial x_1(t_i)}{\partial x(t)} & \frac{\partial x_1(t_i)}{\partial y(t)} & \frac{\partial x_1(t_i)}{\partial z(t)} \\ \frac{\partial y_1(t_i)}{\partial x(t)} & \frac{\partial y_1(t_i)}{\partial y(t)} & \frac{\partial y_1(t_i)}{\partial z(t)} \\ \frac{\partial z_1(t_i)}{\partial x(t)} & \frac{\partial z_1(t_i)}{\partial y(t)} & \frac{\partial z_1(t_i)}{\partial z(t)} \end{bmatrix} = I_{3 \times 3} \quad (3.74)$$

$$\frac{\partial h_{21}}{\partial r(t)} \triangleq \begin{bmatrix} \frac{\partial x_2(t_i)}{\partial x(t)} & \frac{\partial x_2(t_i)}{\partial y(t)} & \frac{\partial x_2(t_i)}{\partial z(t)} \\ \frac{\partial y_2(t_i)}{\partial x(t)} & \frac{\partial y_2(t_i)}{\partial y(t)} & \frac{\partial y_2(t_i)}{\partial z(t)} \\ \frac{\partial z_2(t_i)}{\partial x(t)} & \frac{\partial z_2(t_i)}{\partial y(t)} & \frac{\partial z_2(t_i)}{\partial z(t)} \end{bmatrix} = I_{3 \times 3} \quad (3.75)$$

$$\frac{\partial h_{31}}{\partial r(t)} \overset{\Delta}{=} \begin{bmatrix} \frac{\partial x_3(t_i)}{\partial x(t)} & \frac{\partial x_3(t_i)}{\partial y(t)} & \frac{\partial x_3(t_i)}{\partial z(t)} \\ \frac{\partial y_3(t_i)}{\partial x(t)} & \frac{\partial y_3(t_i)}{\partial y(t)} & \frac{\partial y_3(t_i)}{\partial z(t)} \\ \frac{\partial z_3(t_i)}{\partial x(t)} & \frac{\partial z_3(t_i)}{\partial y(t)} & \frac{\partial z_3(t_i)}{\partial z(t)} \end{bmatrix} = I_{3 \times 3}$$

(3.76)

onde $I_{3 \times 3}$ é a matriz identidade de ordem 3.

ii) Derivada de $h(t)$ em relação ao estado

$$v(t) = (v_x(t), v_y(t), v_z(t)).$$

$$\frac{\partial h_{11}}{\partial v(t)} \overset{\Delta}{=} \begin{bmatrix} \frac{\partial x_1(t_i)}{\partial v_x(t)} & \frac{\partial x_1(t_i)}{\partial v_y(t)} & \frac{\partial x_1(t_i)}{\partial v_z(t)} \\ \frac{\partial y_1(t_i)}{\partial v_x(t)} & \frac{\partial y_1(t_i)}{\partial v_y(t)} & \frac{\partial y_1(t_i)}{\partial v_z(t)} \\ \frac{\partial z_1(t_i)}{\partial v_x(t)} & \frac{\partial z_1(t_i)}{\partial v_y(t)} & \frac{\partial z_1(t_i)}{\partial v_z(t)} \end{bmatrix} = 0_{3 \times 3}$$

(3.77)

$$\frac{\partial h_{21}}{\partial v(t)} \overset{\Delta}{=} \begin{bmatrix} \frac{\partial x_2(t_i)}{\partial v_x(t)} & \frac{\partial x_2(t_i)}{\partial v_y(t)} & \frac{\partial x_2(t_i)}{\partial v_z(t)} \\ \frac{\partial y_2(t_i)}{\partial v_x(t)} & \frac{\partial y_2(t_i)}{\partial v_y(t)} & \frac{\partial y_2(t_i)}{\partial v_z(t)} \\ \frac{\partial z_2(t_i)}{\partial v_x(t)} & \frac{\partial z_2(t_i)}{\partial v_y(t)} & \frac{\partial z_2(t_i)}{\partial v_z(t)} \end{bmatrix} = 0_{3 \times 3}$$

(3.78)

$$\frac{\partial h_{31}}{\partial v(t)} \overset{\Delta}{=} \begin{bmatrix} \frac{\partial x_3(t_i)}{\partial v_x(t)} & \frac{\partial x_3(t_i)}{\partial v_y(t)} & \frac{\partial x_3(t_i)}{\partial v_z(t)} \\ \frac{\partial y_3(t_i)}{\partial v_x(t)} & \frac{\partial y_3(t_i)}{\partial v_y(t)} & \frac{\partial y_3(t_i)}{\partial v_z(t)} \\ \frac{\partial z_3(t_i)}{\partial v_x(t)} & \frac{\partial z_3(t_i)}{\partial v_y(t)} & \frac{\partial z_3(t_i)}{\partial v_z(t)} \end{bmatrix} = 0_{3 \times 3} \quad (3.79)$$

onde O_3 é a matriz nula de ordem 3.

iii) Derivada de $h(t_i)$ em relação ao estado $a = (a_x, a_y, a_z)$

$$\frac{\partial h_{11}}{\partial a} \overset{\Delta}{=} \begin{bmatrix} \frac{\partial x_1(t_i)}{\partial a_x} & \frac{\partial x_1(t_i)}{\partial a_y} & \frac{\partial x_1(t_i)}{\partial a_z} \\ \frac{\partial y_1(t_i)}{\partial a_x} & \frac{\partial y_1(t_i)}{\partial a_y} & \frac{\partial y_1(t_i)}{\partial a_z} \\ \frac{\partial z_1(t_i)}{\partial a_x} & \frac{\partial z_1(t_i)}{\partial a_y} & \frac{\partial z_1(t_i)}{\partial a_z} \end{bmatrix} = 0_{3 \times 3} \quad (3.80)$$

$$\frac{\partial h_{21}}{\partial a} \overset{\Delta}{=} \begin{bmatrix} \frac{\partial x_2(t_i)}{\partial a_x} & \frac{\partial x_2(t_i)}{\partial a_y} & \frac{\partial x_2(t_i)}{\partial a_z} \\ \frac{\partial y_2(t_i)}{\partial a_x} & \frac{\partial y_2(t_i)}{\partial a_y} & \frac{\partial y_2(t_i)}{\partial a_z} \\ \frac{\partial z_2(t_i)}{\partial a_x} & \frac{\partial z_2(t_i)}{\partial a_y} & \frac{\partial z_2(t_i)}{\partial a_z} \end{bmatrix} = 0_{3 \times 3} \quad (3.81)$$

$$\frac{\partial h_{31}}{\partial a} \overset{\Delta}{=} \begin{bmatrix} \frac{\partial x_3(t_i)}{\partial a_x} & \frac{\partial x_3(t_i)}{\partial a_y} & \frac{\partial x_3(t_i)}{\partial a_z} \\ \frac{\partial y_3(t_i)}{\partial a_x} & \frac{\partial y_3(t_i)}{\partial a_y} & \frac{\partial y_3(t_i)}{\partial a_z} \\ \frac{\partial z_3(t_i)}{\partial a_x} & \frac{\partial z_3(t_i)}{\partial a_y} & \frac{\partial z_3(t_i)}{\partial a_z} \end{bmatrix} = 0_{3 \times 3}$$

(3.82)

onde $O_{3 \times 3}$ é a matriz nula de ordem 3.

vi) Derivada de $z(t)$ em relação ao estado

$$q(t) = (q_1(t), q_2(t), q_3(t), q_4(t)).$$

Antes do cálculo de $\delta h(t)/\delta q(t)$, será rescrita neste item a expressão (3.72) sem o produto matricial. Definido-se,

$$R_{11} = q_1^2 - q_2^2 - q_3^2 + q_4^2$$

$$R_{12} = 2(q_1 q_2 + q_3 q_4)$$

$$R_{13} = 2(q_1 q_3 - q_2 q_4)$$

$$R_{21} = 2(q_1 q_2 - q_3 q_4)$$

$$R_{22} = -q_1^2 + q_2^2 - q_3^2 + q_4^2$$

$$R_{23} = 2(q_2 q_3 + q_1 q_4)$$

$$R_{31} = 2(q_1 q_3 + q_2 q_4)$$

$$R_{32} = 2(q_2 q_3 - q_1 q_4)$$

$$R_{33} = -q_1^2 - q_2^2 + q_3^2 + q_4^2$$

então pode-se escrever $R[q]$ em função dos R_{ij} , isto é,

$$R[q] = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}$$

Tendo feito esta simplificação de notação, pode-se, então, rescrever (3.72) sem o produto matricial de uma maneira compacta.

$$\begin{bmatrix} x_{n1}(t_i) \\ y_{n1}(t_i) \\ z_{n1}(t_i) \\ \cdot \\ \cdot \\ \cdot \\ x_{n3}(t_i) \\ x_{n3}(t_i) \\ x_{n3}(t_i) \end{bmatrix} = \begin{bmatrix} R_{11}\hat{s}_{1x} + R_{12}\hat{s}_{1y} + R_{13}\hat{s}_{1z} + x(t) \\ R_{21}\hat{s}_{1x} + R_{22}\hat{s}_{1y} + R_{23}\hat{s}_{1z} + y(t) \\ R_{31}\hat{s}_{1x} + R_{32}\hat{s}_{1y} + R_{33}\hat{s}_{1z} + z(t) \\ \cdot \\ \cdot \\ \cdot \\ R_{11}\hat{s}_{3x} + R_{12}\hat{s}_{3y} + R_{13}\hat{s}_{3z} + x(t) \\ R_{21}\hat{s}_{3x} + R_{22}\hat{s}_{3y} + R_{23}\hat{s}_{3z} + y(t) \\ R_{31}\hat{s}_{3x} + R_{32}\hat{s}_{3y} + R_{33}\hat{s}_{3z} + z(t) \end{bmatrix} + \begin{bmatrix} n_{x1}(t_i) \\ n_{y1}(t_i) \\ n_{z1}(t_i) \\ \cdot \\ \cdot \\ \cdot \\ n_{x3}(t_i) \\ n_{y3}(t_i) \\ n_{z3}(t_i) \end{bmatrix}$$

(3.83)

Uma vez definida (3.83), com uma modificação em (3.72) pode-se dar continuidade ao cálculo de $H = \frac{\partial h(t_i)}{dq}$.

$$\frac{\partial h_{11j}}{\partial q(t)} \triangleq \begin{bmatrix} \frac{\partial x_1(t)}{\partial q_1(t)} & \frac{\partial x_1(t)}{\partial q_2(t)} & \frac{\partial x_1(t)}{\partial q_3(t)} & \frac{\partial x_1(t)}{\partial q_4(t)} \\ \frac{\partial y_1(t)}{\partial q_1(t)} & \frac{\partial y_1(t)}{\partial q_2(t)} & \frac{\partial y_1(t)}{\partial q_3(t)} & \frac{\partial y_1(t)}{\partial q_4(t)} \\ \frac{\partial z_1(t)}{\partial q_1(t)} & \frac{\partial z_1(t)}{\partial q_2(t)} & \frac{\partial z_1(t)}{\partial q_3(t)} & \frac{\partial z_1(t)}{\partial q_4(t)} \end{bmatrix} = \begin{bmatrix} \frac{\partial R_{x1j}}{\partial q_1(t)} & \frac{\partial R_{x1j}}{\partial q_2(t)} & \frac{\partial R_{x1j}}{\partial q_3(t)} & \frac{\partial R_{x1j}}{\partial q_4(t)} \\ \frac{\partial R_{y1j}}{\partial q_1(t)} & \frac{\partial R_{y1j}}{\partial q_2(t)} & \frac{\partial R_{y1j}}{\partial q_3(t)} & \frac{\partial R_{y1j}}{\partial q_4(t)} \\ \frac{\partial R_{z1j}}{\partial q_1(t)} & \frac{\partial R_{z1j}}{\partial q_2(t)} & \frac{\partial R_{z1j}}{\partial q_3(t)} & \frac{\partial R_{z1j}}{\partial q_4(t)} \end{bmatrix}, j = 1..3$$

(3.84)

onde,

$$\frac{\partial \mathcal{R}_x 1j}{\partial q_1(t)} = 2 \quad q_1(t) \quad \hat{s}_{xj} + 2 \quad q_2(t) \quad \hat{s}_{yj} + 2 \quad q_3(t) \quad \hat{s}_{zj}$$

$$\frac{\partial \mathcal{R}_x 1j}{\partial q_2(t)} = -2 \quad q_1(t) \quad \hat{s}_{xj} + 2 \quad q_1(t) \quad \hat{s}_{yj} - 2 \quad q_4(t) \quad \hat{s}_{zj}$$

$$\frac{\partial \mathcal{R}_x 1j}{\partial q_3(t)} = -2 \quad q_3(t) \quad \hat{s}_{xj} + 2 \quad q_4(t) \quad \hat{s}_{yj} + 2 \quad q_1(t) \quad \hat{s}_{zj}$$

$$\frac{\partial \mathcal{R}_x 1j}{\partial q_4(t)} = 2 \quad q_4(t) \quad \hat{s}_{xj} + 2 \quad q_3(t) \quad \hat{s}_{yj} - 2 \quad q_1(t) \quad \hat{s}_{zj}$$

$$\frac{\partial \mathcal{R}_y 1j}{\partial q_1(t)} = 2 \quad q_2(t) \quad \hat{s}_{xj} - 2 \quad q_2(t) \quad \hat{s}_{yj} + 2 \quad q_4(t) \quad \hat{s}_{zj}$$

$$\frac{\partial \mathcal{R}_y 1j}{\partial q_2(t)} = 2 \quad q_1(t) \quad \hat{s}_{xj} + 2 \quad q_2(t) \quad \hat{s}_{yj} + 2 \quad q_3(t) \quad \hat{s}_{zj}$$

$$\frac{\partial \mathcal{R}_y 1j}{\partial q_3(t)} = -2 \quad q_4(t) \quad \hat{s}_{xj} - 2 \quad q_3(t) \quad \hat{s}_{yj} + 2 \quad q_2(t) \quad \hat{s}_{zj}$$

$$\frac{\partial \mathcal{R}_y 1j}{\partial q_4(t)} = -2 \quad q_3(t) \quad \hat{s}_{xj} + 2 \quad q_4(t) \quad \hat{s}_{yj} + 2 \quad q_1(t) \quad \hat{s}_{zj}$$

$$\frac{\partial \mathcal{R}_z 1j}{\partial q_1(t)} = 2 \quad q_3(t) \quad \hat{s}_{xj} - 2 \quad q_4(t) \quad \hat{s}_{yj} - 2 \quad q_1(t) \quad \hat{s}_{zj}$$

$$\frac{\partial \mathcal{R}_z 1j}{\partial q_2(t)} = 2 \quad q_4(t) \quad \hat{s}_{xj} + 2 \quad q_3(t) \quad \hat{s}_{yj} - 2 \quad q_2(t) \quad \hat{s}_{zj}$$

$$\frac{\partial \mathcal{R}_z 1j}{\partial q_3(t)} = 2 \begin{bmatrix} q_1(t) & \hat{s}_{xy} + 2 & q_2(t) & \hat{s}_{xy} + 2 & q_3(t) & \hat{s}_{zj} \end{bmatrix}$$

$$\frac{\partial \mathcal{R}_z 1j}{\partial q_4(t)} = 2 \begin{bmatrix} q_1(t) & \hat{s}_{xy} - 2 & q_1(t) & \hat{s}_{xy} + 2 & q_4(t) & \hat{s}_{zj} \end{bmatrix}$$

v) Derivada de $h(t)$ em relação ao estado $w = (w_x, w_y, w_z)$

$$\frac{\partial h_{11}}{\partial v(t)} \stackrel{\Delta}{=} \begin{bmatrix} \frac{\partial x_1(t_i)}{\partial v_x(t)} & \frac{\partial x_1(t_i)}{\partial v_y(t)} & \frac{\partial x_1(t_i)}{\partial v_z(t)} \\ \frac{\partial y_1(t_i)}{\partial v_x(t)} & \frac{\partial y_1(t_i)}{\partial v_y(t)} & \frac{\partial y_1(t_i)}{\partial v_z(t)} \\ \frac{\partial z_1(t_i)}{\partial v_x(t)} & \frac{\partial z_1(t_i)}{\partial v_y(t)} & \frac{\partial z_1(t_i)}{\partial v_z(t)} \end{bmatrix} = 0_{3 \times 3}$$

(3.85)

$$\frac{\partial h_{21}}{\partial v(t)} \stackrel{\Delta}{=} \begin{bmatrix} \frac{\partial x_2(t_i)}{\partial v_x(t)} & \frac{\partial x_2(t_i)}{\partial v_y(t)} & \frac{\partial x_2(t_i)}{\partial v_z(t)} \\ \frac{\partial y_2(t_i)}{\partial v_x(t)} & \frac{\partial y_2(t_i)}{\partial v_y(t)} & \frac{\partial y_2(t_i)}{\partial v_z(t)} \\ \frac{\partial z_2(t_i)}{\partial v_x(t)} & \frac{\partial z_2(t_i)}{\partial v_y(t)} & \frac{\partial z_2(t_i)}{\partial v_z(t)} \end{bmatrix} = 0_{3 \times 3}$$

(3.86)

$$\frac{\partial h_{31}}{\partial v(t)} \stackrel{\Delta}{=} \begin{bmatrix} \frac{\partial x_3(t_i)}{\partial v_x(t)} & \frac{\partial x_3(t_i)}{\partial v_y(t)} & \frac{\partial x_3(t_i)}{\partial v_z(t)} \\ \frac{\partial y_3(t_i)}{\partial v_x(t)} & \frac{\partial y_3(t_i)}{\partial v_y(t)} & \frac{\partial y_3(t_i)}{\partial v_z(t)} \\ \frac{\partial z_3(t_i)}{\partial v_x(t)} & \frac{\partial z_3(t_i)}{\partial v_y(t)} & \frac{\partial z_3(t_i)}{\partial v_z(t)} \end{bmatrix} = 0_{3 \times 3}$$

(3.87)

vi) Finalmente, após a execução dos itens i a v , pode-se definir a derivada de $h(t)$ em relação ao estado $x(t)$. Isto pode ser feito empregando-se as expressões (3.73) a (3.87). Então,

$$H = \frac{\partial h}{\partial x(t)} = \begin{bmatrix} I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & \frac{\partial h_{11}}{\partial q(t)} & 0_{3 \times 3} \\ I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & \frac{\partial h_{12}}{\partial q(t)} & 0_{3 \times 3} \\ I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & \frac{\partial h_{13}}{\partial q(t)} & 0_{3 \times 3} \end{bmatrix} \quad (3.88)$$

A expressão (3.88) é a aproximação linear de $h[x(t)]$ no ponto de operação $x(t)$, conforme o critério de linearização de sistemas mencionados na Seção III.3.4. Com esta expressão e a matriz de transição (3.69), pode-se substituir o sistema contínuo de equações dinâmicas (3.56) por um sistema discreto capaz de representar o processo em estudo. Assim, tem-se:

$$\begin{cases} x(k+1) = \phi(k; x = x_{op}) \quad x(k) + G \quad w(k) \\ z(k) = H(k; x = x_{op}) \quad x(k) + v(k) \end{cases} \quad (3.89)$$

onde k é o indexador no tempo dos estados $x(k)$ ou das medidas $z(k)$ e $x = x_{op}$ é o ponto de operação do sistema que neste trabalho será as estimativas feitas pelo IEKF no passo anterior ao indexador temporal k .

III.3.5 Equações do IEKF

Na implementação das equações do IEKF [38], para se estimar os estados do sistema (3.3.1.7) devem ser usadas as aproximações lineares F e H desenvolvidas nas seções anteriores, nos respectivos pontos de operação. Este procedimento é seguido

devido à natureza não-linear dessas equações. O ponto de operação dessas aproximações será a estimativa feita $x(k)$ imediatamente após uma medida ter sido feita, conforme mencionado anteriormente. O conjunto de equações que caracteriza o IEKF, segundo a referência [38], pode ser dividido em dois grupos. Um deles faz a propagação do estado estimado $x(k)$ e da matriz de covariância $P(k)$. O outro as devidas atualizações em $P(k)$, no ganho do filtro de Kalman $K(k)$ e no estado estimado.

O IEKF, de acordo com os comentários feitos nas referências [1] e [38], é um estimador recursivo, que antes de fazer a atualização na matriz de covariância $P(k)$ faz inúmeras iterações (com quantidade previamente definida) no cálculo do ganho $K(k)$ e da estimativa de $x(k)$. Essas iterações, segundo as referências [1] e [38] melhoram o desempenho do filtro diante da não-linearidade existente na equação de medição.

A seguir, as equações do IEKF, conforme as referências [1] e [38].

A primeira a ser considerada é a equação de atualização dos estados estimados. Assim, tem-se:

(3.90)

$$\hat{x}(k/k)_{n+1} = \hat{x}(k/k-1) + K(k)_n [z(k) - h(x_{op}) - H(x_{op})(\hat{x}(k/k-1) - \hat{x}(k/k)_n)]$$

onde,

$$x_{op} = \hat{x}(k/k)_n;$$

$$n;$$

$$K(k)_n$$

são respectivamente o ponto de operação, o índice de iteração local e o ganho local. A variável k está relacionado com a evolução no tempo do filtro. A iteração começa com:

$$\hat{x}(k/k)_0 = \hat{x}(k/k-1)$$

O ganho de iteração local $K(k)_n$ é definido como,

$$K(k)_{n+1} = P(k/k-1)H^T(x_{op})[H(x_{op})P(k/k-1)H^T(x_{op}) + R_k]^{-1} \quad (3.91)$$

onde R_k é a matriz de covariância do ruído de medição $v(k)$.

Finalmente, após não mais haver melhoramento no valor de $x(k/k)_n$ em (3.90), isto é, quando as estimativas de $x(k/k)$ avaliadas nos passos $n+1$ e n diferirem por um número ϵ ($\epsilon \cong 0$), após sucessivas iterações, o que segundo a referência [4] ocorre em 3 ou 5 iterações, pode-se fazer a atualização da matriz de covariância $P(k)$. Então, tem-se:

$$P(k/k) = [I - K(k)_n H(x_{op})]P(k/k-1) \quad (3.92)$$

onde são usados os últimos valores de $K(k)$ e $H(x_{op})$ das iterações.

A estimativa de $x(k)$ e a matriz de covariância $P(k)$ são propagadas pelas seguintes equações:

$$\hat{x}(k+1/k) = F[\hat{x}(k/k)]\hat{x}(k/k) \quad (3.93)$$

$$P(k/k-1) = F[\hat{x}(k/k)]P(k-1/k-1)F^T[\hat{x}(k/k)] + G(k)Q(k)G^T(k) \quad (3.94)$$

onde $Q(t)$ é a matriz de covariância do ruído de estado existente no modelo do processo.

O conjunto formado pelas expressões (3.90) a (3.92), (3.93) e (3.95) constituem as equações do IEKF.

Neste capítulo foram analisados os modelos necessários à implementação do estimador de estado (IEKF), quais sejam, modelos representativo do processo em questão (modelos dos movimentos de rotação e translação e o modelo da estrutura do objeto), modelos de medição de pontos característicos de imagens (observador de

pontos ruidosos) e o modelo de projeção de imagens. Ainda faltam examinar o processo de aquisição de imagens, a modelagem das ferramentas de segmentação de imagens, bem como a definição de um modelo que represente as influências das fontes de ruído no processo de digitalização de imagens no experimento com imagens reais. O seguinte algoritmo resume os passos do procedimento de estimação.

<início>

quadro = 0;

while quadro < 50

{

simulação de movimento do objeto de teste;

aquisição de imagens (feita pelas câmeras direita e esquerda);

execução de modelos de segmentação de imagens;

execução de modelos de movimento;

estimação da forma do objeto;

estimação de estados;

avaliação do desvio das estimativa em relação aos valores reais;

}

<fim>

CAPÍTULO IV

SISTEMA DE AQUISIÇÃO, DIGITALIZAÇÃO E PROCESSAMENTO DE IMAGENS

IV.1 Comentários iniciais

Neste capítulo será feita a descrição detalhada do Sistema de Aquisição e Digitalização de Imagens (SADI) empregado para adquirir as informações contidas nas seqüências de imagens estéreo para serem usadas na estimação de parâmetros de movimento. Também é objeto deste capítulo o estudo dos métodos de segmentação de imagens, ou seja, métodos de detecção de cantos. Em um deles, o modelo de detecção é obtido usando-se a definição de gradiente associado a uma imagem. No outro, a modelagem do detetor de cantos é feita a partir da resolução de um problema de otimização, resolvido empregando-se cálculo variacional. Será ainda examinado neste capítulo as diferentes fontes de ruído tais como ruído de digitalização, erro na detecção de cantos, etc., no experimento com imagens reais e as influências dessas fontes no procedimento de estimação paramétrica realizado, bem como a modelagem de duas técnicas de verificação da correspondência entre imagens adquiridas por câmeras distintas (imagens estéreo). Uma dessas técnicas manipula as imagens originalmente adquiridas, considerando todos os atributos existentes, enquanto que a outra necessita da prévia segmentação das imagens para a partir daí, considerar somente certos atributos – cantos de imagens.

IV.2 Montagem do SADI

O SADI é uma parte extremamente significativa para no procedimento de estimação paramétrica, pois é a partir dele que os dados (imagens digitalizadas) são adquiridos e disponibilizados para os modelos e estimador escolhido. Portanto, o processo de montagem do SADI deve ser realizado de maneira bastante criteriosa, isto é, utilização de: suportes apropriados e de fácil manuseio; computadores equipados com *software* (sistema operacional, linguagem de programação e respectivo compilador) e *hardware* (configuração da CPU e placa digitalizadora de imagem) que possibilitem a realização do experimento com imagem real; e, obviamente, duas câmeras, pois pretende-se adquirir seqüências de imagens estéreo.

O SADI foi montado em uma mesa de teste. Ele é composto de duas câmeras, uma placa digitalizadora de imagem e um micro-computador 386. As imagens adquiridas pelas câmeras são digitalizadas com uma resolução de 480x512 pixels na placa DT-2851. As imagens digitalizadas são, então, processadas em um algoritmo configurado no micro-computador. Este algoritmo, implementado em linguagem de programação "C", faz o gerenciamento do processo de aquisição de imagens, fazendo a leitura e o armazenamento das imagens gravadas, utilizando o *software* básico da placa digitalizadora de imagens.

A seguir são resumidos os componentes do SADI, conforme mostrado na Figura 4.1:

- a) Suporte para câmera;
- b) Câmera Sony HVM - 302;
- c) Placa digitalizadora DT- 2851
- d) Um micro computador 386;
- e) Posição da câmera no eixo x 55cm;
- f) Posição da câmera no eixo y: 0cm;
- g) Posição da câmera no eixo z: 30cm;
- h) Ângulo de visada da câmera: - 30 com eixo z

A Figura 4.1 mostra a montagem utilizada na aquisição de imagens.

Fig. 4.1 - Montagem do sistema de aquisição de imagens.

IV.3 Primeiro método de detecção de cantos

Nesta seção será examinado o primeiro dos métodos de detecção de cantos (e respectivo algoritmo) considerados neste trabalho. O algoritmo relativo ao método que será estudado nesta seção não necessita de informações acerca das características geométricas do objeto, por utilizar máscaras de convolução vertical e horizontal (operadores de segmentação) no processo de detecção.

Um canto pode ser definido como o lugar geométrico definido pela interseção de duas ou mais linhas (bordas) retas. Eles são de grande utilidade no cálculo do fluxo óptico e de estimativas da dinâmica dos elementos contidos em uma cena. Nesta seção, será apresentado o algoritmo detetor de cantos, proposto em [10], que usa um modelo matemático para um canto. Outros algoritmos na mesma linha, podem ser encontrado em Canny [40], Zuniga e Haralick [41], Kitechen e Rosenfeld [9], Dreschler e Nagel [42].

O processo de detecção de cantos consiste na segmentação da imagem em regiões e representação da fronteira do objeto como uma cadeia de código (borda). Os cantos são identificados como sendo a região onde a direção da borda muda rapidamente. A taxa de mudança da direção é representada pelo gradiente. Sendo assim, cantos são pontos onde a taxa de mudança da direção do gradiente é máxima. Uma breve descrição desse processo será feita a seguir.

A partir de uma seqüência de imagens estéreo, um par de quadros é selecionado e o correspondente par de imagens. Desse par, uma imagem é destacada. Máscaras de Sobel 3x3 [43]-[45] são usadas no cálculo dos componentes vertical e horizontal do gradiente da imagem, ou seja, Grad_y e Grad_x , respectivamente. A direção do gradiente é calculada através do inverso da função tangente da razão entre os componentes vertical e horizontal. O módulo associado a esses componentes é calculado. O produto entre o módulo da direção do gradiente e o módulo do gradiente da imagem produz um valor que, após comparado com um limiar, define o canto a ser localizado na imagem original. Esse mesmo procedimento é aplicado, em seguida, na outra imagem do par.

IV.3.1 Descrição do algoritmo

O presente algoritmo está organizado em três estágios: Inicialmente, ele identifica as bordas do(s) objeto(s) contido(s) na imagem, calcula a direção do gradiente e a taxa de variação ocorrida e, finalmente, aplica um limiar para a taxa de mudança do ângulo do gradiente, identificando, dessa forma, os cantos. A qualidade da estimação está relacionada com os erros (numéricos) ocorridos em cada estágio.

Identificação de borda

Neste estágio do algoritmo é feita uma identificação das bordas horizontais e verticais do objeto. Esta identificação é feita através da operação de convolução da imagem com máscaras de detecção horizontal e vertical. Essa operação caracteriza o cálculo do gradiente nas direções x e y . As máscaras utilizadas neste algoritmo são matrizes 3×3 , cujos elementos são selecionados em função da identificação que se pretende realizar. É interessante selecionar números que caracterizem a transição existente, em termos de níveis de cinza, entre o objeto e o fundo da cena. Como o presente objetivo é individualizar as fronteiras do objeto em relação aos outros elementos contidos na cena, isto é, suas bordas, três desses elementos, conforme as expressões (4.1) e (4.2), são zeros e os elementos à direita (acima) e à esquerda (abaixo) são iguais, mas com sinais opostos.

$$S_H = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} \quad (4.1)$$

$$S_V = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad (4.2)$$

Analiticamente, pode-se escrever:

$$Grad_x I = I * S_H \quad (4.3)$$

$$Grad_y I = I * S_V \quad (4.4)$$

onde I é a intensidade da imagem do objeto e $Grad_x I$ e $Grad_y I$ são respectivamente os componentes vertical e horizontal do gradiente da imagem.

Cálculo dos módulos dos gradientes da imagem e da direção da borda identificada

O mapeamento 2-D do módulo do gradiente da imagem é obtido calculando-se a raiz quadrada da soma dos quadrados dos componentes vertical e horizontal. O gradiente da direção da borda identificada é obtido calculando-se a tangente inversa da razão entre os componentes vertical e horizontal do gradiente da imagem. Para i no intervalo $[1, m]$ e j em $[1, n]$, com m e n definindo as dimensões da imagem, pode-se obter:

$$g_m(i, j) = \sqrt{(Grad_x I(i, j))^2 + (Grad_y I(i, j))^2} \quad (4.5)$$

$$\varphi(i, j) = tg^{-1} \frac{Grad_y I(i, j)}{Grad_x I(i, j)} \quad (4.6)$$

Definindo

$$Grad_x A = \varphi S_H \quad (4.7)$$

$$Grad_y A = \varphi S_V \quad (4.8)$$

pode se obter

$$g_{dm}(i, j) = \sqrt{(Grad_x A(i, j))^2 + (Grad_y A(i, j))^2} \quad (4.9)$$

onde $g_m(.,.)$ e $g_{dm}(.,.)$ denotam os módulos do gradiente e da direção do gradiente, respectivamente.

Detetor de limiar da taxa de variação do ângulo do gradiente

O limiar da taxa de variação do ângulo do gradiente utilizado neste algoritmo é um número V_r , previamente definido, que deverá ser menor que o produto de (4.9) por (4.5). O canto é localizado na imagem na posição onde for encontrado o limiar. Analiticamente:

$$V(i, j) = \sqrt{[(Grad_x I(i, j))^2 + (Grad_y I(i, j))^2][(Grad_x A(i, j))^2 + (Grad_y A(i, j))^2]} \quad (4.10)$$

Limiar = $V(i, j)$, se $V < V_r$

IV.4 Segundo método de detecção de cantos

Dando continuidade à análise dos métodos de detecção de cantos considerados neste trabalho, passa-se a examinar nesta seção os fundamentos e o algoritmo do segundo detetor de cantos referenciado na Seção IV.1. Aqui, o modelo de detecção de cantos será tratado como um problema de otimização, resolvido usando-se cálculo variacional. A resolução do problema de otimização tem como consequência duas equações diferenciais. A resolução dessa equação será feita na Seção IV.4.2. Também nesta seção, será examinado um exemplo em que as constantes desconhecidas das referidas equações são definidas e, conseqüentemente, o modelo de detecção, determinado.

IV.4.1 Formulação do problema de otimização

Seja $g(x, y)$ o operador a ser determinado e $n(x, y)$ o ruído branco Gaussiano. Então, a função que considera uma imagem contaminada por ruído é:

$$F(x, y) = I(x, y) + n(x, y)$$

O canto será determinado, primeiro fazendo a convolução de $F(x, y)$ com $g(x, y)$, uma função desconhecida, chamada de operador de canto, e em seguida calculando o resultado máximo dessa operação, ou seja, determinando o valor máximo de $M(x, y) = F(x, y) * g(x, y)$, onde $*$ é o operador de convolução.

A função $g(x, y)$ deve satisfazer alguns critérios, quais sejam:

- 1 - O operador não deve ser sensível ao ruído;
- 2 - O canto detectado pode, também, ser um ponto de borda;
- 3 - O canto detectado deve ter, pelo menos, duas vizinhanças que têm gradiente de direção diferente do seu.

O objetivo é encontrar um detector ótimo que satisfaça os critérios mencionados acima. Para tanto, serão utilizadas técnicas de otimização para encontrar a expressão de $g(x, y)$, tal que o valor de $M(x, y)$ seja máximo. Segundo [10], isto é feito, determinando-se Z , razão sinal ruído, e Λ , desempenho do detector. As expressões de Z e Λ são mostradas a seguir.

$$Z = \frac{A \int_0^{+\infty} \int_{-LIM}^{LIM} g(x, y) dx dy}{\sigma_n^2 \sqrt{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g^2(x, y) dx dy}} \quad (4.11)$$

$$\Lambda = \frac{\sigma_n^2 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g_x^2(x, y) dy dx}{\left(A \int_0^{+\infty} \int_{-LIM}^{LIM} g_{xx}^2(x, y) dy dx \right)^2} + \frac{\sigma_n^2 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g_y^2(x, y) dy dx}{\left(A \int_0^{+\infty} \int_{-LIM}^{LIM} g_{yy}^2(x, y) dy dx \right)^2} \quad (4.12)$$

onde σ_n^2 é a variância do ruído $n(x, y)$ e $g(x, y)$ o operador a ser determinado, g_x , g_y , g_{xx} , g_{yy} são derivadas parciais de g , A é o contraste do canto (admitido ser constante neste trabalho) e LIM é o limite de integração da borda do objeto. Para o detector de

canto ideal o quociente entre Z e Λ , representado por Σ , deve ser máximo. Então, tem-se:

$$\Sigma = \frac{\frac{A \int_{-\infty-LIM}^{+\infty+LIM} \int_{-\infty-LIM}^{+\infty+LIM} g(x,y) dy dx}{\sigma_n^2 \sqrt{\int_{-\infty-\infty}^{+\infty+\infty} \int_{-\infty-\infty}^{+\infty+\infty} g^2(x,y) dy dx}}{\frac{\sigma_n^2 \int_{-\infty-\infty}^{+\infty+\infty} \int_{-\infty-\infty}^{+\infty+\infty} g_x^2(x,y) dy dx}{(A \int_{-\infty-LIM}^{+\infty+LIM} \int_{-\infty-LIM}^{+\infty+LIM} g_{xx}^2(x,y) dy dx)^2} + \frac{\sigma_n^2 \int_{-\infty-\infty}^{+\infty+\infty} \int_{-\infty-\infty}^{+\infty+\infty} g_y^2(x,y) dy dx}{(A \int_{-\infty-LIM}^{+\infty+LIM} \int_{-\infty-LIM}^{+\infty+LIM} g_{yy}^2(x,y) dy dx)^2}} \quad (4.13)$$

IV.4.2 Resolução do problema de otimização a partir do modelo gráfico de um canto

Nesta seção será dada continuidade à resolução do problema de otimização iniciado na seção anterior. Para tanto, admite-se que um canto é formado pela interseção de duas retas com ângulos de inclinação conforme mostrado na Figura 4.2. Nesta figura, m é a inclinação positiva da reta e $-m$ a inclinação negativa. A função que descreve os níveis de cinza ao redor do canto é:

$$I(x,y) = \begin{cases} A & \text{se } x > 0 \text{ e } -mx < y < +mx \\ 0 & \text{caso contrário} \end{cases}$$

Fig 4.2 - Modelo gráfico de um canto.

O próximo passo é usar o cálculo variacional para maximizar uma das integrais (4.13), mantidas as outras constantes. Entretanto, devido ao fato das integrais terem limites diferentes, a manipulação algébrica torna-se bastante difícil. Há a necessidade, portanto, de se levar todos os limites de integração de (4.13) ao infinito para simplificar os cálculos. Uma outra simplificação é definir $H(x, y)$ como segue:

$$H(x, y) = \begin{cases} 1 & \text{se } x > 0 \text{ e } -mx < y < mx \\ 0 & \text{caso contrário} \end{cases}$$

Então, com a consideração feita, os limites de integração podem ser levados ao infinito, permitindo escrever:

$$A \int_0^{+\infty} \int_{-LM}^{+LM} k(x, y) dy dx = A \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} k(x, y) g(x, y) dy dx =$$

e a expressão a ser maximizada é então:

$$\Sigma = \frac{A \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) H(x, y) dy dx}{\sigma_n^2 \sqrt{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g^2(x, y) dy dx}} \\ + \frac{\frac{\sigma_n^2 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g_x^2(x, y) dy dx}{(A \int_0^{+\infty} \int_{-LM}^{+LM} g_{xx}^2(x, y) dy dx)^2} + \frac{\sigma_n^2 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g_y^2(x, y) dy dx}{(A \int_0^{+\infty} \int_{-LM}^{+LM} g_{yy}^2(x, y) dy dx)^2}}{2}$$

(4.14)

Resolvendo (4.14), utilizando os métodos proposto em [45], obtém-se o seguinte sistema de equações diferenciais:

$$g(x, y) = c_1 \operatorname{sen}\left(\frac{m\pi x}{w}\right) \left[- (e^{zw} + e^{-zw}) + e^{zy} + e^{-zy} \right] \quad (4.15)$$

$$g(x, y) = c_2 \operatorname{sen}\left(\frac{n_1\pi x}{w}\right) \operatorname{sen}\left(\frac{n_2\pi y}{w}\right) \quad (4.16)$$

onde W é a dimensão da máscara de convolução e c_1, c_2, m, n_1, n_2 e z são constantes. (4.15) e (4.16) são aplicáveis, respectivamente na região (I) e (II).

A partir de considerações feitas em [10], tem-se que a máscara ótima para detecção do canto da Figura 4.2 é a mostrada a seguir:

$$W = \begin{bmatrix} -3 & -5 & -5 & -3 & 0 & -3 & -5 & -5 & -3 \\ -5 & -9 & -9 & -5 & 0 & -5 & -9 & -9 & -5 \\ -5 & -9 & -9 & -5 & 0 & -5 & -9 & 8 & 5 \\ -3 & -5 & -5 & -3 & 0 & 6 & 9 & 9 & 6 \\ 0 & 0 & 0 & 0 & 0 & 6 & 10 & 10 & 6 \\ -3 & -5 & -5 & -3 & 0 & 6 & 9 & 9 & 6 \\ -5 & -9 & -9 & -5 & 0 & -5 & -9 & 8 & 5 \\ -5 & -9 & -9 & -5 & 0 & -5 & -9 & -9 & -5 \\ -3 & -5 & -5 & -3 & 0 & -3 & -5 & -5 & -3 \end{bmatrix}$$

(4.17)

Percebe-se nesta máscara que a transição nos elementos (mudança de sinal) define uma figura geométrica bastante semelhante à região II da Figura 4.2.

IV.4.3 Generalização do modelo

O modelo estudado na seção anterior é bastante simples e limitado. Ele assume que cantos são a resultantes de intercessões de retas com o eixo x . Isto significa que a direção do canto é fixa. Mas, sabemos que em uma imagem real, um canto pode aparecer com orientação arbitrária. Uma outra limitação do modelo da Seção IV.4.2 é o fato da configuração da máscara de convolução depender da forma geométrica do objeto.

É um trabalho difícil enumerar todas as possibilidades de ocorrência de cantos e gerar máscaras de detecção para cada um deles. Contudo, em [10] é dada uma solução para esse problema. A solução consiste em usar 12 sub máscaras de detecção, configuradas conjuntamente. Esta abordagem, conhecida como detetor ótimo de canto, é capaz de detectar cantos com orientação arbitrária.

As Seções IV.4.2 e IV.4.3 trataram dos dois métodos de detecção de cantos utilizados neste trabalho. No Capítulo V serão analisados os resultados obtidos com esses métodos e as influências dos mesmos na estimação de parâmetros.

IV.5 Técnicas de verificação da correspondência entre imagens adquiridas por câmeras distintas

Na recuperação de informações tridimensionais contidas em uma cena filmada por um par de câmeras (visão estereoscópica), inevitavelmente dois problemas têm que ser resolvidos. Um deles é a determinação das coordenadas tridimensionais da posição dos pontos contidos na cena a partir das imagens adquiridas e projetadas em cada câmera (problema já resolvido na Seção III.2.2). O outro é mais complexo. Trata-se do problema da verificação da correspondência existente entre aquelas imagens. Tal problema pode ser enunciado da seguinte forma: Considerando um ponto imagem em uma câmera, como o ponto imagem conjugado, na outra câmera pode ser localizado?

Nesta seção serão apresentadas duas técnicas de localização de pontos conjugados que respondem a questão formulada anteriormente, resolvendo o problema da correspondência existente entre um par de imagens estéreo. A primeira é

fundamentada na manipulação de todas as informações existentes nas imagens originais, enquanto que a segunda manipula atributos (pontos salientes, como bordas ou cantos), requerendo, portanto, a prévia segmentação das imagens originais.

IV.5.1 Primeiro método de verificação da correspondência

O método que será apresentado nesta seção pertence a uma classe de técnicas de Visão Computacional denominada Técnica Estéreo Baseada em Área, [43] - [44] e [46]. Neste método o problema da correspondência é resolvido a partir do cálculo da similaridade entre regiões correspondentes entre as imagens direita e esquerda. Nessas regiões, pontos são selecionados com base na verificação do nível de cinza em setores vizinhos. A similaridade é determinada calculando-se o somatório das diferenças (D) entre as referidas regiões, conforme a expressão mostrada a seguir.

$$D = \sum_{j=1}^n \sum_{i=1}^m [I_L(i, j) - I_R(i, j)]^2 \quad (4.18)$$

onde $I_L(i, j)$ é um ponto imagem ou pixel na imagem da câmera esquerda e $I_R(i, j)$ é o pixel conjugado na imagem da câmera da direita. As constantes m e n definem as dimensões das regiões consideradas, denominada aqui de rc .

Os valores de D são determinados sequencialmente, empregando-se (4.18), em toda rc de cada câmera. A condição de aceitação ou rejeição da similaridade é caracterizada, definindo-se L (valor limite) que tem a função de julgar se os valores de D são maiores ou menores que o valor limite. Durante a avaliação de D nas regiões, o valor de L pode ser fixo ou variar em alguns setores das regiões. Caso ocorra essa última condição, o valor médio de L (L_m), é considerado como valor limite. Em resumo, a similaridade é aceita ou rejeitada de acordo com as seguintes condições:

se $D(rc) < L(rc)$
então
a similaridade é aceita

caso contrário

a similaridade não é aceita

fim do teste

Tendo-se determinado a característica da similaridade (aceita ou rejeitada), tem-se resolvido o problema da correspondência empregando-se o presente método, isto é, as imagens serão correspondentes se a similaridade é aceita. No caso de rejeição da similaridade a correspondência não está assegurada.

Será considerado a seguir, dando seguimento à análise objeto deste item, a sensibilidade ao efeito de contrastes e a utilização de câmeras com características distintas influenciam no cálculo da similaridade.

Tanto o efeito de contrastes e a utilização de câmeras diferentes alteram a expressão original do cálculo de D e conseqüentemente a determinação da similaridade, pois quando da ocorrência de um ou de outro, ([46]), a variância da intensidade da luz σ^2 e a média da imagem μ em rc passam a ser termos de (4.18). As expressões de σ_k^2 (com $k = L$ - câmera esquerda ou R - câmera direita) são mostradas a seguir.

$$\sigma_L^2 = \frac{1}{mn} \sum_{j=1}^n \sum_{i=1}^m [I_L(i, j) - \mu_L]^2 \quad (4.19)$$

$$\sigma_R^2 = \frac{1}{mn} \sum_{j=1}^n \sum_{i=1}^m [I_R(i, j) - \mu_R]^2 \quad (4.20)$$

μ_k , com $k = L$ ou R (respectivamente para as câmeras esquerda e direita), é determinado com segue:

$$\mu_k = \frac{1}{nm} \sum_{j=1}^n \sum_{i=1}^m I_k(i, j) \quad (4.21)$$

São mostradas a seguir as expressões de D quando da ocorrência do efeito de contraste nas imagens consideradas. Na expressão, em que D foi substituído por D_{ec} para fazer alusão ao referido efeito (ec), as variâncias σ_L^2 e σ_R^2 , calculadas por (4.19) e (4.20) respectivamente, são os elementos normalizadores do D original.

$$D_{ec} = \frac{1}{\sigma_L \sigma_R} \sum_{j=1}^n \sum_{i=1}^m [I_L(i, j) - I_R(i, j)]^2 \quad (4.22)$$

Se na aquisição das imagens estéreo for utilizada câmeras com características de iluminação distintas, isto é, ocorrência de intensidades diferentes de luz para um mesmo ponto, o cálculo da similaridade é feito pela expressão mostrada a seguir, que foi obtida da expressão original de determinação de D pela inclusão das variâncias σ_L^2 e σ_R^2 ((4.19) e (4.2)) e das médias μ_k ($k = L$ ou R) em cada r_c .

$$D_{ec} = \sum_{j=1}^n \sum_{i=1}^m \left[\frac{1}{\sigma_L} (I_L(i, j) - \mu_L) - \frac{1}{\sigma_R} (I_R(i, j) - \mu_R) \right]^2 \quad (4.23)$$

Nesta seção foi descrito um método de resolução do problema da correspondência entre um par de imagens estéreo proposto em [46]. O método tratado aqui considera todas as informações contidas nas imagens conjugadas (pertence à classe Estereoscopia Baseada em Área) a partir da delimitação de regiões de exploração, onde sequencialmente são avaliadas os valores de D (ou (4.18) ou (4.22) ou (4.23)), elemento que permite medir as similaridades nas referidas regiões, calculando as diferenças das intensidade dessas imagens. A correspondência entre as imagens é,

imagens é, então, verificada a partir das similaridade, conforme descrição feita na presente subseção. O que foi dito neste parágrafo pode, ainda, ser resumido no seguinte algoritmo:

- 1- Definição das constantes m e n ;
- 2- Início da verificação da correspondência;
- 3- Delimitação da região de pesquisa da correspondência;
- 4- Teste da sensibilidade do efeito de contrastes e câmeras diferentes;
 - caso as imagens não sejam sensíveis ao efeito de contrastes: execução seqüencial do cálculo de D utilizando a expressão (4.18);
 - caso as imagens sejam sensíveis ao efeito de contrastes: execução seqüencial do cálculo de D utilizando a expressão (4.22);
 - caso a intensidade de luz para um mesmo ponto seja diferente: execução seqüencial do cálculo de D utilizando a expressão (4.23);
- 5- Testar as similaridades das imagens;
- 6- Verificar a correspondência;
- 7 - Voltar ao passo 1;

Na próxima subseção será examinado o outro método de resolução do problema da correspondência entre imagens estéreo considerado neste trabalho e proposto em [44] - [46].

IV.5.2 Segundo método de verificação da correspondência

O método de verificação da correspondência que será descrito nesta subseção pertence a uma classe de métodos aplicáveis em Visão Computacional (VC) denominada: Estereoscopia Baseada em Atributos. A principal característica do método em questão é a não consideração de todas as informações existentes na imagem mas sim, parte delas que são suficientes à análise de um determinado processo. Essas partes recebem em VC o nome de atributos, que são exatamente os aspectos salientes existentes na cena, as bordas dos elementos de interesse ou os pontos de canto. Essa

característica faz com que o método seja mais rápido na verificação da correspondência que o analisado na Seção IV.5.1 (baseado em área), pois possui menos pontos conjugados passíveis de serem correspondentes.

Por considerar somente os atributos contidos nas imagens, o método requer a prévia segmentação da imagem para que o mesmo seja implementado. As características deste método são:

- A correspondência é menos sensível a variações fotométricas dos imageamentos, pois atributos representam propriedades específicas da cena;
- Os pares de correspondência somente são encontrados se ambas as atributos forem extraídas de ambas as imagens;
- É mais rápido que o método baseado em área por necessitar de menos pontos (pixel) candidatos.

A descrição matemática deste método é feita utilizando-se basicamente as mesmas expressões do método examinado na seção anterior. Contudo, a execução de seu algoritmo é mais rápida por razões já mencionadas nos parágrafos anteriores. Um algoritmo para este método é apresentado a seguir:

- 1- Início da verificação da correspondência;
- 2- Segmentação das imagens (Seção IV.3.1.1);
 - obtenção dos componentes verticais através do operador (4.1);
 - obtenção dos componentes horizontais através do operador (4.2);
- 3- Extração das atributos necessárias a individualização do objeto em relação ao resto da cena;
- 4- Execução seqüencial do cálculo de D (4.18) a partir dos atributos extraídos em 4;
- 5- Verificação das similaridades;
- 6- Verificação da correspondência;
- 7- Volta ao passo 1.

IV.6 Erro de digitalização no experimento com imagens reais

Nesta seção são examinadas as influências das diferentes fontes de ruído no experimento com imagens reais, fato intimamente ligado com erros na detecção de cantos, ruído de digitalização, e outras fontes de incertezas relacionadas com o processo de aquisição, processamento e segmentação de imagens. Esta abordagem possibilita lidar com a ocorrência de imagens ruidosas. Nos experimentos com imagem real é bastante difícil a modelagem das várias fontes de incerteza nos sistemas de imageamentos. Os efeitos dessas fontes no processo de medição são caracterizados aqui, segundo [4], como ruído de medição associado à resolução das câmeras empregadas no experimento.

Na formulação do sistema de equações dinâmicas (3.56) que representa a situação abordada (movimento de um objeto no espaço 3-D), o ruído é considerado Gaussiano, porém, como os erros de detecção dos cantos podem ser modelados como ruído de digitalização, o ruído é considerado distribuído uniformemente, ao invés de apresentar uma distribuição Gaussiana. Neste caso, a solução do sistema (3.56) passa a apresentar a ser ótima dentre as soluções lineares. As estatísticas do ruído (média e variância) são determinadas, admitindo que as mesmas não dependem do tempo e que o ruído de medida são independentes, o que permite escrever $R_k = R = \sigma_n^2 I$, onde R é matriz de covariância do ruído, σ^2 é a varância e I é a matriz identidade.

IV.6.1 Obtenção de R

Conforme dito anteriormente, R é associada à resolução das câmeras empregadas no experimento com imagens reais.

Sejam $1/h_x$ e $1/h_y$ eventos independentes consideradas nos intervalos $[0; T]$ e $[0; S]$. Os comprimentos desses intervalos são respectivamente T e S .

$1/h_x$: $[0; T]$ evento na direção horizontal;

$1/h_y$: $[0; S]$ evento na direção vertical.

As probabilidades de ocorrência dos eventos independentes nos intervalos considerados, ou seja, nas direções x e y , são $p(1/hx) = 1/T$ e $p(1/hy) = 1/S$. As médias desses evento μ_x e μ_y são determinadas como segue:

$$\begin{aligned}\mu_x &= \frac{T}{2} \\ \mu_y &= \frac{S}{2}\end{aligned}\tag{4.24}$$

Então a variância do evento $1/h_x (1/h_y)$ é $\sigma^2 = E\{x^2\} - E^2\{x\}$, com $x = 1/hx$, $y = 1/hy$ e necessariamente,

$$\begin{aligned}E^2\{x\} &= \mu_x^2 \\ E^2\{y\} &= \mu_y^2\end{aligned}\tag{4.25}$$

e

$$\begin{aligned}E\{x^2\} &= \int_0^T x^2 p_x dx = \int_0^T x^2 \frac{1}{T} dx = \frac{T^2}{3} \\ E\{y^2\} &= \int_0^T y^2 p_y dy = \int_0^T y^2 \frac{1}{S} dy = \frac{S^2}{3}\end{aligned}\tag{4.26}$$

Então, de (4.24), (4.25) e (4.26), tem-se:

$$\sigma_x^2 = E_x\{x^2\} - E^2\{x\} = \frac{T^2}{3} - \frac{T^2}{4} = \frac{T^2}{12}\tag{4.27}$$

$$\sigma_y^2 = E_y\{y^2\} - E^2\{y\} = \frac{S^2}{3} - \frac{S^2}{4} = \frac{S^2}{12}\tag{4.28}$$

As expressões (4.27) e (4.28) permitem o cálculo das variâncias das variáveis x e y respectivamente.

Finalmente, pode-se escrever:

$$R = \sigma_n^2 I, \quad \text{com } n \text{ assumindo o valor } x \text{ ou } y \quad (4.29)$$

R é a matriz de covariância de todo o ruído associado ao processo de aquisição, processamento e digitalização de imagens.

Até este ponto, foi resolvido o problema da determinação da matriz R , elemento da equação (3.58), modelo representativo da covariância do ruído de medição $v_n(t)$. O próximo passo é estabelecer os procedimentos relativos à determinação da matriz Q (3.57).

IV.7 Seleção da matriz de covariância de ruído de estado Q

Não é uma tarefa fácil selecionar valores da matriz Q apropriados que proporcionem um bom desempenho do filtro estimador de estados (IEKF) [4]. A má seleção de Q pode ocasionar ou a divergência do filtro ou a convergência para valores errados. A natureza não-linear do problema abordado é uma forte aliada dessas ocorrências. A convergência do filtro é assegurada quando a matriz de ganho K , (3.91), decai em magnitude com taxa apropriada (verificada experimentalmente neste trabalho), ou seja, taxa de decaimento não extremamente alta tampouco bastante baixa [39].

A matriz de covariância do erro de estimativas P , (3.92), é função da covariância do erro inicial de estimativa $P(0)$, da matriz do erro de medição R , pois P depende de K e R é parâmetro de K , e da matriz de ruído do modelo do processo Q . Já que $P(0)$ e R são fixas no tempo, conforme definido anteriormente, a única forma de controlar a taxa de decaimento da matriz K é justamente através da matriz Q . Conforme a expressão (3.94), o produto matricial $G(t)Q(t)G^T(t)$ será utilizado como elemento controlador de decaimento da matriz K . Esse produto matricial pode ser representado pela relação Iq , [4], onde I é a matriz identidade e q escalares a serem determinados segundo [4], baixos valores de q levam a um rápido decaimento de K , o que faz com que em (3.90), expressão do cálculo das estimativas, as estimativas sejam determinadas

independentemente do processo de medição. Por outro lado, elevados valores de q provocam o efeito inverso, ou seja, as estimativas passam a ser bastantes sensíveis aos valores medidos. Portanto, para evitar esses dois extremos indesejáveis, é interessante selecionar valores de q não muito baixos tampouco bastante elevados.

De acordo com o que foi dito no parágrafo anterior a determinação da matriz de ruído de estado Q é um procedimento eminentemente experimental (tentativa e erro), pois depende da seleção dos melhores valores de q que favorecem a uma taxa apropriada de decaimento da matriz de ganho K . Esses valores serão determinados experimentalmente (a partir de simulações sucessivas até que sejam encontrados os valores apropriados), apresentados no Capítulo V, bem com o desempenho do filtro estimador associado a cada um deles.

A presente seção discutiu as fontes de ruído no experimento com imagens reais. Felizmente, a forma como o assunto foi tratado é um procedimento bastante comum em trabalhos científicos que dependem de sistemas desse tipo ([4] e [1], por exemplo).

Tendo analisado os modelos de movimentos, os modelos de observação de pontos de canto, os modelos de projeção de imagens e o filtro de Kalman, análises feitas no Capítulo III e estudado as técnicas, algoritmos de segmentação, verificação de correspondência e as fontes de ruído associadas ao processo de digitalização de imagens (estudos feitos no presente capítulo), já é possível apresentar a versão final do algoritmo da estimação paramétrica. Esta versão leva em consideração os modelos analisados no Capítulo III, bem como todos os estudos, técnicas, e modelagens examinados neste capítulo. O seguinte algoritmo resume o processo.

<início>

quadro = 0;

Do while (quadro < NUMERO MÁXIMO DE QUADROS)

{

Aquisição de imagens via SADI;

Verificação da correspondência;

Detecção de cantos;

Modelagem de projeção;

Modelagem de observação de ponto ruidosos;
Obtenção da estrutura do objeto
Modelagem das equação de estados;
estimação de estados via filtro de Kalman;
}
Incrementa quadro ;
<fim>

A Figura 4.3 mostra a configuração do SADI utilizada nos experimentos com imagens reais

Fig. 4.3 - Configuração do SADI.

A Figura 4.4 mostra na forma de diagrama os passos do algoritmo de estimação paramétrica estudado neste trabalho.

Fig. 4.4 Procedimento de estimação paramétrica.

CAPÍTULO V

RESULTADOS DE EXPERIMENTOS

V.1 Comentários iniciais

Neste capítulo são analisados os resultados obtidos com o procedimento de estimação paramétrica apresentado, objeto deste trabalho. Os algoritmos elaborados foram implementados em linguagem de programação “C” (DOS/UNIX). O desempenho da estimação, foi avaliado com seqüências de imagens reais. No experimento atributos das imagens adquiridas para de um objeto de teste (a ser caracterizado nas seções seguintes), foram obtidos para cada cena filmada (quadros).

Também neste capítulo, são avaliados os desempenhos dos métodos de detecção de cantos, das diferentes técnicas de resolução do problema da correspondência.

V.2 Avaliação dos métodos de detecção considerados

Nesta seção são apresentados os resultados experimentais obtidos com os dois métodos de detecção estudados no Capítulo IV, ou seja, o detector fundamentado nos gradientes horizontal e vertical e o detector ótimo de cantos. Para que fosse possível comparar os desempenhos de cada um dos métodos de detecção, foram realizados testes com imagens simuladas. Nesses testes, os esforços computacionais, medido pelo tempo de CPU, consumidos na execução dos respectivos algoritmos, requerido pelos dois métodos, foram avaliados. Também foi objeto desses testes, a determinação da condição do casamento existente entre as imagens obtidas após a aplicação dos detectores e as imagens originais (casamento).

Após selecionar, dentre os métodos de detecção considerados, aquele que apresentou o menor tempo de execução, ou seja, o menor esforço computacional (menor tempo de CPU), foram feitos os testes com os MOPR(s).

A tabela 5.2 mostra os tempos gastos na execução dos algoritmos de detecção de cada um dos métodos estudados no Capítulo IV, bem como a condição do casamento na localização do canto na imagem original. As imagens de teste usadas pelos dois métodos são as mostradas na tabela 5.1. Uma delas, a imagem triangular, é a mesma utilizada em [10]. A outra foi considerada para cobrir as diferentes possibilidades de ocorrências das faces do objeto de teste no plano-imagem das câmeras quando as imagens são projetadas. O computador utilizado nos experimentos foi um PC-386 (8Mbytes-40MHz).

Tabela 5.1 - Imagens sintéticas usadas no experimento.

Banco de imagens sintéticas	
Triângulo isósceles	Retângulo

Essas imagens foram geradas por um programa simples, implementado em linguagem de “C”, aproveitando-se da biblioteca gráfica do compilador associado a essa linguagem. Cada uma delas apresenta um fundo de cor preta. O processo de detecção foi realizado usando-se os dois métodos de detecção e as referidas imagens foram contaminadas com ruído de intensidade variável, gerando-se dessa forma uma imagem ruidosa, para representar o processo de digitalização de imagens. O gerador de ruído também é um programa implementado em “C” e consiste em variar, de maneira

aleatória, os níveis de cinza das imagens em relação ao fundo. O nível de cinza máximo nos experimentos foi 255 e o mínimo, 0. Assim, as imagens podiam apresentar-se fortemente ruidosas (nível de cinza = 255) ou com pouca influência de ruído de digitalização, o que corresponde ao nível de cinza igual a zero.

A avaliação da qualidade da detecção apresentada por cada um dos métodos, ou seja, a localização do canto na imagem original, foi feita, verificando-se o casamento entre as imagens obtidas após a aplicação dos detectores, que somente contêm os cantos detectados, e a imagem original. Ocorrerá o casamento (detecção considerada muito boa) se os cantos detectados coincidirem com as posições das interseções das bordas, atributos considerados na imagem original.

Na aplicação do detector fundamentado nos gradientes vertical e horizontal foram utilizadas em todas as imagens as máscaras apresentadas na Seção IV.3, isto é, as expressões (4.1) e (4.2) e todos os princípios de implementação daquele detector lá examinado. Já na aplicação do detector ótimo de cantos, foram utilizadas diferentes máscaras de detecção, da seguinte forma:

1. a máscara W_1 em aplicação com imagem triangular;
2. a máscara W_2 em aplicação com imagem retangular.

No preenchimento das máscaras W_1 e W_2 foram utilizadas as equações (4.15) e (4.16). W_1 foi preenchida usando-se (4.15) na região triangular e (4.16) na região não triangular. A máscara W_2 foi preenchida de maneira similar e a mesma é considerada como a máscara ótima de detecção de cantos em imagens cujas bordas são ortogonais [10].

$$w_1 = \begin{bmatrix} -3 & -5 & -5 & -3 & 0 & -3 & -5 & -5 & -3 \\ -5 & -9 & -9 & -5 & 0 & -5 & -9 & -9 & -5 \\ -5 & -9 & -9 & -5 & 0 & -5 & -9 & 8 & 5 \\ -3 & -5 & -5 & -3 & 0 & 6 & 9 & 9 & 6 \\ 0 & 0 & 0 & 0 & 0 & 6 & 10 & 10 & 6 \\ -3 & -5 & -5 & -3 & 0 & 6 & 9 & 9 & 6 \\ -5 & -9 & -9 & -5 & 0 & -5 & -9 & 8 & 5 \\ -5 & -9 & -9 & -5 & 0 & -5 & -9 & -9 & -5 \\ -3 & -5 & -5 & -3 & 0 & -3 & -5 & -5 & -3 \end{bmatrix} \quad (5.1)$$

$$w_2 = \begin{bmatrix} -6 & -11 & -11 & -6 & 0 & 4 & 7 & 7 & 4 \\ -11 & -18 & -18 & -11 & 0 & 8 & 13 & 13 & 8 \\ -11 & -18 & -18 & -11 & 0 & 10 & 17 & 17 & 10 \\ -6 & -11 & -11 & -6 & 0 & 12 & 19 & 19 & 12 \\ 0 & 0 & 0 & 0 & 0 & 12 & 20 & 20 & 12 \\ -6 & -11 & -11 & -6 & 0 & -6 & -11 & -11 & -6 \\ -11 & -18 & -18 & -11 & 0 & -11 & -18 & -18 & -11 \\ -11 & -18 & -18 & -11 & 0 & -11 & -18 & -18 & -11 \\ -6 & -11 & -11 & -6 & 0 & -6 & -11 & -11 & -6 \end{bmatrix} \quad (5.2)$$

A tabela 5.2 mostra os tempos de execução dos dois métodos de detecção de cantos abordados, bem como a sobreposição (casamento) do canto detectado na imagem original. As Figuras 5.1 a 5.2 mostram as imagens resultantes após o processo de detecção de cantos.

Tabela 5.2 - Tempo de execução dos algoritmos de detecção de cantos e do casamento.

Detector	Imagem triangular		Imagem retangular	
	Tempo de CPU	casamento	Tempo de CPU	casamento
Método do Gradiente	360s	aceitável	480s	Muito bom
Detector ótimo	600s	Muito bom	900s	Muito bom

(1)

(2)

Fig. 5.1 - Detecção de cantos pelo Método do Gradiente. Cantos detectados: (1) imagem triangular e (2) imagem retangular . No processo de detecção os limiares utilizados em cada caso foram respectivamente 0.002 e 0.07, valores menores que estes não melhoraram o desempenho do detector.

(1)

(2)

Fig. 5.2 - Imagens resultantes do detector ótimo de canto. (1) os cantos detectados na imagem triangular e (2) na imagem retangular.

Analisando a tabela 5.2 mostrada anteriormente, percebe-se que a segunda abordagem, ou seja, o detector ótimo de cantos é o que requer o maior esforço computacional qualquer que seja a imagem de teste considerado. O tempo necessário à execução de seu algoritmo é sempre maior que o do primeiro detector (método do gradiente). Este fato está diretamente relacionado com a dimensão da máscara de

convolução empregada na implementação do algoritmo ($M_{9 \times 9}$). Entretanto, os elevados tempos de processamento são acompanhados de uma maior precisão de localização do canto na imagem original em relação ao método do gradiente. Foi observado, também, que o casamento relativo ao detector ótimo tinha maior qualidade que o da outra abordagem, independente da imagem testada. Todavia, o detector ótimo de canto requer um prévio conhecimento da forma geométrica do objeto, para com esta definir o tipo de máscara a ser utilizada ou configurar máscaras que são um conjunto de sub máscaras, quando a geometria do objeto é desconhecida [10]. Essas sub máscaras devem representar todas as possibilidades de ocorrência de cantos na imagem do objeto, o que acaba por aumentar o tempo de execução desse detector.

O método dos gradientes é mais rápido que a outra abordagem, mas é sensível à forma geométrica da imagem considerada. A sua rapidez é justificada pelo fato de seu algoritmo manipular máscaras de convolução que individualmente apresentam dimensão ($SV_{3 \times 3}$ e $SH_{3 \times 3}$) bem menor do que aquela do detector ótimo ($M_{9 \times 9}$). No que diz respeito a sua sensibilidade, foi observado nas simulações que o presente detector tem dificuldade em resolver (detectar) imagens onde ocorrem bordas formando ângulos menores do que 90° . Por exemplo, na imagem triangular utilizada nas simulações, todos os ângulos medem 60° ; observou-se que o casamento relativo à detecção do canto localizado nos vértices desses ângulos não apresenta um bom casamento (vide Figuras 5.1 e 5.2). Em alguns casos (15° e 10° , por exemplo) verificou-se inconsistência ou ausência de detecção.

Conclui-se, portanto que: o detector ótimo de cantos poderá ser sempre utilizado no processo de segmentação de imagens utilizado na estimação, contanto que sejam configuradas sub máscaras capazes de prever todas as ocorrências de cantos na imagem em questão. Contudo, o custo da sua utilização como modelo de detecção de cantos é, conforme comentários feitos anteriormente, a lenta execução da estimação. O detector fundamentado no método do gradiente poderá ser utilizado no processo de segmentação de imagens com eficiência similar a do detector ótimo de canto, caso nas imagens consideradas não ocorram bordas formando ângulos menores do que 90° , conforme comentários feitos. Em se verificando esta última condição, o procedimento

de estimação com esse detector de cantos, será bem mais rápida que a obtida a partir da utilização do detector ótimo.

V.3 Avaliação das técnicas de verificação da correspondência entre imagens

São examinados, nesta seção, os resultados obtidos com as duas técnicas de resolução do problema da correspondência entre imagens estéreo abordadas neste trabalho. Uma delas é fundamentada na manipulação de todas as informações existentes nas imagens originais, enquanto que a outra manipula atributos (pontos salientes, como bordas ou cantos) existentes nas imagens, requerendo, portanto, que as mesmas sejam previamente submetidas ao processo de localização dos atributos. Os modelos necessários à implementação dos algoritmos de uma e outra técnica foram analisados no Capítulo IV. Foram feitas simulações com um par de imagens reais, adquirida por câmeras distintas (imagens estéreo). As imagens em questão são de dois quadros aleatórios (referentes à mesma cena) da seqüência de quadros obtida após a filmagem dos movimentos realizados pelo objeto de teste. Um quadro, mostrado a seguir, foi utilizado nas simulações com as duas técnicas, possibilitando comparar os resultados obtidos.

Fig. 5.3 - Quadros utilizados no estudo da correspondência. Atributos não pertencentes ao objeto de teste são os reflexos da fonte de luz na mesa de teste. O objeto de teste encontra-se mais próximo da câmera da esquerda.

O desempenho das referidas técnicas será analisado em termos do tempo de processamento. Conforme se percebe observando a Figura 5.3, as imagens são monótonas, ou seja, é sempre o objeto de teste (cor branca) mergulhado num cenário de cor preta. Primeiramente são analisados os resultados com a técnica pertencente à classe Estereoscopia Baseada em Área e em seguida será feita a análise dos resultados obtidos com o técnica que pertence à classe Estereoscopia Baseada em Atributos

V.3.1 Avaliação da técnica pertencente à classe Estereoscopia Baseada em Área

A concepção do modelo matemático desta técnica já foi examinada no Capítulo IV (item IV.5.1). Dadas as características das imagens utilizadas nos experimentos, isto é, a não sensibilidade ao efeito de contraste e o fato de serem filmadas por câmeras idênticas, o parâmetro D , elemento de fundamental importância na determinação da similaridade, foi calculado empregando-se a expressão (4.18). As constantes m e n assumiram, respectivamente os valores 500 e 200 que foram suficientes para definir

regiões nas imagens com alta probabilidade de ocorrência de atributos do objeto de teste. O limite máximo de limiar L foi mantido fixo durante a avaliação das similaridades. utilizou-se $L = \sigma_n^2 = 1.38 \times 10^{-6}$, porque a presença de ruído relativo ao processo de aquisição e processamento das imagens poderia interferir na resolução do problema da correspondência. Os pontos-imagem localizados nas faces comuns aos campos de visão das câmeras, ou seja, faces superior e frontal apresentaram-se correspondentes. Os pontos-imagem pertencentes às faces filmadas exclusivamente pelas câmeras, por estarem numa região sombreada em relação às outras duas faces, foram considerados não correspondentes, porque nessa região os valores de D superaram a magnitude de L .

V.3.2 Avaliação da técnica pertencente à classe Estereoscopia Baseada em atributos

A modelagem relativa a esta técnica encontra-se examinada no Capítulo IV (item IV.5.2). De acordo com o que lá foi delineado, existe a necessidade da prévia localização dos atributos das imagens antes da aplicação da presente técnica. Nas simulações, tal localização foi feita, empregando-se o detector de cantos fundamentado no cálculo dos gradientes (vertical e horizontal, por ser o de execução mais rápida e devido o fato da não ocorrência de ângulos menores do que 90°). Nas imagens resultantes após a localização dos atributos foram definidas as coordenadas do ponto de partida para os cálculos das similaridades (D). A exemplo do procedimento adotado no item V.3.1, a expressão (4.18) foi empregada no cálculo do parâmetro D . O resultado das simulações revelou que dos sete cantos detectados seis deles eram correspondentes. O canto não correspondente localiza-se na face sombreada do objeto de teste e é a resultante da interseção das bordas horizontais inferiores com a borda vertical posterior.

Nesse experimento foi constatado que a opção por utilizar uma ou outra técnica está diretamente relacionada com a velocidade de execução dos algoritmos utilizados. O tempo de processamento requerido pela técnica pertencente à classe Estereoscopia

Baseada em atributos é inferior ao necessário ao processamento da outra técnica (pertencente à classe Estereoscopia Baseada em Área).

V.4 Resultados experimentais com um modelo de observação de pontos ruidosos

Nesta seção são apresentados os resultados obtidos com os modelos de observação de pontos ruidosos (MOPR's) empregados neste trabalho. Na obtenção dos resultados foram utilizadas algumas imagens reais do objeto de teste, já submetidas ao algoritmo de verificação da correspondência, objetivando testar as potencialidades dos MOPR's na recuperação de informações tridimensionais do referido objeto.

A tabela 5.3 mostra um quadro comparativo entre os diferentes MOPR's mencionados anteriormente. Para se elaborar esse quadro, dez realizações foram feitas, isto é, dez imagens do objeto de teste foram adquirida para que fosse feito um teste de desempenho dos dois MOPR's estudados no Capítulo III. O método de detecção de cantos utilizado nos testes foi o que necessitou menor esforço computacional, ou seja, o detector fundamentado nos gradientes (vertical e horizontal) abordado no Capítulo IV.

Os testes foram realizados nas seguintes condições:

- 1 - Um ponto do objeto de teste, cuja posição na mesa e teste era conhecida, foi considerado;
- 2 - O objeto foi filmado pelas câmeras, obtendo as projeções do referido ponto nos planos imagem das câmeras;
- 3 - Os pontos projetados, ou seja, os pares (X, Y) foram usados nos MOPR's;
- 4 - Os MOPR's calculam a posição daquele ponto a partir desses pares;
- 5 - Os erros relativos médios, ee_{mx} , ee_{my} e ee_{mz} , existentes entre a posição calculada pelos MOPR's e a posição real do objeto (medida na mesa de testes) são determinados;
- 6 - Um outro ponto foi considerado e os passos de 1 a 5, repetidos.

Tabela 5.3 - Comparação entre os desempenhos dos diferentes MOPR's empregados na recuperação de informações tridimensionais.

MOPR(s)	ee_{mx}	ee_{my}
A partir do cálculo dos mínimos quadrados, item III.2.2	2.33%	1.23%
A partir da determinação do "range", item III.2.2.1	1.65%	0.34%

Observando a tabela 5.3, nota-se que o segundo MOPR's é o mais confiável na recuperação de pontos tridimensionais a partir de informações bidimensionais, resultantes do processo de projeção de imagens de uma cena filmada. Isto se deve ao fato da segunda abordagem, mas não a primeira, considerar a correspondência entre o ponto localizado no plano imagem de uma câmera e o correspondente ponto conjugado, localizado no plano imagem da outra câmera.

V.5 Resultados experimentais com seqüência de imagens reais

Considera-se nesta seção o problema de se adquirir uma seqüência de imagens, constituída de 50 quadros, relativos ao movimento de um objeto de teste. Nessa seqüência constam imagens que contêm as diferentes posições ocupadas pelo objeto em uma mesa de testes. Tais imagens são adquiridas e gravadas em arquivos pelo sistema de aquisição e digitalização de imagens (SADI), descrito no Capítulo IV. Após isto, elas são utilizadas, como se fossem adquiridas a cada instante em que o objeto de teste realiza um movimento na mesa, em algoritmos processadores de imagens, detectores de cantos e verificadores de correspondência para que certas informações contidas nelas, cantos, sejam utilizadas no estimador de parâmetros IEKF. Pretende-se com esse procedimento, analisar a possibilidade de aquisição de imagens em tempo real, objetivando a estimação de parâmetros que descrevam o movimento do objeto, quais sejam, posição, velocidade linear, aceleração e velocidade angular.

V.5.1 Movimento do objeto de teste no espaço tridimensional

Empregou-se como objeto de teste, um cubo de cor branca com comprimento 4.1cm, altura 2.3cm e largura 4.7cm. Os movimentos desse cubo foram realizados em uma mesa plana de cor preta, que por contrastar com a cor do objeto, favoreceu bastante a qualidade das imagens adquiridas. O centro de massa do cubo seguiu uma trajetória definida pela tabela 5.4 mostrada a seguir, obtida utilizando-se a equação (3.69). Inicialmente, o objeto movimentou-se para longe da origem do sistema referencial inercial (movimento progressivo). Em seguida, o objeto mudou o sentido do movimento, passando a aproxima-se da origem do referido sistema (movimento retrógrado).

A descrição física do local onde os movimentos foram realizados e as condições de movimentação do objeto de teste são as seguintes:

- Um espaço tridimensional referencial inercial, representado fisicamente por um sistema de eixos coordenados Ox , Oy , e Oz , montados em uma mesa de testes;
- Os eixos Ox e Oy possuem 250.0cm de comprimento, enquanto que o eixo Oz mede 150cm;
- O centro de massa do objeto de teste realizou um movimento retilíneo e uniformemente variante no tempo (MRUV), partindo do repouso, isto é, $v = (0,00; 0,00; 0,00)$;
- Pontos não pertencentes à linha que passa pelo centro de massa do objeto de teste realizaram movimentos curvilíneos uniformes (MCU) em, relação a essa linha;
- Iluminação da cena (área de movimentação do objeto) com luz difusa;
- A aceleração do centro de massa do objeto de teste apresentou o valor constante de 0.2cm/min^2 ;
- A velocidade angular do objeto de teste $10^\circ/\text{min}$ (em relação ao centro de massa);

- O objeto de teste apresentou movimento acelerado e progressivo até 32min de movimentação, retardado e progressivo entre 34 e 60min e entre 62 e 98min o movimento foi retardado e retrógrado.

Tabela 5.4 - Trajetória seguida pelo centro de massa do objeto teste.

Tempo (min)	Posições ocupadas pelo centro de massa do objeto de teste (cm)	Tipo do movimento
0	(10,00; 10,00; 1,15) $\pm$ 0,05	acelerado/progressivo
2	(10,40; 10,40; 1,15) $\pm$ 0,05	acelerado/progressivo
4	(11,60; 11,60; 1,15) $\pm$ 0,05	acelerado/progressivo
6	(13,60; 13,60; 1,15) $\pm$ 0,05	acelerado/progressivo
8	(16,40; 16,40; 1,15) $\pm$ 0,05	acelerado/progressivo
10	(20,00; 20,00; 1,15) $\pm$ 0,05	acelerado/progressivo
12	(24,40; 24,40; 1,15) $\pm$ 0,05	acelerado/progressivo
14	(29,60; 29,60; 1,15) $\pm$ 0,05	acelerado/progressivo
16	(35,60; 35,60; 1,15) $\pm$ 0,05	acelerado/progressivo
18	(42,40; 42,40; 1,15) $\pm$ 0,05	acelerado/progressivo
20	(50,00; 50,00; 1,15) $\pm$ 0,05	acelerado/progressivo
22	(58,40; 58,40; 1,15) $\pm$ 0,05	acelerado/progressivo
24	(67,60; 67,60; 1,15) $\pm$ 0,05	acelerado/progressivo
26	(77,60; 77,60; 1,15) $\pm$ 0,05	acelerado/progressivo
28	(88,40; 88,40; 1,15) $\pm$ 0,05	acelerado/progressivo
30	(100,00; 100,00; 1,15) $\pm$ 0,05	acelerado/progressivo
32	(112,40; 112,40; 1,15) $\pm$ 0,05	acelerado/progressivo
34	(125,60; 125,60; 1,15) $\pm$ 0,05	retardado/progressivo
36	(138,00; 138,00; 1,15) $\pm$ 0,05	retardado/progressivo
38	(149,60; 149,60; 1,15) $\pm$ 0,05	retardado/progressivo
40	(160,40; 160,40; 1,15) $\pm$ 0,05	retardado/progressivo
42	(170,40; 170,40; 1,15) $\pm$ 0,05	retardado/progressivo
44	(179,60; 179,60; 1,15) $\pm$ 0,05	retardado/progressivo
46	(188,00; 188,00; 1,15) $\pm$ 0,05	retardado/progressivo
48	(195,60; 195,60; 1,15) $\pm$ 0,05	retardado/progressivo
50	(202,40; 202,40; 1,15) $\pm$ 0,05	retardado/progressivo
52	(208,40; 208,40; 1,15) $\pm$ 0,05	retardado/progressivo
54	(213,60; 213,60; 1,15) $\pm$ 0,05	retardado/progressivo
56	(218,00; 218,00; 1,15) $\pm$ 0,05	retardado/progressivo
58	(221,60; 221,60; 1,15) $\pm$ 0,05	retardado/progressivo
60	(224,40; 224,40; 1,15) $\pm$ 0,05	retardado/progressivo
62	(226,40; 226,40; 1,15) $\pm$ 0,05	mudança de sentido
64	(226,00; 226,00; 1,15) $\pm$ 0,05	acelerado/retrógrado
66	(224,80; 224,80; 1,15) $\pm$ 0,05	acelerado/retrógrado
68	(222,80; 222,80; 1,15) $\pm$ 0,05	acelerado/retrógrado
70	(220,00; 220,00; 1,15) $\pm$ 0,05	acelerado/retrógrado
72	(216,40; 216,40; 1,15) $\pm$ 0,05	acelerado/retrógrado
74	(212,00; 212,00; 1,15) $\pm$ 0,05	acelerado/retrógrado
76	(206,80; 206,80; 1,15) $\pm$ 0,05	acelerado/retrógrado

78	(200,80; 200,80; 1,15) $\pm$ 0,05	acelerado/retrógrado
80	(194,00; 194,00; 1,15) $\pm$ 0,05	acelerado/retrógrado
82	(186,40; 186,40; 1,15) $\pm$ 0,05	acelerado/retrógrado
84	(178,00; 178,00; 1,15) $\pm$ 0,05	acelerado/retrógrado
86	(168,80; 168,80; 1,15) $\pm$ 0,05	acelerado/retrógrado
88	(158,80; 158,80; 1,15) $\pm$ 0,05	acelerado/retrógrado
90	(148,00; 148,00; 1,15) $\pm$ 0,05	acelerado/retrógrado
92	(136,40; 136,40; 1,15) $\pm$ 0,05	acelerado/retrógrado
94	(124,00; 124,00; 1,15) $\pm$ 0,05	acelerado/retrógrado
96	110,80; 110,80; 1,15) $\pm$ 0,05	acelerado/retrógrado
98	(96,80; 96,80; 1,15) $\pm$ 0,05	acelerado/retrógrado

V.5.2 Processo de aquisição de sequência de imagens estéreo

A cada nova posição ocupada pelo objeto de teste, o mesmo era filmado pelo SADI. O SADI, cuja configuração está mostrada na Figura (4.1), é constituído de:

- "Duas câmeras" Sony HVM-302 fixas. Uma na posição (250,00, 0,00, 30,00) $\pm$ 0,05 cm e a outra na posição (-50,00; 0,00; 30,00) $\pm$ 0,05cm.
- Uma placa digitalizadora de imagem - DT2851- com capacidade de armazenamento de imagens de 480 x 512 pixel.
- Um micro-computador do tipo 386 equipado com *software* que permite a configuração de algoritmos de manipulação de arquivos de imagens.

As imagens adquiridas pelo SADI foram armazenadas em 100 arquivos (um arquivo por câmera) de imagens digitalizadas (480 x 512 pixels). As imagens foram adquiridas para cada posição ocupada pelo objeto, de acordo com a tabela 5.4.

V.5.3 Processo de calibração das câmeras utilizadas no SADI

O processo de calibração das câmeras consiste em transformar as coordenadas dos pontos imagem, medidas em pixels, em coordenadas em unidade métrica (cm). Isto é feito a partir do cálculo dos fatores de escala da câmera nas direções x_c e y_c do plano

imagem. Denotando-se esses fatores por h_x e h_y respectivamente, os mesmos podem ser determinados por um método simples que utiliza a proporcionalidade entre as dimensões do objeto e a sua respectiva imagem formada na câmera. O diagrama da Figura 5.4 mostra a representação geométrica da câmera utilizada nos experimentos.

O diagrama mostrado na Figura 5.5 auxilia na compreensão do método empregado. Este diagrama é obtido após simplificações feitas naquela representação geométrica. Estas simplificações produzem uma situação muito particular de calibração da câmera, mas que, na situação explorada neste trabalho, em que o objeto segue uma trajetória que cruza diagonalmente o campo visual da câmera, passando pelo eixo ótico desta, são admissíveis, ou seja, não há a necessidade de se determinar todos os parâmetros mostrados na Figura 5.4 para o processo de calibração da câmera utilizada no SADI. Determina-se, somente, os mostrados na Figura 5.5.

Fig. 5.4 - Diagrama do modelo generalizado da câmera.

A partir da Figura 5.4 e de comentários feitos em [33], tem-se:

$$x = \frac{(X - X_o) \cos(\theta) + (Y - Y_o) \sin(\theta) - r_1}{-(X - X_o) \sin \theta \sin \alpha + (Y - Y_o) \cos \theta \sin \alpha - (Z - Z_o) \cos \alpha + r_3 + f} \quad (5.1)$$

$$y = \frac{(X - X_o) \sin \theta \sin \alpha + (Y - Y_o) \cos \theta \cos \alpha - (Z - Z_o) \sin \alpha - r_2}{-(X - X_o) \sin \theta \sin \alpha + (Y - Y_o) \cos \theta \sin \alpha - (Z - Z_o) \cos \alpha + r_3 + f} \quad (5.2)$$

onde θ e α são, respectivamente, o ângulo entre x e X , e y e Y .

As expressões (5.1) e (5.2) permitem calcular as projeções no plano imagem da câmera, isto é, definem o modelo generalizado de obtenção das projeções (x, y) , feitas nas câmera, a partir dos pontos (x, y, z) localizados no objeto .

As simplificações a serem feitas no presente modelo são as seguintes:

- $X_o = Y_o = Z_o$;
- $r_1 = r_2 = r_3 = 0$;
- $\theta = \alpha = 0^\circ$.

Feito isto, tem-se:

$$x = \frac{f X}{f - Z} \quad (5.3)$$

$$y = \frac{f Y}{f - Z} \quad (5.4)$$

As equações (5.3) e (5.4) permitem o seguinte modelo básico de processamento de imagem:

Fig. 5.5 - Diagrama do fator de escala.

A partir do diagrama, tem-se:

$$h_i = \frac{a_i}{A} \frac{d}{f}$$

(5.5)

onde,

- A: Lado do objeto de teste;
- d: Distância de A até o referencial da câmera. O valor de d é equivalente a $(f - Z)$ em (5.3) e (5.4);
- a: Medida em pixel do lado do objeto de teste na imagem;
- f: distância focal da câmera;
- i: eixo do plano de imagem das câmeras.

A calibração foi realizada com os seguintes dados:

$$A = 5\text{cm};$$

$$d = 37,6\text{cm};$$

$$f_c = 1,1\text{cm};$$

$$a_x = 40 \text{ pixels};$$

$$a_y = 39 \text{ pixels}$$

Substituindo esses valores na expressão (5.5) e a imagem do objeto adquirida pela placa DT2851, obtém-se os seguintes valores de fatores de escala:

$$h_x = 335,64 \text{ pixels/cm} \quad (5.6)$$

$$h_y = 267,08 \text{ pixels/cm} \quad (5.7)$$

V.5.4 Estimação paramétrica

Será testada nesta seção a possibilidade de se usar os modelos (e respectivos algoritmos) analisados nos Capítulos III e IV, na descrição do movimento do objeto de teste. O objeto realizou movimentos na mesa de teste segundo a equação (3.69), as posições por ele ocupadas ao longo da trajetória, cujos valores encontram-se reproduzidos na tabela 5.3. As imagens das cenas foram adquiridas e armazenadas em arquivos. Os arquivos foram, então, manipulados para que os cantos das imagens fossem extraídos das mesmas e utilizados no processo de estimação dos parâmetros que descrevem o movimento do objeto de teste e os que os permitem definir a forma a sua geométrica.

A validação da procedimento foi feita da seguinte forma:

- O objeto ocupou posições sucessivas na mesa de teste, conforme descrito anteriormente;
- Cada nova posição (x, y, z) teve o respectivo valor anotado e a correspondente imagem do objeto, adquirida.

- Após adquiridas todas as imagens e anotadas as respectivas posições ocupadas pelo objeto, o processo de medição estava concluído;
- O conjunto de imagens gerado no item anterior foi, então, usado para que os parâmetros fossem identificados;
- A validação é finalmente feita, avaliando-se o desvio existente entre as estimativas obtidas e os valores observados experimentalmente.

Antes de apresentar os resultados experimentais obtidos, os valores de alguns parâmetros, fundamentais à implementação do filtro de Kalman ((3.90) a (3.94)), necessitam ser determinados. Esses parâmetros são a matriz de covariância associada ao ruído de medida R , a matriz de covariância do ruído de estado Q e a matriz de covariância inicial do erro inicial de estimativa P . Assim, qualquer condição de simulação, de caráter investigatório, que interfira no desempenho desses modelos e/ou técnicas (aumentando-o ou diminuindo-o), necessariamente refletirá no desempenho do procedimento como um todo. As influências de alguns modelos e/ou técnicas já foram analisadas nas seções precedentes (Seções V.2 a V.4); a subseção seguinte exibirá alguns resultados relativos à sensibilidade das estimativas feitas com o filtro de Kalman perante à manipulação de certos parâmetros.

Seleção das matrizes R , Q e P

A matriz R foi determinada a partir das expressões da modelagem da condição de iluminação do experimento destinado à aquisição de imagens reais examinada na Seção IV.6.1, isto é, expressões (4.27) e (4.28), bem como da calibração da câmera analisada na subseção (V.5.3), ou seja, (5.6) e (5.7). A matriz Q , conforme subseção IV.6.2, foi obtida a partir dos valores de q e a matriz P assumiu o valor $1000000I_n$, onde I_n é a matriz identidade de ordem 16, pois, esse é o número de estados a serem estimados. Então, tem-se:

i) Seleção de R

$h_x = 335,64$ pixels/cm, então $1/h_x$,(resolução horizontal da câmera) tem valor $2,98 \times 10^{-3}$

$h_y = 267,08$ pixels/cm, então $1/h_y$,(resolução horizontal da câmera) tem valor $3,74 \times 10^{-3}$

De (4.27) e (4.28) tem-se ,

$$\sigma_x^2 = \frac{(2.98 \times 10^{-3})^2}{12} = 7.4 \times 10^{-7}$$

$$\sigma_y^2 = \frac{(3.74 \times 10^{-3})^2}{12} = 1.17 \times 10^{-6}$$

Em que T e S são substituídos respectivamente por $1/h_x$ e $1/h_y$ nas expressões (4.27) e (4.28).

As covariâncias determinadas acima estão associadas respectivamente às resoluções nas direções horizontal e vertical. A covariância resultante que considera as influências simultâneas dessa resoluções é obtida calculando-se a norma 2 das duas contribuições. Então, tem-se:

$$\sigma_n^2 = \sqrt{(\sigma_x^2)^2 + (\sigma_y^2)^2} \quad \text{ou}$$

$$\sigma_n^2 = 1.38 \times 10^{-6}$$

O que permite escrever que $R = 1.38 \times 10^{-6} I_0$.

ii) Seleção de Q

De acordo com a subseção IV.7, $Q = I_{16}$. Nas simulações, diversos valores de q , pertencentes ao intervalo $[10^{-9}; 10^{-1}]$, serão utilizados e a correspondente consequência na qualidade das estimativas, analisadas.

iii) Seleção de P

Segundo [41], P deve ser uma matriz positiva definida durante todo o processo de filtragem (obtenção das estimativas) feito pelo IEKF. Um procedimento que garante essa condição é iniciar o processo de filtragem com valores elevados de P . Então, nas simulações o valor inicial de P adotado é $P(0) = 10^6 I_{16}$.

Estimativas obtidas

A seguir, a influência da escolha da matriz Q no desempenho do Filtro de Kalman será analisada. Também serão mostrados resultados de simulação nas quais se examina a qualidade das estimativas em função das condições iniciais do Filtro de Kalman, uma vez definido o valor ideal da matriz Q . Os resultados serão exibidos através de gráficos das três componentes (x , y e z) dos estados estimados comparados com os gráficos desses mesmos estados (componentes x , y e z) admitidos como reais (observados experimentalmente), isto é, serão exibidos gráficos, contendo os desvio (erros) das estimativas feitas em relação aos valores observados a cada movimento do objeto de teste.

As Figuras 5.6 a 5.9 mostram os gráfico dos estados observados experimentalmente (medidos na mesa de testes) do centro de massa do objeto de teste, descrevendo uma trajetória que cruza diagonalmente o eixo ótico da câmera. Os gráficos são da posição, da velocidade e da aceleração do centro de massa (os três primeiros). Mostra-se também a seguir, o gráfico da velocidade angular dos cantos do objeto de teste em relação ao centro de rotação OB , paralelo ao eixo z (o quarto gráfico).

Fig. 5.6 – Posição do centro de massa do objeto. Os componentes x e y são idênticas, pois o objeto seguiu uma trajetória diagonal, cruzando o eixo ótico da câmera. O componente z é igual a 1,15cm, pois não houve movimento na direção oz .

Fig. 5.7 – Velocidade do centro de massa do objeto. Os componentes v_x e v_y são idênticos, pois o objeto seguiu uma trajetória diagonal, cruzando o eixo ótico da câmera. O componente v_z é igual a 0, pois não houve movimento na direção oz .

Fig. 5.8 – Aceleração do centro de massa do objeto. Os componentes a_x , a_y e a_z são idênticos e iguais a $0,2\text{cm/min}^2$.

Fig. 5.9 – Velocidade angular do objeto de teste. Os componentes w_x , w_y e w_z são idênticos e iguais a $10,0\text{graus/min}$.

Influências dos valores de q nas estimativas

Os melhores valores de q serão determinados observando-se os desvios de estimativas em relação às medidas determinadas experimentalmente durante a

processo de determinação dos valores de q pertencentes ao intervalo $[10^{-10}; 10^{-1}]$. São considerados quatro casos

1º caso, $q = 10^{-9}$.

Condições iniciais: $r_e(0) = (13,80; 14,50; 1,15)$, $v_e(0) = (2,8; 1,8; 1,5)$, $a_e(0) = (0,2; 0,2; 0,2)$ e $w_e = (12,00; 13,00; 10,00)$.

Fig. 5.10 - Desvio entre as estimativas dos componentes da posição do centro de massa e os valores desses componentes observados experimentalmente.

Fig. 5.11 - Desvio entre as estimativas dos componentes da velocidade do centro de massa e os valores desses componentes observados experimentalmente.

Observando os gráficos das Figuras 5.10 a 5.13, percebe-se que as estimativas dos estados feitas pelo filtro de Kalman começam a convergir para os respectivos estados observados experimentalmente a partir de 20 minutos de simulação, pois, é aproximadamente nesse instante que os desvios de estimativas, isto é, a diferença entre

os valores estimados e os valores observados experimentalmente aproximam-se de um valor mínimo. Esse valor mínimo é aproximadamente 0,25 para as estimativas dos componentes do vetor posição do centro de massa; 0,2 para as estimativas do vetor velocidade do centro de massa e 0,35 para as estimativas feitas tanto dos componentes do vetor aceleração do centro de massa quanto as estimativas da velocidade angular dos pontos característicos do objeto de teste. Entretanto, a partir dos instantes 85 ou 92min os desvios de estimativas que até então estabilizavam nos respectivos valores mencionados anteriormente, passam a aumentar de magnitude, afastando-se dos mesmos, em outras palavras, o filtro de Kalman tende a divergir a partir daqueles instantes, confirmando tal tendência nos instantes seguintes. Este resultado mostra que $q = 10^{-9}$, valor selecionado para a definição da matriz Q , conforme comentários feitos na Seção IV.6.2, é um valor que não garante um bom desempenho do filtro de Kalman. Portanto, dentro do intervalo de pesquisa dos valores de q ($[10^{-10}; 10^{-1}]$) a serem utilizados na implementação do filtro de Kalman (IEKF), os resultados das simulações revelam que valores de q localizados no extremo inferior daquele intervalo, ou seja, valores de q pequenos interferem de maneira desfavorável no desempenho do IEKF, fazendo com que o mesmo, ou não obtenha qualquer estimativa dos estados em questão - o IEKF diverge - ou encontre estimativas bastantes diferentes dos valores esperados, isto é, o IEKF converge para valores errados. Esta última condição não foi observada nas simulações, mas a primeira sim, a partir dos instantes referidos anteriormente. São mostrados a seguir os resultados obtidos com q assumindo valores localizados no outro extremo do intervalo de pesquisa.

2º caso, $q = 10^{-1}$

Condições iniciais: $r_e(0) = (10,00; 15,50; 3,55)$, $v_e(0) = (-2,5; -2,5; -2,5)$, $a_e(0) = (2,0; 2,0; 2,0)$ e $w_e = (12,00; 11,00; 12,00)$.

Fig. 5.14 - Desvio entre as estimativas dos componentes da posição do centro de massa e os valores desses componentes observados experimentalmente.

Fig. 5.15 - Desvio entre as estimativas dos componentes da velocidade linear do centro de massa e os valores desses componentes observados experimentalmente.

Fig. 5.16 - Desvio entre as estimativas dos componentes da aceleração linear do centro de massa e os valores desses componentes observados experimentalmente.

Fig. 5.17 - Desvio entre as estimativas dos componentes da velocidade angular de três pontos característicos do objeto de teste (*ec1*, *ec2*, *ec3*) e os valores desses componentes observados experimentalmente.

Observando os gráficos das Figuras 5.14 a 5.17, percebe-se que a partir de 40 minutos de simulação, as estimativas dos estados feitas pelo filtro de Kalman começam a convergir para o respectivo estado observado no processo (movimento do objeto de

teste), pois verifica-se nesse instante que os desvios de estimativas tendem a um valor mínimo. O valor mínimo em questão é de aproximadamente 0.2 cm para as estimativas feitas dos componentes do vetor posição do centro de massa; 0.45 cm/min para as estimativas obtidas do vetor velocidade do centro de massa e 0.25 cm/min² (graus/min) para as estimativas feitas tanto das componentes do vetor aceleração do centro de massa quanto das estimativas da velocidade angular dos cantos do objeto de teste. Entretanto, a exemplo do que acontece no caso anterior, partir do instante 80 min os desvios de estimativas deixam de estabilizar nos respectivos valores mencionados anteriormente e passam afastar-se daqueles valores, porém seguindo um sentido oposto ao verificado no caso anterior. É desnecessário dizer que, também aqui, o filtro de Kalman diverge a partir daquele instante. Este resultado revela que $q = 10^{-1}$ é um valor que, a exemplo de $q = 10^{-9}$, não garante um bom desempenho do filtro de Kalman. Portanto, dentro do intervalo de pesquisa dos valores de $q = ([10^{-10}; 10^{-1}])$, valores pequenos ou elevados desse parâmetro interferem de maneira desfavorável no desempenho do IEKF, fazendo com que o mesmo deixe, a partir de um certo instante, de obter qualquer estimativa dos estados em questão, ou seja, é muito provável que o filtro venha a divergir após algum tempo de simulação se aqueles valores de q forem considerados na definição da matriz Q . É mostrado a seguir os resultados obtidos com o valor de q que proporcionou o melhor desempenho do IEKF.

3º caso, $q = 10^{-4}$, mas o processo de estimação dos parâmetros é iniciado com conhecimento das condições iniciais dos valores iniciais dos estados relativos ao movimento do objeto de teste.

Fig. 5.18 - Desvio entre as estimativas dos componentes da posição do centro de massa e os valores desses componentes observados experimentalmente.

Fig. 5.19 - Desvio entre as estimativas dos componentes da velocidade linear do centro de massa e os valores desses componentes observados experimentalmente.

Fig. 5.20 - Desvio entre as estimativas dos componentes da aceleração linear do centro de massa e os valores desses componentes observados experimentalmente.

Fig. 5.21 - Desvio entre as estimativas dos componentes da velocidade angular de três pontos característicos do objeto de teste (*ec1*, *ec2*, *ec3*) e os valores desses componentes observados experimentalmente.

O conhecimento das condições iniciais é assegurado quando $xe(t)$ é inicializado ($xe(0)$) com valores bastantes próximos dos estados iniciais observados experimentalmente. Sendo assim, para o estudo do presente caso, nas simulações realizadas, considerou-se $xe(0) = (8.50, 8.50, 1.10, 0.01, 0.01, 0.00, 0.15, 0.15, 0.15, 1.00, 0.00, 0.00, 0.00, 9.00, 9.00, 9.00)$ que, por ser bastante próximo das condições iniciais verificada no processo, apresentando desvios relativos abaixo de 20%, permite ao IEKF iniciar a obtenção das estimativas dos estados com um relativo conhecimento dos estados que realmente se observa no processo.

Conforme pode ser observado nos gráficos das Figuras 5.19 a 5.22, iniciando-se a estimação com o $xe(0)$ em questão, os desvios de estimativas iniciais já são bem menores que os observados no caso anterior. Observa-se também que 40 minutos (20 quadros) de simulação ou menos é suficiente para que as estimativas feitas pelo IEKF comecem a convergir para os estados observados experimentalmente, pois, a partir desse instante o desvio absoluto entre as estimativas dos componentes da posição do centro de massa converge para $\approx 0,09$, permanecendo na vizinhança desse valor até o final da simulação ($t = 98\text{min}$). Também nas mesmas condições as estimativas dos componentes das velocidades do centro de massa, as estimativas dos componentes da aceleração do centro de massa e os valores estimados das velocidades angulares começam a convergir para os respectivos valores medidos experimentalmente, apresentando, até o final das simulações, respectivamente os seguintes desvios absolutos $\pm 0.06, 0.07$, e 0.05 (aproximadamente). Este resultado mostra que iniciar a filtragem (estimação) conhecendo-se as condições iniciais do processo, isto é, inicializar os estados a serem estimados com valores suficientemente próximos dos valores iniciais efetivamente verificados no processo, tem o efeito de antecipar a convergência das estimativas para os estados esperados (observados no processo durante a simulação), ou seja, essa condição faz com que a estimação de parâmetros feita pelo algoritmo de estimação paramétrica seja mais rápida. Quanto às magnitudes dos erros absolutos verificadas após o IEKF convergir para os estados esperados, essas foram as menores encontrados nas simulações (tanto neste quanto no caso anterior), utilizando o melhor valor de q , pesquisado no intervalo $[10^{-9}; 10^{-1}]$.

4º caso, $q = 10^{-4}$, mas o processo de estimação dos parâmetros é iniciado sem qualquer conhecimento das condições iniciais dos valores iniciais dos estados relativos ao movimento do objeto de teste.

Condições iniciais: $r_e(0) = (10,00; 10,00; 1,15)$, $v_e(0) = (0,0; 0,0; 0,0)$, $a_e(0) = (0,2; 0,2; 0,2)$ e $w_e = (10,00; 10,00; 10,00)$.

Fig. 5.22 - Desvio entre as estimativas dos componentes da posição do centro de massa e os valores desses componentes observados experimentalmente.

Fig. 5.23 - Desvio entre as estimativas dos componentes da velocidade linear do centro de massa e os valores desses componentes observados experimentalmente.

Fig. 5.24 - Desvio entre as estimativas dos componentes da aceleração linear do centro de massa e os valores desses componentes observados experimentalmente.

Fig. 5.25 Desvio entre as estimativas dos componentes da velocidade angular de três pontos característicos do objeto de teste (ec1, ec2, ec3) e os valores desses componentes observados experimentalmente.

Neste caso, a estimação do estado é iniciada sem qualquer informação acerca das condições iniciais do processo considerado, pois $x_e(t)$, vetor dos estados

estimados pelo IEKF, é inicializado ($x_e(0)$) com valores bastante distantes dos estados iniciais, observados experimentalmente, ou seja, $x_e(0) = (p_m(0), v_m(0), a_m(0), q_m(0), w_m(0))$, onde o índice m diz respeito aos valores medidos (conhecidos) dos estados a serem estimados. O vetor inicial dos componentes dos quatêrnios ($q_m(0)$) devem ser selecionados contanto que a norma do vetor seja unitária, isto é, $\|q_m(0)\| = 1$. Esse vetor não é um estado relacionado com o movimento. Ele é somente um estado utilizado na propagação do movimento de rotação do objeto de teste (vide capítulo III). Sendo assim, para o estudo do presente caso, nas simulações realizadas, considerou-se $x_e(0) = (3.50, 3.50, 2.10, -1.5, -1.8, 1.8, 0.35, 0.35, 0.37, 1.00, 0.00, 0.00, 0.00, 4.00, 4.00, 4.00)$ que, de acordo com as considerações feitas na seção V.5.1, apresenta desvios relativos acima de 50%.

Conforme pode ser observado nos gráficos das figuras 5.23 a 5.26, ao iniciar a estimação dos estados com o $x_e(0)$ mencionado anteriormente, os desvios de estimativas iniciais são elevados, mas, a partir de 70 minutos (35 quadros) de simulação, esses mesmos desvios apresentam uma sensível redução nas respectivas magnitudes (quando comparadas com os valores iniciais), significando que as estimativas feitas pelo IEKF começam a convergir para os estados observados experimentalmente, pois, a partir desse instante, o desvio absoluto entre as estimativas das componentes da posição do centro de massa converge para ≈ 0.08 , estabilizando nesse valor até o final da simulação ($t=98\text{min}$). Nas mesmas condições, as estimativas das componentes das velocidades do centro de massa, as estimativas das componentes da aceleração do centro de massa e os valores estimados das velocidades angulares começam a convergir para os respectivos valores observados experimentalmente, apresentando, até o final das simulações, respectivamente, os seguintes desvios absolutos ± 0.05 , 0.09 , e 0.05 (aproximadamente). Este resultado mostra que ao iniciar a filtragem (estimação) sem conhecer as condições iniciais do processo, isto é, inicializar os estados a serem estimados com valores bastante distintos dos valores iniciais verificados no processo, retarda a convergência das estimativas obtidas pelo IEKF para os estados observados no processo, ou seja, a

condição examinada neste caso faz com que a obtenção das estimativas de parâmetros feitas pelo algoritmo de estimação paramétrica seja mais lenta.

Os resultados dos experimentos com imagens reais examinados nos casos 1 a 4, utilizando o algoritmo de estimação paramétrica implementado neste trabalho, revelam a real possibilidade de utilização do mesmo no rastreamento de um objeto que se movimenta em uma mesa de testes, filmado por um par de câmeras. Para que seja possível o rastreamento, estimativas dos estados relacionados com o modelo de movimentação do objeto são obtidas a cada mudança de posição verificada nas cenas filmadas. O componente do algoritmo encarregado de obter as referidas estimativas é o IEKF. Sendo assim, a possibilidade ou não de se rastrear o objeto depende do tempo gasto por esse componente para convergir as estimativas feitas para os valores realmente observados no processo. Também a magnitude dos desvios verificados após o IEKF convergir é uma condição bastante relevante no desempenho de algoritmos rastreadores de objetos. Os casos examinados nessa seção revelam algumas condições que interferem no desempenho do IEKF ou na possibilidade de um rastreamento bem sucedido. Essas condições são as seguintes:

- Se a estimação dos parâmetros é iniciada sem qualquer conhecimentos das condições iniciais verificadas no processo, o IEKF necessita de mais tempo (mais quadros) para convergir para o estado observado no processo. Este fato faz com que o algoritmo implementado seja lento no rastreamento, podendo, contudo ser aplicado, para rastrear objetos que se movimentam com baixa velocidade. Nas simulações conduzidas para o estudo dos casos 1 a 4 a velocidade máxima atingida pelo objeto de teste foi 7.7cm/min, ou seja, o objeto movimentou-se lentamente.

- Se os desvios entre as estimativas e o estado realmente observado no processo for de magnitude elevada, o desempenho do algoritmo implementado no rastreamento do objeto também é afetado negativamente, pois nessa situação haverá uma baixa precisão, por exemplo, na localização instantânea do objeto na mesa de teste ou, se essa estimativa for utilizada em dispositivos capazes de decidir alguma

ação de controle a partir do movimento do objeto, falha desse dispositivos. Os desvios verificados nas simulações são elevados (em média 0,06), mas permitem o algoritmo examinado rastrear com uma boa precisão o movimento do objeto de teste dada a suas dimensões e as imposições (limitações) feitas na Seção V.5.1.

CAPÍTULO VI

CONCLUSÃO

Neste trabalho considera-se o problema do emprego dos recursos utilizados em visão computacional, particularmente em visão estereoscópica, na estimação dos parâmetros que descrevem os movimentos de translação e rotação de um objeto. O objetivo é obter estimativas desses parâmetros a partir de um modelo de espaço de estado representativo do movimento de um corpo rígido no espaço 3-D e de uma seqüência de imagens estéreo. Das imagens são retirados atributos que são utilizadas no processo de filtragem feita pelo IEKF. Para tanto, foi desenvolvido um procedimento que reúne recursos de visão computacional, modelo de espaço estado, manipulação de seqüência de imagens e filtragem de parâmetros. As principais contribuições foram:

- a) utilização, no procedimento adotado, de métodos não considerados em [1], isto é, método de correspondência entre imagens estéreo, detetor ótimo de cantos e triangularização de imagens utilizando o conceito de *range* ([8], [10], [40] e [41]);
- b) melhoria da qualidade das estimativas dos parâmetros, por uso de um método de obtenção de pontos ruidosos (Seção III 2.1) que, por estabelecer uma correspondência entre o ponto localizado no plano imagem de uma câmera e o correspondente ponto conjugado, localizado no plano imagem da outra câmera, recupera pontos tridimensionais com mais eficiência que o método sugerido em [1].
- c) estudo de dois métodos de detecção de cantos, analisando a influência de cada um deles no procedimento adotado;

- d) proposta de um ambiente para aplicação em tempo real de um dispositivo de visão estéreo;

O desempenho do algoritmo desenvolvido em situações realísticas foi verificado nos testes com seqüência de imagens estéreo reais, conforme descrição e apresentação de resultados feitos na Seção V.5. Nesses testes, algumas restrições quanto ao movimento do objeto de teste foram feitas (vide item V.5.1). Estas restrições ajudaram a especificar o problema abordado, limitando as magnitudes dos valores esperados das estimativas a serem obtidas. Na obtenção da trajetória do objeto de teste utilizada nos experimento com imagens reais (tabela 5.3), vários ensaios foram realizados (levando em consideração as dimensões da mesa de testes, os comprimentos dos suportes das câmeras empregados e a especificação feita para o problema abordado) até que fosse obtida a trajetória apresentada na tabela 5.3. Poderiam ser utilizadas outras trajetórias, porém os ensaios revelaram que em algumas delas, os valores iniciais da velocidade linear do centro de massa fariam com que o objeto de teste ocupasse posições superiores as dimensões da mesa de teste.

Os testes com seqüência de imagens reais revelaram que a natureza recursiva dos métodos empregados na configuração do algoritmo desenvolvido faz com que o mesmo seja bastante flexível no que diz respeito à manipulação dos dados necessários a sua execução. Esse fato foi facilmente confirmado neste trabalho, pois nas simulações com imagens reais (ou mesmo com imagens sintéticas), era, por exemplo, livre a escolha da quantidade dos pontos característicos ou arbitrária a quantidade de imagens a serem manipuladas pelos algoritmos implementados. Tal flexibilidade representa, portanto, uma vantagem do algoritmo desenvolvido sobre outras abordagens que não utilizam métodos recursivos.

Contudo, esta mesma natureza recursivas do procedimento desenvolvido, particularmente do filtro de Kalman empregado na configuração da mesma (IEKF), pode levar as estimativas obtidas pelo IEKF a divergir, estimando estados bastante diferentes dos realmente observados, por má escolhas das condições iniciais dos estados a serem estimados. Neste trabalho tal ocorrência pode ser verificada nos

estudos dos casos 1 a 4, quando se pesquisava os valores do parâmetro q capazes de influenciar a qualidade das estimativas obtidas.

Na definição da matriz Q , utilizada na implementação do IEKF, utilizou-se como procedimento a pesquisa do parâmetro q mencionado anteriormente num intervalo previamente definido (item V.5.4). Existem, contudo, outros procedimentos envolvendo técnicas de estimação adaptativa [41] em os elementos da matriz Q são estimados à proporção que as estimativas dos estados vão sendo obtidas. Técnicas com essa característica não foram abordadas neste trabalho, mas é possível configurar o procedimento desenvolvido com técnicas desse tipo.

Na implementação do procedimento desenvolvido foram considerados outros métodos, examinados nos Capítulos III e IV, não explorados em [1], e a maneira como os mesmos afetam o desempenho do procedimento desenvolvido. Foram feitas várias simulações usando tais métodos no Capítulo V e os resultados destas simulações mostram as vantagens e as desvantagens de usar um ou outro método na configuração do procedimento desenvolvido.

Conforme ressaltado no Capítulo IV, pretendia-se neste trabalho verificar a possibilidade de aplicação desse procedimento em robótica, isto é, aqui seria examinada a possibilidade de usá-la como dispositivo de visão de um manipulador robótico. Esta aplicação, entretanto, necessita de interações com trabalhos dedicados ao assunto. Contudo, em função dos resultados apresentados no Capítulo V, não será uma tarefa difícil usar o procedimento desenvolvido como um dispositivo de visão de um manipulador robótico. Essa afirmativa encontra forte apoio nas condições em que o procedimento foi validado (confronto das estimativas com os estados observados no processo), pois, durante a validação foram consideradas situações realísticas tais como: modelo generalizado da câmera e respectivas simplificações, iluminação das imagens adquiridas e estimação dos parâmetros do objeto feita em cada cena filmada.

As estimativas obtidas desses parâmetros apresentaram desvio aceitáveis em relação aos estados realmente observados durante a utilização da trajetória de teste. Se esses parâmetros forem devidamente comunicados ao manipulador robótico procedimento desenvolvido, então, será possível a execução de tarefas, com o conhecimento dos movimento realizado pelo objeto de teste, pois nesta situação, o

manipulador estará atuando a partir da realimentação constante das diferentes posições ocupadas pelo mesmo, das velocidades desenvolvidas e da aceleração.

As sugestões para trabalhos futuros podem ser divididas em duas classes. Em uma delas, sugere-se inovações no processo de obtenção das estimativas dos parâmetros que descrevem o movimento do objeto. Na outra, faz-se sugestões acerca da atitude das câmeras durante o movimento do objeto.

Como exemplo de sugestões tem-se:

- Utilização de um filtro de Kalman adaptativo no processo de estimação paramétrica. Esse filtro deverá estimar parâmetros, bem como obter as estimativas dos ruídos de processo e de medição.. Nessa abordagem, as equações matemáticas do filtro são sistematicamente concebidas, à luz da teoria dos filtros adaptativos, no sentido de fazer com que os resíduos de predições, avaliados em tempo real, determinem o nível de ruído apropriado que proporcione um desempenho aceitável do filtro, isto é, durante a obtenção das estimativas, os ruídos são adaptados aos resíduos de predições;
- Utilização de câmeras móveis no rastreamento do objeto. O movimento das câmeras poderá ser definido a partir do treinamento em tempo real de uma rede neural, realimentada com as atuais posições do objeto.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] G. S. Young, and R. Chellappa, "3-D Motion Estimation Using a Sequence of Noisy Stereo Images: Model, Estimation and Uniqueness Result". *IEEE Transactions on Pattern Analysis and Machine*, vol. 12 no. 2, pp. 735-759, 1990.
- [2] R.Y. Tsai, "Determining the Instant. Direction of Motion from Optical Flow Generated by a Curvilinearly Moving Observer". In *Proc. IEEE Conf. Pattern Recog. Image Proc.* Dallas, pp. 109-114, 1981.
- [3] W.B. Tompso .and V. A. Brerins, "Dynamic Occlusion Analysis in Optical Flow Fields". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. vol. - 7, no. 4, pp. 135-141, 1985.
- [4] T. J. Broida, "Recursive 3-D Motion Estimation from a Monocular Image Sequence". *IEEE Transactions on Aerospace and Electronic Systems* vol. 26, no. 4, pp. 639 to 750, 1990.
- [5] J.Q. Fang and T. S. Hung, "Solving Three-Dimensional Small Rotation Equation". *IEEE Conf. Computer. Vision. on Pattern Recog.* Washington, DC, pp. 253-258, 1983.
- [6] S. Karl, "Estimating 3-D Movement of a Rigid Object: Experimental Result" *Int. Joint Conf. Artif. Intell.* West Germany, pp. 1035-1037, 1983.
- [7] A. R. Brus and B. K P. Horn, *Passive Navigation*. Mass. Inst. Technology, A.I. Memo 662, 1981.
- [8] Y. C Shah, R Chapman and R.B. Mahani, "A New Technique to Range Information from Stereo Images". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, no 7, pp. 768 - 773. 1989.
- [9] A. Kitchen and A. Rosenfeld, "Edge Evaluation using Local Coherence". *IEEE Transactions Systems Man. Cybernetic*, vol. 11, no 9, pp. 597 - 605, 1981.

- [10] K. Rangarajan, M. Shah and D. V. Bracke, "Optimal Corner Detector". *Computer Vision, Graphics and Image Processing*, pp. 230 - 245, 1989.
- [11] S. Ullman, "Analysis of Visual Motion by Biological and Computer Systems Computer". *MIT Artificial Intelligence Laboratory*, vol. 14 pp. 57-59, 1981.
- [12] H.H. Nagel, "On the Derivation of 3-D Rigid Point Configurations from Image Sequences". *IEEE Conf. Pattern Recong. Image Proc.*, Dallas, TX, pp. 103-108. 1981.
- [13] H.H. Nagel and B. Newman, "On 3-D Reconstruction from Two Perspective Views". *In Proc. Int. Joint Conf. Artif. Intell*, Vancouver, Canada, pp. 661-663, 1981.
- [14] K. Prazdny, "Motion and Relative Depth Map from Optical Flow". *Biol. Cybern*, vol. 36, pp. 87-102, 1980.
- [15] H. C. Longuet - Higgins, "A Computer Algorithmics for Reconstructing a Scene from Two Projections". *Nature*, vol. 293, pp.133-135, 1981.
- [16] R.Y. Tsai and T.S. Huang, "Uniqueness and Estimation of 3-D Motion Parameter of Rigid Objects with Curved Surfaces". *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 6, pp. 13-27, 1984.
- [17] J.M. Roach and J.K. Aggarwal, "Determining the Movement Objects for a Sequence of Images". *IEEE Trans. Pattern Analysis. Machine Intell.*, vol. 2, pp. 554 - 562, 1980.
- [18] H.C. Longuet and K. Prazdny, "The Interpretation of a Moving Retinal Image". *In Proc. Roy. Soc. London*, vol. 208, pp. 385-397, 1980.
- [19] S. Ullman, "The Interpretation of Visual Motion". *Cambridge. MA: MIT Press*, pp. 110-118, 1979.
- [20] L. Hungen, "Fast Hough Transformation": *A Hierarchical Approach. Computer Vision, Graphics and Image Processing*, pp. 139-161, 1986.
- [21] R.Y. Tsai, and T.S. Huang, "Estimating Three-Dimensional Motion Parameter of Planar Patch, II: Singular Value Decomposition". *IEEE Transactions on Acoustic Speech and Signal Processing*, vol. 30, no 4. 1982.
- [22] J.Q. Fang, and T. S. Hung "Some Experiments on Estimating The 3-D Motion Parameters of Rigid Body from Two Consecutive Image-Frames". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 545-554, 1984.

- [23] R.Y. Tsai and T.S. Hunag, "Estimating Three-Dimensional Motion Parameter of Planar Patch". *IEEE Transactions on Acoustic Speech and Signal Processing*. pp. 554 – 562, 1982.
- [24] O. Gibson, J.J. Gibson, O. W., Smith, and H. Flock "Motion parallax as a Determinant of Perceived Depth" *J. Exo. Psych*, vol. 45. 1959.
- [25] A. Azarbayejani, B. Horowitz and Pentland, "Recursive estimation of structure and motion using relative orientation constraints". *Proc. CVPR*, NEW YORK, 1993.
- [26] A. Black, M. Taylor and A Cox, *Grasping visual symmetry*. Proc. of the ICCV, 1993.
- [27] W. Daywansa and X. Wang, "Real Time Visual Tracking for Surveillance and Path Planing". *Proc. of the ECCV*, 1992.
- [28] M. Swain and M. Stricker, "Promising directions in active vision". *NSF Active Vision Workshop*, University of Chicago. 1991
- [29] S. Stefano and G. Picci, "Motion Estimation via Dynamic Vision". Procedure of the 33rd Conference on Decision and Control. Lake Buena Vista, pp. 3253 - 3258. FL 1994
- [30] C.W. Chang, and A. S. Morse, "Control of a Rigid Robot Using an Uncalibrated Stereo Vision System". *Center for Computational Vision and control Department of Electrical Engineering*, Yale University. AACC. 1997.
- [31] C.W. Chang, A. S. Morse and G. D. Hanger, "A Calibration-Free, Self-Adjusting Stereo Visual Control System". *In Proc. of the 13th World Congress, International Federation of Automatic Control*, vol. A, pp. 33-348. IFAC., 1996.
- [32] A. S. Morse, "Control Using Logic-Based Switching". In A. Isidori, editor, *trends in Control: A European Perspective*, pp. 69-113. Springer Verlag, 1995.
- [33] T.D. Williams, "Computer Interpretation of a Dynamic Image from a Moving Vehicle". PhD. Dissertation, Dep. Comput. Inform. SCI, Univ. Mass, pp. 81-82, 1981.
- [34] S. Karl, "Estimating 3-D Movement of a Rigid Object: Experimental Result". In *Proc. Int. Joint Conf. Artif. Intell*, Karlsruhe, West Germany, pp. 1105 - 1127, 1985.
- [35] H. H. Nagel, "On the Derivation of 3-D Rigid Point Configurations from Image Sequences". In *Proc. IEEE Conf. Pattern Recong. Image Proc.* Dallas, TX, pp.103-108, 1981.

- [36] K. S. Fu and R. C. Gonzales, *Robotics: Control, Sensing, Vision and Intelligence*. MacGraw - Hill Book Company, 1987.
- [37] B. Friland, "Analysis Strapdown Navigation Using Quaternios". *IEEE Transactions on Aerospace and Electronic Systems*, vol. 14, no 5, pp 164-168, 1978.
- [38] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York Academic Press, 1970.
- [39] C. T. Shen, *Linear System Theory and Design*. State University of New York at Stony Brook. 1970.
- [40] J. F. Canny, "Finding Edges and lines in images - Master's thesis", *MIT Press*, pp.41-50. 1983.
- [41] A. Zuniga and R. Haralick, "Corner detection using the fact model". *IEEE CPVR*. pp. 30 - 37, 1983.
- [42] H. Drechler and H. Nagel, "Voluntaric Model and 3-D Trajectory of a Moving Car Derived from Monocular Tv - Frame of Street Scene". *IJCAI*, pp. 692-697, 1981.
- [43] A. Singh and M. Shneier, *Computer Vision, Graphics and Image Processing*. pp. 54-69, 1990.
- [44] D. H. Ballard & C. M. Brown, *Computer vision*. Prentice Hall, 1982.
- [45] K. Rangarajan and M. Shah, "Optmal Corner Detector". *Computer Vision, Graphics and Image Processing*, pp. 230-245, 1989.
- [46] W. E. L. Grimsom, *Object Recognition by Computer: Role of Geometric Constrains*. MIT Press 1990.

APÊNDICE – MATRIZES DO SISTEMA DINÂMICO LINEARIZADO

Este apêndice mostra os valores numéricos de algumas variáveis utilizadas no processo de obtenção das estimativas dos estados (parâmetros que descrevem o movimento do objeto) do sistema de equações dinâmicas (3.59). Aqui são mostrados os valores numéricos assumidos pelas matrizes de linearização do sistema dinâmico (3.56), ou seja, são apresentadas as matrizes $H(x_{op})$ e $F(x_{op})$. Denomina-se x_{op} o ponto de linearização. Nas simulações, x_{op} assumiu o valor do estado estimado no instante imediatamente anterior à realização de uma nova medição no processo. Também são apresentados as matrizes de covariâncias dos ruídos de medição, ruído de estado e a matriz de covariância do erro de estimação (R , Q , e P respectivamente). Como a equação matricial obtida após a linearização de (3.59) tem dimensão 16×16 e nas simulações feitas neste trabalho foram consumidos 50 quadros, implicando na avaliação dessa equação 50 vezes, somente serão apresentados os valores numéricos correspondentes à primeira estimativa dos estados.

Na apresentação dos valores numéricos, serão considerados os resultados obtidos no 3º caso analisado no item V.5.4.

1 – Valores numéricos das variáveis correspondentes ao primeiro quadro

- estimativa inicial dos estados:

$$x_e(0) = (8,5; 8,5; 1,10; 0,01; 0,01; 0,00; 0,15; 0,15; 0,15; 1,00; 0,00; 0,00; 0,00; 9,00; 9,00; 9,00)$$

- valores iniciais dos estados observados experimentalmente:

$$x_0(0) = (10,00; 10,00; 1,15; 0,00; 0,00; 0,00; 0,20; 0,20; 0,20; 1,00; 0,00; 0,00; 0,00; 10,00; 10,00; 10,00)$$

- matriz de covariância do ruído de medição (expressão (4.30)):

$$R_{9 \times 9} = \begin{bmatrix} 1,38 \times 10^{-6} I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & 1,38 \times 10^{-6} I_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & 1,38 \times 10^{-6} I_{3 \times 3} \end{bmatrix}$$

- matriz de covariância do processo de ruído de estado:

$$P(0) = 10^6 I_{16 \times 16}$$

- matriz $H(x_{op})$ (expressão (3.88)), linearização de $h[x(t_i); t]$ do sistema (3.56) no ponto de operação ($x_{op} = x_e(0)$):

$$H(x_{op})_{9 \times 16} = \begin{bmatrix} I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & \frac{\partial h_{111}}{\partial q(t)} & 0_{3 \times 3} \\ I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & \frac{\partial h_{112}}{\partial q(t)} & 0_{3 \times 3} \\ I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & \frac{\partial h_{113}}{\partial q(t)} & 0_{3 \times 3} \end{bmatrix}$$

onde,

$$\frac{\partial h_{111}}{\partial q(t)} = \begin{bmatrix} 0,0 & 0,0 & 7,8 & 0,0 \\ 0,0 & 0,0 & 0,0 & 7,8 \\ -7,8 & 0,0 & 0,0 & 0,0 \end{bmatrix}$$

$$\frac{\partial h_{112}}{\partial q(t)} = \begin{bmatrix} 4,2 & -4,2 & 0,0 & 0,0 \\ 0,0 & 4,2 & 0,0 & 0,0 \\ 0,0 & 0,0 & 4,2 & 4,2 \end{bmatrix}$$

$$\frac{\partial h_{113}}{\partial q(t)} = \begin{bmatrix} 0,0 & 8,0 & 0,0 & 0,0 \\ 0,0 & 0,0 & 0,0 & 0,0 \\ 0,0 & 0,0 & 0,0 & -8,0 \end{bmatrix}$$

- matriz $F(x_{op})$ (expressão (3.61)), linearização de $f[x(t); t]$ do sistema (3.56) no ponto de operação:

[illegible]